Learning from others through testimony and statistics

By

Jane Hu

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Psychology
in the
Graduate Division
Of the
University of California, Berkeley

Committee in charge:
Professor Fei Xu, chair
Professor Alison Gopnik
Professor Thomas Griffiths
Professor Terry Regier

Fall 2014
Abstract

Learning from others through testimony and statistics

by

Jane Hu

Doctor of Philosophy in Psychology

University of California, Berkeley

Fei Xu, Chair

In learning about the world, children have at least two types of information available to them: information they learn from their personal experiences, and information they receive from others. This dissertation examines how children use both of these types of information to make inferences about others. In chapter 1, I discuss the role of this work in the context of previous developmental psychology research. In chapter 2, I present a set of empirical studies in which children inferred an agent’s graded preferences from observing his choice actions. In chapter 3, I present a second set of empirical studies, in which children’s endorsement of majority testimony differs based on domain type and amount of personal experience available. In chapter 4, I present a third set of empirical studies in which children assessed informants’ knowledge sources and chose options endorsed by informants who received their knowledge from more reliable sources of knowledge. In chapter 5, I discuss the implications of this work and suggest future directions. Overall, the empirical work included in this dissertation suggests preschoolers’ inferences abilities are more sophisticated than previously demonstrated: they can use contextual information in conjunction with statistical information, and can use that to make inferences about others’ mental states and knowledge.
Acknowledgments

I’d first like to thank my advisor, Fei Xu, for her guidance, support, and patience over the last five years. I would also like to thank the other members of my committee – Tom Griffiths, Alison Gopnik, and Terry Regier – for their many helpful comments and stimulating discussions. I am extremely grateful to have had such a fantastic collection of minds to help steer my dissertation work.

I would also like to thank my collaborators – Daphna Buchsbaum, Chris Lucas, Andrew Whalen, and Falk Lieder – for their dedication and invaluable input, as well as my labmates and fellow CPD students Stephanie Denison, Brian Waismeyer, Zi Lin Sim, Azzurra Ruggieri, Justine Hoch, Maddy Hanson, Caren Walker, Audun Dahl, Sophie Bridgers, and Mikel Delgado for all their great ideas, honest opinions, constructive criticism, and moral support. You have made even my windowless office in Tolman a fun place to be.

Additionally, I’d like to thank all my RAs for their hard work, willingness to learn, and tireless enthusiasm: Chelsea Tuomi, Eri Nagase, Stephanie Jones, Dea Garic, Anamita Guha, Nina Alpert, and Gabi Espinoza.

This dissertation would not be possible without the support of my family. Mom and Dad: you’ve always given me 100% of your hearts; this is a debt I will never be able to repay. All those times I complained about MST or TIP: you were right. I quite literally wouldn’t be here without you, and I would not have gotten so far without your love and support. Mom, thank you for always making time for me, whether that was for after school activities or chats on a Saturday morning. Dad, thank you for instilling in me your insatiable curiosity, and for always pushing me to be the best I can. Thanks to you, there will be two Dr. Hus in our family. Patty and Gigi: I picked the right family. Thank you for your constant love and support, for always being willing to lend a sympathetic ear, and for giving us an excuse to drive to Nepenthe. Last but certainly not least: Nate, I could not have done this without you. Your love and support have meant everything to me, and I still agree to disagree about the usefulness of solving computational-level problems.

Thanks also to the many participants and their parents for volunteering their time and brainpower. Lastly, I’d like to thank the National Science Foundation Graduate Research Fellowship for supporting my research.
For Nate and Maeby
# Table of Contents

**List of figures**........................................................................................................................................v

**List of tables**.........................................................................................................................................vi

## 1. Introduction...........................................................................................................................................1

1.2 Background.........................................................................................................................................1
  1.2.1 Inferring others’ personality traits and preferences...............................................................2
  1.2.2 Assessing informants’ accuracy.................................................................................................3
  1.2.3 Evaluating multiple informants and source information.........................................................4

1.3 Goals of the dissertation......................................................................................................................4

## 2. Using statistical information as cues to mental states: Preschoolers’ understanding of graded preferences

2.1 Introduction...........................................................................................................................................5

2.2 Experiment 1: Inferring graded preferences.....................................................................................6
  2.2.1 Method......................................................................................................................................7
    2.2.1.1 Participants.......................................................................................................................7
    2.2.1.2 Materials and procedure.................................................................................................7
  2.2.2 Results.......................................................................................................................................8
  2.2.3 Interim discussion...................................................................................................................11

2.3 Experiment 2: Inferring graded preferences for objects chosen some of the time.........................11
  2.3.1 Method...................................................................................................................................11
    2.3.1.1 Participants.......................................................................................................................11
    2.3.1.2 Materials and procedure.................................................................................................12
  2.3.2 Results.....................................................................................................................................12
  2.3.3 Interim discussion...................................................................................................................12

2.4 General discussion.............................................................................................................................13

## 3. When does the majority rule? Preschoolers’ trust in majority informants varies by task domain

3.1 Introduction.......................................................................................................................................16

3.2 Experiment 1: Comparing socially constructed and non-socially constructed task types.............17
  3.2.1 Methods...................................................................................................................................18
    3.2.1.1 Participants.......................................................................................................................18
    3.2.1.2 Materials............................................................................................................................18
    3.2.1.3 Procedure..........................................................................................................................18
  3.2.2 Results.....................................................................................................................................19
  3.2.3 Discussion................................................................................................................................20

3.3 Experiment 2: The effect of feedback..............................................................................................20
  3.3.1 Methods...................................................................................................................................20
    3.3.1.1 Participants.......................................................................................................................20
    3.3.1.2 Materials and procedure.................................................................................................21
  3.3.2 Results.....................................................................................................................................21
4. Tracking informants’ knowledge sources

4.1 Introduction

4.2 Experiment 1: Direct knowledge vs. hearsay

4.2.1 Methods

4.2.2 Results

4.3 Experiment 2: Replication with new domain

4.3.1 Methods

4.3.2 Results

4.4 Experiment 3: Hearsay from multiple sources vs. one source

4.4.1 Methods

4.4.2 Results

4.5 Discussion

5. Conclusions
List of figures

2.1 Stimuli used in Experiments 1 and 2 ................................................................. 8
2.2 Children’s responses in Experiment 1 test questions ........................................ 9
2.3 Children’s responses in Experiment 2 test questions. ..................................... 12
3.1 Examples of video clips ......................................................................................... 18
3.2 Results from Experiment 1 .................................................................................. 20
3.3 Results from Experiment 2 .................................................................................. 22
3.4 Results from Experiment 3. ................................................................................ 23
3.5 Overview of average responses from Experiments 1-3. ................................. 24
## List of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Children’s rate of selection of stimulus objects in baseline questions from Experiments 1 &amp; 2</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Experiment 1, question 2 response patterns of children who chose the 100% object over 70% object in question</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Experiment 1, question 2 response patterns of all children in Experiment 1</td>
<td>10</td>
</tr>
<tr>
<td>2.4</td>
<td>Experiment 2, question 2 response patterns of children who chose the 83% object over 63% object in question</td>
<td>13</td>
</tr>
<tr>
<td>2.5</td>
<td>Experiment 2, question 2 response patterns of all children in Experiment 2</td>
<td>13</td>
</tr>
<tr>
<td>3.1</td>
<td>Number of majority and minority responses in Experiments 1-3</td>
<td>25</td>
</tr>
<tr>
<td>4.1</td>
<td>Children’s choices in Experiments 1 and 2</td>
<td>29</td>
</tr>
<tr>
<td>4.2</td>
<td>Children’s choices in Experiment 3</td>
<td>32</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

At the root of developmental psychology is a single big question: how do humans grow from helpless infants to capable adults? Aristotle famously described the newborn human as a *tabula rasa*, a blank slate, while William James made a slightly more generous attribution to infants as experiencing the world as a “blooming, buzzing confusion”. We begin our lives with very little, so how is it that we develop into adults who are capable of extremely complex achievements, like sending man into space?

An obvious starting point to answering this question is to assess children’s perceptual and cognitive skills. Through trial and error, children can learn some basic life skills: how to crawl, which foods taste good, what happens when you flip a light switch on and off. From classic psychologists like Jean Piaget to modern researchers Susan Carey and Alison Gopnik, psychologists have long suspected children are like “little scientists”, who make hypotheses, test them, and, if necessary, revise them in light of new evidence. In the past several decades, developmental psychologists have begun to uncover the richness of children’s inferential abilities. Children have been shown to make complex inferences from small amounts of data in a variety of domains, from language (Kemp, Perfors, & Tenenbaum, 2007; Regier & Gahl, 2007; Saffran, Aslin, & Newport, 1996; Smith & Yu, 2008; Swingley, 2007; Vouloumanos & Werker, 2009; Xu, Dewar, & Perfors, 2009; and Xu & Tenenbaum, 2007) to causality (Buchsbaum, Gopnik, & Griffiths, 2010; Gopnik et al., 2001; Schulz, Gopnik, & Glymour, 2007; Schulz, Hooppell, & Jenkins, 2008).

As sophisticated as children’s inferential skills are, children cannot learn about the world on their own. Researchers have posited that our ability to imitate and teach are uniquely human skills (Tomasello et al., 2005; Warneken & Tomasello, 2009). Boyd, Richerson, & Henrich (2011) suggest that these skills – part of what they refer to as the “cultural niche” – are responsible for our success as a species. They argue that the complex set of skills necessary for survival cannot be learned through trial and error, but rather, have been passed down through generations. For instance, in one example, they give details about the innumerable skills the Inuit must master to survive in the harsh Arctic environment, from the specific stitches used to make sturdy snow shoes, to the tools and techniques necessary for hunting seal through ice. They pose the question whether one could venture into the Arctic and survive alone, without social help. Their conclusion: “We don’t think so.” Learning these skills, they say, only happens through trial and error, and it’s more likely that one would perish in such harsh conditions before figuring out the intricacies of hunting, tool-making, and shelter-building necessary to survive.

Learning from others is a boon to children’s development. Looking to others for information can be a useful shortcut for young children; rather than learning through strictly trial and error, which can be time-intensive and effortful, children can discover new things from peers or adults. Furthermore, some information – such as word meanings or social conventions – can only be learned through interaction with others who share the child’s cultural background.

Recent research has uncovered the wide variety of domains in which children use social information. Through simply observing others’ actions, children have been found to make inferences about a variety of personal attributes.

1.1 Background
1.1.1 Inferring others’ personality traits and preferences

Children use statistical information from people’s actions to predict personal traits and characteristics. Boseovski and Lee (2006) found that three- to six-year-old children interpret frequency of positive or negative behaviors as an indication of an actor’s likelihood to act positively or negatively in the future. Seiver, Gopnik, and Goodman (2012) found that four- and six-year-old children can combine frequency information with prior knowledge about others to determine whether a person’s actions are more likely attributable to a personality trait or a situational factor; for instance, if Sally consistently backed away from riding a skateboard while Josie approached it, children inferred Sally was afraid and Josie was brave, but if Josie and Sally both backed away from riding a skateboard but readily approached a bicycle, children did not infer personality traits about Josie and Sally, but rather, inferred that the skateboard was intrinsically scary while the bicycle was not.

Children also infer others’ preferences from observing their actions. Kushnir, Xu, and Wellman (2010) found that children use statistical sampling as a cue to others’ preferences. Children in this study watched as a puppet retrieved five identical toys (e.g., purple triangles) from a box. Three conditions varied in what the box contained: a population of 100% purple triangles, 50% purple triangles, or 18% purple triangles. Children in the 18% condition were most likely to infer that the puppet preferred the purple triangles, because the puppet’s choice violated random sampling; children in the 100% condition did not infer the puppet had a preference for the triangles, given that random sampling would have produced the same choices.

Furthermore, children infer that like personality traits, others’ preferences are relatively stable and can be used to predict a person’s future actions. Fawcett & Markson (2010) found that children used this information to determine whether actors’ preferences would match their own. Two-year-olds were asked about their preferences about familiar objects like toys and books, and were then introduced to the preferences of two actors, one of whom shared the child’s preferences, and one of whom expressed the opposite preferences. The actors then reacted to novel objects negatively or positively; some of these novel objects were in the same category as the familiar objects, e.g. toys and books, whereas others were from unfamiliar categories. When choosing a novel object from a familiar category, children were more likely to follow the lead of the actor who had previously shared their preferences, but children did not look to these actors for information about the unfamiliar object, presumably because they had no metric of how well their preferences in those categories would overlap.

A model by Lucas et al. (2014) also suggests that children use statistical information to make inferences about others’ preferences. Lucas and colleagues adapted an econometrics model to explain children’s inferences; the model assumes that children believe preferences are relatively stable, and that children make predictions from others’ actions about the value they assign to features of objects, like being purple or triangular (rather than just the objects themselves). This model successfully predicted children’s responses in several preference studies, suggesting children’s inferences could mirror the model’s assumptions.

1.1.2 Assessing informants’ accuracy

Studies showing children’s inferences about personality traits and preferences show that they can use statistical information to learn about people; children’s inferential skills can also help them determine from whom they should learn. Recent research has uncovered that children are quite discerning about their sources. Koenig, Clement, and Harris (2004) showed three- and
four-year-old children testimony from two actors: one who was consistently correct in naming familiar objects like a shoe or ball, and another who was consistently wrong. When the two actors gave testimony about the name of a novel object, children endorsed the label given by the informant who had a record of being correct. In a similar study, Koenig and Harris (2005) found that when children were given the opportunity to choose an informant to give testimony about the name of a novel object, they tended to ask the informant who had a record of being correct over the informant who had a record of being incorrect.

While the previous studies contrasted informants who were consistently accurate or consistently inaccurate, Pasquini et al. (2007) explored whether children use statistical information to assess informants’ relative accuracy. In addition to examining children’s assessment of one informant who was 100% accurate and another who was 0% accurate, Pasquini and colleagues contrasted informants with varying rates of accuracy: 100% accurate informant with one who was 25% accurate; a 75% accurate informant with a 0% accurate informant; and a 75% accurate informant with a 25% accurate informant. The two informants then gave conflicting labels to a novel object. Three-year-olds were inflexible in their assessments; they chose informants who were 100% accurate, and guessed at random between trials with two informants who were less than 100% accurate. However, four-year-olds selected informants with the higher relative accuracy.

Additional studies have found that an informant’s record of accuracy is more important than an informant’s age (Jaswal & Neely, 2007), and that children use this information in domains beyond than language learning, such as object function (Birch, Vauthier, & Bloom, 2008).

1.1.3 Evaluating multiple informants and source information

When evaluating others, children also take into account other information. One evaluation children can make is whether an informant’s receives support from others. Previous research suggests that preschoolers may have a preference for informants who are not dissenters (Fusaro & Harris, 2008; Corriveau, Fusaro, & Harris, 2009). In a series of studies, Harris and colleagues found that children endorse object labels given by the majority of informants rather than one given by a dissenting informant. In a 2010 paper, Corriveau and Harris found that children also defer to majority opinions in the classic Asch (1956) task. Together, this work suggests children can weigh information from multiple informants, and often base their responses on popular opinion.

Beginning at three years, children also understand that visual information can provide informants with knowledge, and that informants can learn new things by looking at them (O’Neill, Astington, & Flavell, 1992; Pillow, 1989; Sodian & Wimmer, 1987). The literature on children’s source monitoring suggests that they are poor at explicitly reporting how they or others obtained knowledge (Gopnik & Graf, 1988; O’Neill & Gopnik, 1991; O’Neill & Chong, 2001; Wimmer, Hogrefe, & Perner, 1988). However, children do succeed at tasks that require them to apply the information they’ve learned from others, rather than making explicit judgments about how they know something. As suggested by Haigh and Robinson (2009), “Children might reveal working understanding… even if they fail to report explicitly how they know something.” Children may not to be able to report to others where their knowledge came from, but are able to use it to inform choices.

For instance, Povinelli & deBlois (1992) found that when finding a hidden prize, preschoolers prize testimony from an informant who hid the prize over testimony from an
informant who was not in the room when the prize was hidden, even though they are unable to explain how they knew where to look. Similarly, Robinson, Champion, and Mitchell (1999) found that three-and-a-half year olds use testimony from an informant if it contradicts their own experiences, as long as the informant’s source of knowledge is more reliable than their own (e.g., an informant sees what object is in a container, while the child has only felt it), but children were inaccurate in reporting the sources of their information and the informant’s. This body of work suggests that visual access and others types of knowledge sources can help children assess informants’ knowledge.

1.2 Goals of the dissertation

The goal of this dissertation is to explore how children learn from and about other people using statistical inference skills. Learning from others requires children to make inferences by combining statistical information with prior knowledge and knowledge about others’ minds. Together, this body of work demonstrates the breadth of variables children consider when learning from others, and the flexibility with which they employ these strategies.

The next three chapters contain empirical studies that demonstrate the sophistication of children’s social inference skills. Chapter 2 examines how children learn about others through statistical information in lieu of other social cues. In these studies, children observe statistical information from watching others’ choice actions and are asked to infer their preferences.

Chapters 3 and 4 examine how children integrate statistical information with other cues. Chapter 3 investigates how children use testimony from a consensus of informants, and how they reconcile statistical information (number of informants) with their personal observations in two domains: a socially constructed domain (object labeling) versus a non-socially constructed domain (learning causal relationships). The fourth chapter explores how children integrate statistical information with informants’ quality of their knowledge; specifically, whether informants have direct knowledge through personal experience, or indirect knowledge through hearsay. The fifth chapter looks at preschoolers’ and adults’ decisions when faced with two conflicting types of information: statistical information (number of endorsements) and quality (overall valence of endorsements). The sixth and final chapter discusses the implications of this work, and suggests possible follow-up questions.
Chapter 2

Using statistical information as cues to mental states: Preschoolers’ understanding of graded preferences

2.1 Introduction

Inferring others’ mental states from their actions is an indispensable social skill. As adults, we often observe others’ actions to infer what they like or dislike. For example, if our colleague Jim buys an orange soda every day from a vending machine with many other options available, we might infer that Jim has a preference for orange soda. Because preferences are relatively stable dispositions, identifying a person’s preferences can help us predict how an agent may behave in the future (e.g., Jim will likely buy orange soda again), how to become a better social partner (e.g., buying Jim an orange soda as a thank-you for a favor), or how much we'll like something new based on our shared preferences with that person (Fawcett & Markson, 2010).

How does the ability to infer preferences develop? One prerequisite for inferring preferences is the recognition that other people are intentional agents. Infants as young as 3 months old can interpret agents’ reaching behavior as goal-directed (Woodward, 1998; Sommerville, Woodward, & Needham, 2005), and 6-month-olds take into account others’ perceptual access when making judgments about agents’ preference for objects they consistently reach towards (Luo & Baillargeon, 2007; Luo & Johnson, 2009). By 12 months, infants track multiple agents’ behaviors and understand that goals are person-specific (Buresh & Woodward, 2007).

To infer others’ preferences, children must also recognize what cues are informative, and when. Previous studies suggest toddlers and preschoolers infer preferences from emotional or verbal cues (Lumeng et al., 2008; Repacholi & Gopnik, 1997). However, these cues are not always available; people do not always wear their emotions on their sleeves. Even worse, these cues could be unreliable. Our emotional states aren’t always a result of our choices, and could be misleading if taken to represent our attitudes towards our choices. For instance, a coworker who appears irritated as she buys coffee is probably not irritated by her choice of beverage – instead, she may have just had a bad morning.

In the case of inferring preferences, the choices an agent makes are often more reliable cues than their emotional or verbal responses. Actions often reflect preferences, as people tend to choose options they like or avoid options they don’t. Recent studies suggest children use statistical evidence to infer preferences; preschoolers and 20-month-old toddlers infer that a puppet has a preference when its choices are inconsistent with random sampling (Kushnir, Xu & Wellman, 2010; Ma & Xu, 2011).

Past studies have only examined children’s ability to identify a preference for one object over another – either the agent likes object X more than object Y, or Y more than X. However, real-life preferences are more complex. Preferences can be graded, with some stronger than others. You may enjoy all flavors of ice cream, but you may prefer strawberry the most, followed by chocolate, and then vanilla.

In addition to identifying preferences from others’ choices, understanding graded preferences also requires making inferences to compare preference strengths for different objects. Research on children’s transitive inference abilities suggests preschoolers may have
some of the elementary skills necessary for inferring graded preferences. The earliest transitive inference studies found that 4-year-old children struggle with transitive inference in word problems (Piaget, 1928; 1955), but this failure could have been due to memory limitations. More recent studies found that 4-year-olds can make transitive inferences about spatial position with help from visual cues like ordered rods or block towers to aid memory (Bryant & Trabasso, 1971; Halford, 1984; Pears & Bryant, 1990; Andrew & Halford, 1998).

Children’s performance in causal learning tasks also suggests they can make transitive inferences about causes and effects. Children as young as three years have demonstrated an understanding that if X causes Y and Y causes Z, then X causes Z, and five-year-olds can explicitly state the necessity of Y in the relationship between X and Z (Shultz, Pardo, & Altmann, 1982). Furthermore, Schulz, Gopnik and Glymour (2007) found that preschoolers successfully identify transitive relationships between causes. Children were introduced to an electronic toy that contained a switch and two gears, and watched an experimenter intervene on the toy (i.e. turning the switch on or off, or removing one of the gears) to infer the causal relationships between the components of the toy. Depending on the result of the experimenter’s interventions, children inferred different transitive causal relationships between the components – for instance, that the switch causes Gear 1 to spin, which causes Gear 2 to spin, versus the Switch causes Gear 2 to spin, which causes Gear 1 to spin.

However, transitive inference in causal relationships differs from transitive inference in preferences in a crucial way: the relationships between causes change with the omission of a causal element, whereas the relationships between preferences do not. To use the Schulz, Gopnik, and Glymour study as an example, if the switch causes Gear 1 to spin, which causes Gear 2 to spin, Gear 1 is necessary for the transitive relationship between the Switch and Gear 2; i.e., without Gear 1, turning on the Switch may not cause Gear 2 to spin. In transitive inferences about preferences, however, the existence of an intermediate preference doesn’t change the relationship between other preferences. For instance, if one prefers X most, Y next most, and Z least, one could still infer that X is preferred to Z without any knowledge of where Y falls on the spectrum of preferences.

The task we used in the current studies asked children to infer an agent’s graded preferences from their choices, and further investigated children’s transitive inference abilities in a social domain. Experiment 1 examined the inferences children make about an agent’s preferences after observing his choices. This task required children to integrate information from two relational premises – the agent’s choices between objects A and C, and B and C – to make inferences about an agent’s overall preferences for A, B, C, and a novel object D.

### 2.2 Experiment 1: Inferring graded preferences

In Experiment 1, participants watched as Duckie, a puppet, chose one of two objects. When object A and object C were presented, Duckie chose A 5 out of 5 times (100%). When objects B and C were presented, Duckie chose B 7 out of 10 times (70%). Thus overall, object A was chosen 100% of the time over competitors, object B was chosen 70% of the time, and object C was only chosen 20% of the time (3 out of 15 times), consistent with a hierarchy of preferences where A is the most preferred, B is the second most preferred, and C is the least preferred (see Figure 1 for a schematic representation of the procedure). Duckie’s choices were designed to control for mere association effects; in order to succeed at inferring this hierarchy of
preferences, children must infer that the relative proportion of Duckie’s choices, not absolute number of choices, indicate his preferences.

Children were asked to make inferences about Duckie’s preferences in two test questions. In the first test question, participants were asked to give Duckie the object they believed he would like the most: object A (chosen in 100% of the trials it appeared in) or B (chosen in 70% of the trials it appeared in). If children tracked Duckie’s actions in the demonstration phase, they should infer that Duckie prefers object A over B.

In the second test question, participants were asked to infer Duckie’s preferences for objects A, B, and C compared to a novel object D. Participants were asked to give Duckie the object he preferred from the following pairs: A and D, B and D, and C and D. If children inferred that Duckie’s choices in the demonstration represented graded preferences, the rates of their choices for A, B, and C compared to D should also be graded. If children inferred that Duckie’s choices represented the strength of his preferences, they should infer that Duckie was very likely to prefer object A over D, somewhat likely to prefer B over D, and unlikely to prefer C over D.

2.2.1 Method

2.2.1.1. Participants. Participants were 31 preschoolers (mean age = 4 years 5 months; range = 44 – 63 months; 14 female). Participants were recruited in a major metropolitan city by mail and phone calls or from local preschools, and were predominantly Caucasian and middle class. An additional two children were tested, but were uncooperative and thus excluded.

2.2.1.2. Materials and Procedure. A set of four novel objects was used in the study (see Figure 2.1); which objects were designated as A, B, C, or D were counterbalanced.

To minimize novelty effects, the study began with a familiarization phase where children played with the four objects. If children did not spontaneously play with them, the experimenter picked up each object, presented it to the child, and asked, “Did you see this one?”

A strong positive or negative preference for any specific object could affect children’s test responses, so each child was asked a baseline question to assess their preferences. The experimenter presented each participant with two randomly chosen objects and asked, “Which one do you like more?” Whether the baseline question was presented before the demonstration phase or after test questions was counterbalanced.

Next, children watched as Duckie made his choices in two blocks of demonstrations: a 70% versus 30% block and a 100% versus 0% block. The order of trial block presentation was counterbalanced. In the 70% versus 30% demonstration block, each child sat at a table across from the experimenter. The experimenter asked the child, “Should we invite my friend Duckie to play with us?“ and introduced the puppet. Then, the experimenter placed objects B and C on the table, approximately 18 inches apart, and said, “Let’s ask Duckie which one he wants to play with.” Duckie appeared from beneath the table and stood behind it, equidistant from the objects. Before each of Duckie’s choices, the experimenter asked, “Duckie, which one?” Duckie looked back and forth between the objects as if examining his options, then picked one up, held it for one second, and placed it back on the table before returning to his spot behind the table. This was repeated for a total of ten trials. In seven of the ten trials, Duckie chose object B. In the other three trials, Duckie chose object C. The side on which each object appeared was counterbalanced across participants, and the order of Duckie’s looks between objects B and C and the order in which he made his choices were randomized.
The 100% versus 0% demonstration block was the same as the 70% versus 30% block, except that Duckie chose between objects A and C, and made only five choices. Duckie chose object A all five times.

Finally, children were asked two test questions. In Test Question 1, participants were presented with objects A and B and asked, “Which one does Duckie like more?”

In Test Questions 2a, 2b, and 2c, participants were asked about Duckie’s preferences for each of the objects in the demonstration phase (A, B, and C) compared to a novel object D. For each comparison, participants were asked, “Can you give Duckie the one he likes more?” The order in which 2a, 2b, and 2c were asked was randomized.

![Figure 2.1. Stimuli used in Experiments 1 and 2.](image)

### 2.2.2 Results

Binomial tests revealed that no objects were selected at higher-than-chance rates on the baseline trials, suggesting that stimulus objects were comparably appealing to participants. Table 2.1 presents children’s baseline responses in both experiments. Fisher’s exact tests found no effects of sex, demonstration block (70% v. 30% block or 100% v. 0% block first), or the order of questions 2a, 2b, and 2c in responses to test questions.
Table 2.1

*Children’s rate of selection of stimulus objects in baseline questions from Experiments 1 & 2.*

<table>
<thead>
<tr>
<th>Object #</th>
<th>Object description</th>
<th>Number of times object appeared as baseline option</th>
<th>Number of times selected by a participant</th>
<th>Rate of selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Candy dispenser</td>
<td>58</td>
<td>28</td>
<td>48%</td>
</tr>
<tr>
<td>2</td>
<td>Disc</td>
<td>64</td>
<td>28</td>
<td>44%</td>
</tr>
<tr>
<td>3</td>
<td>J-shaped object</td>
<td>32</td>
<td>33</td>
<td>53%</td>
</tr>
<tr>
<td>4</td>
<td>L-shaped object</td>
<td>58</td>
<td>36</td>
<td>62%</td>
</tr>
</tbody>
</table>

Figure 2.2 shows the main results of Experiment 1. In Test Question 1, 25 of 31 children (81%) chose object A over B, binomial test, p < .001. In Test Question 2, 24 of 31 children (77%) chose object A over the novel object D, 19 of 31 children chose (61%) object B over the novel object D, and 12 of 31 children (39%) chose object C over the novel object D. A Cochran’s Q test revealed a statistically significant difference in the rate of selecting the familiar objects (A, B, C) over the novel object D in questions 2a, 2b, and 2c, $\chi^2 (2, N = 31) = 8.07$, p < .01.

![Exp 1 Q1: 100% object v. 70% object](image1)

![Exp 1 Q2: Comparisons to novel object](image2)

Figure 2.2. Children’s responses in Experiment 1 test questions. In Test Question 1, children were asked whether Duckie would prefer object A, chosen in 100% of demonstration trials it appeared in, or object B, chosen in 70% of demonstration trials it appeared in. In Test Question 2, children were asked to infer Duckie’s preferences for each of the three familiar objects seen in demonstration trials to a novel object.

We also examined children’s patterns of responses for additional evidence that *individual* children inferred the gradedness of Duckie’s preferences. Since children were asked for responses to binary questions, data from individual children can reveal indirect yet useful clues about children’s inferences. If children did indeed infer that A > B > C, they would have chosen the 100% object over the 70% object in Test Question 1 (25 of 31 children). Additionally, their
choices in Test Questions 2a, 2b, and 2c should reflect this belief; they should infer that A is at least as likely as B to be preferred over the novel object, and that B should be at least as likely as C to be preferred over the novel object. There are a total of eight possible patterns children’s responses in Test Question 2 could take (see Table 2.2 for a full list), and there are four patterns that could reflect a belief that A > B > C. One would be choosing the 100% object in 2a, the novel object in 2b, and novel object in 2c; if Duckie prefers A the most, he should choose it over a novel object, but would still prefer the novel object over B or C. Another pattern would be to choose the 100% object in 2a, the 70% object in 2b, and the novel object in 2c; this would reflect an understanding that C is least preferred, given that the novel object was chosen over it but not A or B. A pattern of choosing all familiar objects (100%, 70%, 20%) or all novel objects could also reflect this belief.

Of the 25 children who chose the 100% object over the 70% object in question 1, 18 responded with a pattern in question 2 that suggests they inferred A > B > C (binomial test p < .05; see Table 2.2). In an analysis including all participants (including those who chose the 70% object over the 100% object in question 1), 20 of 31 children showed one of the patterns (p = .14; see Table 2.3).

### Table 2.2

**Experiment 1, question 2 response patterns of children who chose the 100% object over 70% object in question 1.**

<table>
<thead>
<tr>
<th>Choice in 2a</th>
<th>Choice in 2b</th>
<th>Choice in 2c</th>
<th>Number of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>70</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>70</td>
<td>novel</td>
<td>11</td>
</tr>
<tr>
<td>100</td>
<td>novel</td>
<td>novel</td>
<td>5</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>novel</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>novel</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>novel</td>
<td>70</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>novel</td>
<td>70</td>
<td>novel</td>
<td>2</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>20</td>
<td>3</td>
</tr>
</tbody>
</table>

*Note: Response patterns suggesting individual children inferred A>B>C are bolded.*

### Table 2.3

**Experiment 1, question 2 response patterns of all children in Experiment 1.**

<table>
<thead>
<tr>
<th>Choice in 2a</th>
<th>Choice in 2b</th>
<th>Choice in 2c</th>
<th>Number of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>70</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>100</td>
<td>70</td>
<td>novel</td>
<td>11</td>
</tr>
<tr>
<td>100</td>
<td>novel</td>
<td>novel</td>
<td>5</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>novel</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>novel</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>novel</td>
<td>70</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>novel</td>
<td>70</td>
<td>novel</td>
<td>3</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>20</td>
<td>3</td>
</tr>
</tbody>
</table>

*Note: Response patterns suggesting individual children inferred A>B>C are bolded.*
2.2.3 Interim Discussion

Children’s choices on test questions suggest they used Duckie’s choices to infer a hierarchy of graded preferences with object A (chosen 100% of the time over competitors) being the most preferred, followed by B (chosen 70% of the time over competitors) and C (chosen 20% of the time over competitors). As a group, children’s responses indicate that they inferred that Duckie preferred A over B. Furthermore, children inferred that the likelihood Duckie would prefer a novel object over an object shown in demonstration was graded with respect to the consistency with which Duckie chose it during demonstrations. Because Duckie chose A consistently, children inferred he would likely continue to prefer it over a novel object, whereas he was only somewhat likely to prefer B over the novel object, and unlikely to prefer C over the novel object.

These results also suggest that children can make transitive inferences based on their observations. In order to succeed in this experiment, children had to integrate several relational premises to make inferences about an agent’s graded preferences. To form a mental hierarchy that reflects these graded preferences, children had to first observe each set of choices to make inferences about their relative value (e.g., A is strongly preferred over C, and B is somewhat preferred over C), and then integrate those relational premises to make inferences about the choices’ relative values: A>>C and B>C, therefore A>B.

A possible alternative explanation for these results is that children used a heuristic to answer test questions without understanding that Duckie’s preferences were graded. Rather than tracking the consistency of Duckie’s choices, children could have used a simpler heuristic to infer Duckie’s preferences. For instance, children could have inferred that object A was strongly preferred because it was chosen all of the time, whereas objects B and C were less preferred because they were chosen only some of the time. Children’s responses to Test Questions 2b (12/31 children chose novel object) and 2c (19/31 chose novel object) suggest that they inferred the gradedness of preferences when comparing objects chosen some of the time, but this issue could be more thoroughly explored. Experiment 2 was designed to address this issue, and to extend and replicate the results of Experiment 1.

2.3 Experiment 2: Inferring graded preferences for objects chosen some of the time

In Experiment 2, we tested preschoolers with a procedure similar to the one used in Experiment 1, except that during demonstrations, Duckie chose all objects only some of the time. If children rely on a simple heuristic like inferring objects chosen all of the time are preferred over objects chosen only some of the time, they would fail at this task. Alternatively, children’s success would provide further support that they are indeed tracking the proportion, not absolute number, of Duckie’s choices to infer his graded preferences.

2.3.1 Method

2.3.1.1. Participants. Participants were 35 preschoolers (mean age = 4 years 1 month; range = 42 – 62 months; 24 girls). Participants were recruited from a major metropolitan city by mail and phone calls or from local preschools, and were predominantly Caucasian and middle class. An additional eight children were tested, but excluded due to experimenter error (6), video malfunction (1), and uncooperativeness (1).
2.3.1.2. Materials and Procedure. The same materials and general procedure from Experiment 1 were used in Experiment 2. The only difference between the two procedures was the set of choices Duckie made during the two demonstration blocks.

The 63% versus 37% demonstration block was identical to the first demonstration block in Experiment 1, except Duckie made a total of 11 choices. In 7 of the 11 trials (63%), Duckie chose object B. In the other four trials, Duckie chose object C.

The 83% versus 17% demonstration block was identical to the second demonstration block in Experiment 1, except that Duckie made a total of 6 choices. Duckie chose object A 5 out of 6 times (83%), and C 1 out of 6 times (17%).

2.3.2 Results

Binomial tests revealed that no objects were selected at higher-than-chance rates on the baseline trials, suggesting that stimulus objects were comparably appealing to participants. Fisher’s exact tests found no effects of sex, demonstration block (63% v. 37% or 83% v. 17% block first), or the order of questions 2a, 2b, and 2c in responses to test questions.

Figure 2.3 shows the main results of Experiment 2. In Test Question 1, 24 of 35 children (69%) chose object A over B, binomial test, p = .04. In Test Question 2, 25 of 35 children (71%) chose object A over the novel object D, 18 of 35 children chose (51%) object B over the novel object D, and 13 of 35 children (37%) chose object C over the novel object D. A Cochran’s Q test revealed a statistically significant difference in the rate of selecting the familiar objects (A, B, C) over the novel object in questions 2a, 2b, and 2c, $\chi^2 (2, N = 35) = 8.34, p < .01$.

2.3.3 Interim discussion

Like in Experiment 1, children’s patterns of choices in Experiment 2 suggest that individual children inferred the gradedness of Duckie’s preferences. If children inferred that Duckie’s preferences were A > B > C, they would have chosen the 83% object over the 63% object in question 1, and in question 2, they should infer that A is more or just as likely to be
preferred over the novel object than B, and that B should be more or just as likely to be preferred over the novel object than C. The patterns that would reflect this belief are: 83% over novel object (2a), 63% over novel object (2b), novel object over 29% object (2c); 83%, novel, novel; 83%, 63%, 29%; novel, novel, novel.

Of the 24 children who chose the 83% object over the 63% object in question 1, 20 children’s responses fit one of those patterns (binomial test p < .001; see Table 2.4). In an analysis including children who chose the 63% object over the 83% object in question 1, 26 of 35 children’s responses matched the patterns (binomial test p < .01; see Table 2.5).

Table 2.4

<table>
<thead>
<tr>
<th>Choice in 2a</th>
<th>Choice in 2b</th>
<th>Choice in 2c</th>
<th>Number of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>63</td>
<td>29</td>
<td>5</td>
</tr>
<tr>
<td>83</td>
<td>63</td>
<td>novel</td>
<td>8</td>
</tr>
<tr>
<td>83</td>
<td>novel</td>
<td>novel</td>
<td>3</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>novel</td>
<td>4</td>
</tr>
<tr>
<td>83</td>
<td>novel</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>novel</td>
<td>63</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>novel</td>
<td>63</td>
<td>novel</td>
<td>0</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>29</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Response patterns suggesting individual children inferred A>B>C are bolded.

Table 2.5

<table>
<thead>
<tr>
<th>Choice in 2a</th>
<th>Choice in 2b</th>
<th>Choice in 2c</th>
<th>Number of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>63</td>
<td>29</td>
<td>7</td>
</tr>
<tr>
<td>83</td>
<td>63</td>
<td>novel</td>
<td>8</td>
</tr>
<tr>
<td>83</td>
<td>novel</td>
<td>novel</td>
<td>6</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>novel</td>
<td>5</td>
</tr>
<tr>
<td>83</td>
<td>novel</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>novel</td>
<td>63</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>novel</td>
<td>63</td>
<td>novel</td>
<td>3</td>
</tr>
<tr>
<td>novel</td>
<td>novel</td>
<td>29</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: Response patterns suggesting individual children inferred A>B>C are bolded.

2.4 General Discussion

In two experiments, we found that children observed an agent’s choices to infer his graded preferences. Children inferred that Duckie preferred an object chosen 100% of the time over an object chosen 70% of the time, or an object chosen 83% of the time over an object chosen 63% of the time. Children also used Duckie’s choices to predict his future actions; their responses suggest that they inferred that objects chosen less consistently were less preferred and
less likely to be preferred compared to a novel object, whereas objects chosen consistently were more preferred and more likely to be preferred compared to a novel object.

These studies provide the first evidence that young children use statistical evidence to infer graded preferences. Without explicit emotional or verbal cues, children can use statistical evidence from choice actions to infer the strength of an agent’s preferences, and use those inferences to make predictions about what an agent would prefer in the future. Children do not simply represent preferences as all or none; instead they can use statistical evidence to establish a hierarchy of preferences.

These results also have important implications for the study of children’s ability to make transitive inferences. Whereas previous studies have found that 4-year-olds succeed in basic transitive inferences (A>B, B>C, ∴ A>C), children’s success in these experiments indicate competence in making more sophisticated, indirect transitive inferences. In our studies, objects A and B were never directly compared. The only available information about the relationship between A and B was each of their relationships with object C: Duckie chose A more consistently over C (A>>C), whereas Duckie chose B only somewhat consistently over C (B>C).

These studies also suggest new questions for future investigation. First, the mechanism children are using to infer Duckie’s preferences could be further explored. There are several strategies children could use to infer the relative strengths of preferences. An easy strategy would be to track the raw frequency of an agent’s choices, and to assume that the number of times an agent chooses an option has a direct relationship to the agent’s preference strength. For instance, if A is chosen five times and B is chosen seven times, then the agent has a stronger preference for B than for A. Previous work has found that infants succeed at tracking frequency information in linguistic input, even in noisy environments (Smith & Yu, 2007; Vouloumanos & Werker, 2009), and preschoolers track the frequency of a person’s positive or negative behaviors to predict their future behaviors and infer their personality traits (Boseovski & Lee, 2006). These studies suggest preschoolers should be adept at tracking the frequencies of different choices.

However, this set of studies suggests that children are not only tracking frequency of choices. Even though children were shown that Duckie chose object B a greater number of times than A (seven times versus five in Experiment 1, or seven times versus four in Experiment 2), children still inferred that Duckie preferred A over B.

A more sophisticated strategy for determining preference strength would require children to go beyond raw frequency counts and consider the agent’s choices relative to the total number of choices made. The number of times an option is chosen only matters in the greater context of how many total choices there were; the proportion of times an object is chosen could indicate preference strength. Indeed, in this study, children’s responses indicate that they inferred that Duckie preferred toys he chose a high proportion of times over toys that he chose a lower proportion of times; for example, they inferred he preferred object A (chosen 100% or 83% over competitors) over object B (chosen 70% or 63% over competitors).

Yet there remains an alternate explanation. Rather than tracking the proportion of Duckie’s choices, children could instead track the frequency with which Duckie rejects a toy. For instance, in Experiment 2, children could have inferred that object C was least preferred because it was rejected the most often (12 out of 17 choices), that object B was next least preferred (rejected 4 out of 11 choices), and that object C was most preferred because it was rejected only once out of 6 choices. To our knowledge, there have been no published studies that suggest children track the frequency of rejected items in choice tasks, but the procedure of these two experiments does not allow us to rule out the possibility that children could be using frequency
of rejections as a cue to Duckie’s preferences. Further investigation is necessary to determine whether children in this study were tracking Duckie’s proportion of choices or frequency of rejections, as both strategies could lead children to infer an A>B>C hierarchy in Experiments 1 and 2.

Additionally, in these studies, Duckie’s choices indicated his preferences, but other factors, like the amount of data observed, can also affect inferences about preference. We might infer that choosing an object 99 out of 100 times (99%) reflects a stronger preference than choosing it 1 out of 1 time (100%), even though it was chosen more consistently in the latter case. The number of observations affects our confidence in making predictions about preferences. In the 1 out of 1 case, we have very little information with which to predict the strength of the agent’s preference and whether (s)he will continue to prefer this object in future choices. However, in the 99 out of 100 case, we have much more information, and the agent’s choices give us more confidence that (s)he strongly prefers the object chosen. Recent Bayesian modeling work suggests that the amount of available information should influence inferences about preferences (Lucas et al., 2014), and future work can investigate whether children share this intuition.

While the current studies focused on children’s use of statistical evidence to make inferences about positive preferences, there has been little investigation into how children infer negative preferences (i.e., dislike). Children interpret choice actions as evidence for a preference. Would they interpret lack of choice as evidence for a negative preference? In our studies, it is unclear whether children interpreted Duckie’s relatively few choices of object C as a very weak preference, or a dispreference. It is also unclear whether children can use choice behavior to infer the strengths of negative preferences.

Lastly, choices are not always indicative of a positive or negative preference. Positive and negative preferences involve positive or negative appraisals, whereas choices can be indicative of a neutral internal mental state, such as a goal. For example, imagine that you observe a colleague picking up a stapler. You are unlikely to infer that she has a preference for staplers. Rather, you might infer that she has a goal of stapling paper together. Future studies may examine when children develop the understanding that choices may reflect different mental states such as goals, preferences, and beliefs.
Chapter 3

When does the majority rule? Preschoolers’ trust in majority informants varies by task domain

3.1 Introduction

We humans are inherently social creatures, and throughout our daily interactions, we openly share our thoughts and opinions with one another. The ubiquity of our social sharing and learning is rare among animal species (Tomasello et al., 2005; Warneken & Tomasello, 2009), and has been cited as an explanation for the robustness of human culture (Boyd, Richerson, & Henrich, 2011). Listening to others who share their knowledge can save precious time and effort, as learning through experience can be difficult and time-consuming. By asking others for their testimony, we can instantly and effortlessly learn how to prepare a dish, where to hunt, or who to hire to fix the kitchen sink.

Learning from others is especially important for young children, who have a relatively small pool of life experiences to draw on in new situations. However, one potential drawback to social learning is the possibility of receiving incorrect or misleading information. Therefore, it would be advantageous for children to employ mechanisms to evaluate the reliability of their information sources. Previous work has found that children use informants’ past accuracy as an indicator of trustworthiness (Sabbagh & Baldwin, 2001, Birch, Vauthier, & Bloom, 2008; Koenig, Clement & Harris, 2004; Pasquini et al., 2004; Corriveau & Harris, 2009) and selectively imitate others’ actions (Gergely, Bekkering, & Kiraly, 2002; Brugger et al., 2007; Buchsbaum et al., 2011; Schulz, Hooppell, & Jenkins, 2008). On the other hand, other studies suggest children’s social learning is sometimes surprisingly unselective and irrational. Lyons, Young, & Keil (2007) showed that preschoolers often “overimitate” by imitating adults’ causal unnecessary actions, like tickling the side of a box with a feather before opening it. Similarly, McGuigan and colleagues (2007) showed that 5-year-olds were actually more likely to overimitate than 3-year-olds.

We can learn not only from reliable individuals, but also from “crowd sourcing” information from a group of people. Adults often turn to others for advice, assuming that opinions held by many must be valid by virtue of their popularity. This intuition echoes the law of large numbers in probability theory: the more individual testimonies, the more likely the collective conclusion of those testimonies is accurate. Corriveau, Fusaro, and Harris (2009) found that three- and four-year-old children view consensus as an indication of reliability; they were more likely to endorse novel object labels that received majority support, and to choose a member of the majority group as an informant about other object labels. The authors concluded that preschoolers prefer information endorsed by the majority, and prefer members of a majority as informants.

If children use purely statistical information to make decisions about whose testimony to endorse, then children should always endorse the majority opinion. However, depending on demands of a task, tracking statistical information may not always be useful. For instance, keeping track of the number of times an opinion is stated may not be useful to children unless the child is also tracking who says what; an opinion stated 3 times by one person is different from an opinion stated 3 times, each by a different person. Simply tracking statistics would not take into account this additional but equally important information about source of knowledge. Thus,
children should be able to flexibly integrate statistical information with outside cues, such as identity of individuals providing statistical information, or task domain. This project explores how children’s endorsement of majority testimony varies depending on domain. In some domains, like causal learning, the relevant knowledge is a directly observable effect, so one could rely on personal observations to make judgments. In other domains – specifically, those that are strongly socially constructed, like object labeling – directly observable effects are not available, and the “correct” answer is entirely determined by majority consensus. In this case, it is rational to rely on others’ testimony.

While previous research has established that children often do favor majority testimony, the extent to which children prefer members of a majority as informants is still unclear. One possibility is that children prefer majority members as informants in all situations where multiple testimonies are available. In this view, children would indiscriminately weigh information from others as the most valuable source of information, perhaps prizing it above their own observations. Social psychologists have discovered that consensus opinions can override adults’ existing opinions (Asch, 1956; Cialdini & Goldstein, 2004). Even when adults don’t believe the majority’s endorsement and merely go along with the group’s consensus to avoid social conflict (called “compliance”), studies have shown that the act of complying can later result in adults believing or “internalizing” that testimony (Kelman, 1958; Nolan et al., 2008).

However, if children are rational learners, they should not always prefer majority testimonies. The demands of the learning task, as well as children’s prior knowledge of a domain’s demands, should affect the weight children place on others’ testimony. When learning new facts that are heavily socially constructed (e.g. object labels or tool use conventions), testimony from others should be highly valuable because the relevant knowledge is transmitted through others, and children cannot learn this type of information on their own. By virtue of the social conventions that dictate object labeling, typically only one label is regarded as correct (Markman, 1989). Alternatively, learning new facts that are not socially constructed (e.g. causal relationships, or naïve physics), children should consider not only informants’ endorsements but also their own knowledge, because this type of knowledge can be gained through personal observation and is not typically bound by social convention. This leaves open the possibility that an effect can have multiple causes.

In this set of studies, we explore how children’s endorsement of majority testimony varies as a function of fact type. Specifically, we compare children’s endorsement of majority testimony in an object labeling task versus a causal learning task. We predict that when given two options – one endorsed by a three-person majority, and one endorsed by a single minority informant – children should be more likely to endorse the majority’s testimony when learning socially constructed facts (object labeling) than when learning non-socially constructed facts (causal learning).

### 3.2 Experiment 1: Comparing socially constructed and non-socially constructed task types

In this study, we present preschoolers with four informants’ conflicting testimony about objects. In the object labeling condition, informants identify the referent of a novel label, and in the causal learning condition, they demonstrate a novel action on the object that results in an effect, e.g., music playing.
3.2.1 Methods

3.2.1.1 Participants Participants were 36 preschoolers, 10 male and 26 female (mean age = 4 years 2 months; range = 38 – 62 months). Participants were recruited in the San Francisco Bay Area by mail and phone calls or from local preschools and museums. An additional four children were tested, but were excluded due to uncooperativeness (3) or experimenter error (1).

3.2.1.2 Materials In the object labeling condition, stimuli were four novel objects. In the causal condition, stimuli were two plush toys, each of which contained a wireless, battery-powered doorbell chime box. The boxes played short melodies when activated by a handheld remote to create the illusion that children’s actions were causally efficacious. Pre-recorded video clips of informants’ testimonies were shown to children on a 13” laptop screen.

3.2.1.3 Procedure Participants were randomly assigned to either the object labeling condition or causal condition. Each participant participated in two test trials of their condition.

In each condition, the experimenter introduced novel objects to participants and explained that they were unknowledgeable about their labels or causes. Participants then watched four video clips of four informants evaluating the objects (see Figure 3.1 for example).

Each clip began with a female informant sitting at a table with the novel objects. She visually inspected them, then picked up one of the toys and called it by the novel label (e.g. modi), or acted on the toy, resulting in the toy playing a short song. In three of the four video clips, the majority informants each endorsed one object as a modi or performed one action to elicit music, and the minority informant endorsed the other object as the modi, or performed an alternate action to elicit music. The minority informant always repeated the novel label or alternate action three times so that each participant heard the label used to refer to each object an equal number of times, or the action 3 times.

After participants watched the video clips, the experimenter presented the child with the objects from the video clips and asked children to identify the referent of the novel label, or to make the toy play music. Participants were invited to name or activate the toy a total of three times, and their gestural or vocal responses were recorded.

Half of the video clips were mirror images of original recordings to control for the location of objects (object labeling condition) and handedness of informants when manipulating toys (causal condition). The trial presented first and identity of the minority informant were also counterbalanced, and the minority informant always presented her opinion last.
3.2.2 Results

Participants were assigned a score (0-6) based on the number of trials in which they endorsed the majority informants’ testimony first responses in the two trials they participated in (see Table 3.1).

Children in the object labeling condition endorsed majority testimony significantly more often than children in the causal condition, \( t(34) = -2.55, p < .01 \). Children endorsed majority testimony over the minority informant’s significantly more often than chance in object labeling trials \( t(17) = 12.12, p < .001 \), but not in causal trials, \( t(17) = 0.77, p = .45 \) (see Figure 3.2).

Out of 108 total responses given by participants in the language condition, 83 (77% of total attempts) were the object labels given by majority informants, while in the causal condition, only 59 (55% of total attempts) were actions performed by majority informants (see Table 3.1). All participants in the causal condition attempted at least one action performed by the minority informant, while only 11/18 children did in the object labeling condition (\( p < .001, \phi = 0.49 \)).

Additionally, though there was no formal coding scheme for children’s spontaneous comments during the study sessions, anecdotal evidence suggests that children’s intuitions matched our model assumptions about mutual exclusivity. In the causal condition, some children expressed a belief that both the majority testimony and the minority informant’s testimony were not mutually exclusive (“Both [actions] make it go!”). Furthermore, all children in this condition attempted an action performed by a minority informant at least once in the study, suggesting that children were open to multiple possibilities when learning about cause and effect. In the object labeling condition, however, several children expressed the belief that there was only one correct answer (“That one isn’t the modi!” about the minority-endorsed object).
Figure 3.2. Results from Experiment 1. Children in the language condition performed significantly more majority responses than children in the causal condition.

3.2.3 Discussion

These results show children were more likely to endorse majority testimony in a socially constructed domain (object labeling) than a non-socially constructed domain (causal learning).

A possible alternative explanation of these results is that children in the causal condition did not use information gained through their personal observations or informants’ testimony; rather, they were simply confused by the task and randomly imitated informants’ responses. To rule out this possibility, we designed another causal condition in which we expected children to endorse the majority testimony.

3.3 Experiment 2: The effect of feedback

In the causal condition of Experiment 1, children indiscriminately imitated the majority and minority informant actions, presumably because they were able to rely on their own observations, which suggested both demonstrated actions were equally effective at activating the toy. This second experiment examines how children behave when they do not have their own observations to rely on, but instead only have information from informants. We predict that when children lack personal observations indicating the efficacy of informants’ testimony, they should rely on the majority.

3.3.1 Methods

3.3.1.1 Participants Participants were 36 preschoolers, 20 male and 16 female (mean age = 4 years 5 months; range = 40–65 months). Participants were recruited in the San
Francisco Bay Area by mail and phone calls or from preschools. An additional four children were tested, but excluded due to uncooperativeness (2) or failure to complete the second trial (2).

### 3.3.1.2 Materials and Procedure

The materials and general procedure of Experiment 2 were identical to those used in Experiment 1. The crucial difference between the two experiments was the content of the video clips participants watched. Participants were randomly assigned to either the experimental “no feedback” condition, or a control condition. Each participant participated in two test trials of their condition.

In the “no feedback” condition, informants in the video clips only mimed the actions they endorsed, and no music played as a result of miming the actions. While in Experiment 1, children watched informants in the video clips perform actions that successfully resulted in the toy playing music, children who viewed Experiment 2’s “no feedback” condition video clips received no information about the efficacy of the informants’ testimony.

This necessitated a change in the script of the videos. Unlike the informants in Experiment 1, who did not verbally describe the action they performed, informants in the video clips shown in the no feedback condition of Experiment 2 explicitly described their endorsed action and its hypothetical causal effect before miming the action, in order to provide context to children about why the action was being mimed: “It plays music if you pull the pink one!”

The control condition of Experiment 2 was designed to rule out the possibility that this change in the video actors’ language accounts for any differences between the Experiment 2 no feedback condition and Experiment 1 causal condition. The procedure for the control condition of Experiment 2 was exactly the same as that of Experiment 1’s causal condition, but with the language used in Experiment 2’s no feedback condition. We predicted that given the feedback provided by the toy playing music, participants in Experiment 2’s control condition would perform similarly to participants in Experiment 1’s causal condition.

### 3.3.2 Results

As in Experiment 1, Experiment 2 participants were given a score of 0-6 based on their responses in test trials. Children were significantly more likely than chance to endorse majority informants’ testimony in the no feedback condition \(t(17) = 3.96, p < .01\), and they were also more likely to endorse majority informants’ testimony in the no feedback condition than in the control condition \(t(34) = -2.16, p < .03\) (see Figure 3.3). They were also significantly more likely to endorse majority informants’ testimony in the no feedback condition than in Experiment 1’s causal condition, \(t(34) = -1.99, p < .05\) (see Figure 3.5).

Out of 108 total attempts made by participants in the no feedback condition, 76 (70% of total attempts) were actions performed by majority informants, while in the control condition, 61 (56% of total attempts) were actions performed by majority informants (see Table 3.1). Unlike in Experiment 1, where all causal condition participants attempted at least one action performed by the minority informant, five of the 18 participants in the no feedback condition attempted only actions endorsed by the majority, \(p < .05, \phi = 0.40\).
Figure 3.3. Results from Experiment 2. Children in the no feedback condition performed significantly more majority responses than children in the control condition.

3.3.3 Interim discussion
In Experiment 1, children learning facts about cause and effect, a non-socially determined domain, were at chance in endorsing the majority testimony and minority testimony. Results from Experiment 2 show that when children were not able to make personal observations about the efficacy of actors’ actions, children were more likely to default to endorsing majority testimony. However, because actors in Experiment 2 were miming actions in the “no feedback” condition, the language in the videos included explanation of the actors’ actions: “It plays music if you…” This language could have implied that actors were knowledgeable, and this implication could have biased children’s responses. We designed Experiment 3 to address this problem.

3.4 Experiment 3: The effect of feedback and pedagogy
In Experiment 2, children in a causal “no feedback” condition imitated the majority informant’s actions at a rate greater than chance, but in the control condition, when they received feedback showing that both majority informants’ and the minority informant’s actions were equally efficacious, they no longer relied on majority’s action to make the toy play music. This result could have been due in part from the pedagogical language used in the videos in Experiment 2, so the goal of Experiment 3 was to examine the effects of pedagogical language on children’s endorsement of majority opinions.
3.4.1 Methods

3.4.1.1 Participants Participants were 36 preschoolers, 19 male and 17 female (mean age = 4 years 5 months; range = 41 – 61 months). Participants were recruited in the San Francisco Bay Area by mail and phone calls or from preschools. An additional nine children were tested, but were excluded due to failure to finish the study (3), uncooperativeness (6), or experimenter error (1).

3.4.1.2 Materials and Procedure The materials and general procedure of Experiment 3 were identical to those used in Experiments 1 and 2. Like in Experiment 2, participants were randomly assigned to either the experimental “no feedback” condition, or a control condition, and each participant participated in two test trials of their condition. The crucial difference between Experiments 2 and 3 was the language used in the video clips participants watched.

In Experiment 2, actors provided a pedagogical explanation before miming or performing their action: “It plays music if you…. [endorsed action]!” In Experiment 3, actors explicitly stated that they were not knowledgeable about the toy, and framed their endorsed action as a guess: “I’ve never played with this toy before! Maybe it plays music if you…”

![Figure 3.4](image_url) Results from Experiment 3. Responses from these two conditions did not differ significantly.

3.4.2 Results

As in the other two experiments, participants were given a score of 0-6 based on their responses in test trials. Children in the no feedback condition were significantly more likely than chance to imitate majority actions, t(17) = 3.96, p < .001, whereas children in the control condition were not, t(17) = 1.68, p = .11. However, an independent samples t-test comparing the mean number of majority responses of these two groups found that responses in the two conditions did not significantly differ from one another, t(34) = 1.28, p = .20 (see Figure 3.4). In
contrast, children were significantly more likely to endorse majority informants’ testimony in Experiment 3’s no feedback condition than in Experiment 1’s causal condition, \( t(34) = 1.99, p < .05 \) (see Figure 3.5).

Out of 108 total attempts made by participants in the no feedback condition, 76 (70% of total attempts) were actions performed by majority informants, while in the control condition, 65 (60% of total attempts) were actions performed by majority informants (see Table 3.1).

**Figure 3.5.** Overview of average responses from Experiments 1-3.
Table 3.1.

<table>
<thead>
<tr>
<th></th>
<th>Number of trials in which children performed action endorsed by majority actors</th>
<th>Number of trials in which children performed action endorsed by minority actor</th>
<th>Percentage of trials in which children performed majority action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp 1 language</td>
<td>83</td>
<td>25</td>
<td>77%</td>
</tr>
<tr>
<td>Exp 1 causal</td>
<td>59</td>
<td>49</td>
<td>55%</td>
</tr>
<tr>
<td>Exp 2 no feedback</td>
<td>76</td>
<td>32</td>
<td>70%</td>
</tr>
<tr>
<td>Exp 2 control</td>
<td>61</td>
<td>47</td>
<td>56%</td>
</tr>
<tr>
<td>Exp 3 no feedback</td>
<td>76</td>
<td>32</td>
<td>70%</td>
</tr>
<tr>
<td>Exp 3 control</td>
<td>65</td>
<td>43</td>
<td>60%</td>
</tr>
</tbody>
</table>

3.5 Discussion

This set of studies suggests that children do not indiscriminately endorse majority opinions; rather, their endorsement of majority opinions varies by domain and availability of alternate sources of knowledge. In Experiment 1, children were significantly more likely to endorse majority testimony when learning about a socially constructed domain (object labels) than a non-socially constructed domain (causal relationships). Experiment 2 found that in the absence of information about the efficacy of informants’ actions, children endorse majority testimony. However, without the cue of pedagogical language, children’s responses in the two conditions of Experiment 3 did not differ significantly.

There are at least two explanations for the results of Experiment 3. The simplest explanation would be that any effect of feedback in this study was quite small, and that our relatively small sample size (n=18 per condition) was not sufficient to detect any significant differences between our experimental and control conditions.

Another possibility is that the language used by the actors in this experiment confused children. Actors explicitly stated that they had never played with the toys in the videos before, yet still ventured a guess about how to make the toy play music. It is possible that children did not believe the actors were guessing or truly “unknowledgeable” – after all, why would an unknowledgeable informant guess that a toy would play music at all, let alone all happen to guess the same thing?

In comparing the results of Experiment 2 to Experiment 3, it seems plausible that both of these explanations could have contributed to the results we obtained. In the “no feedback” conditions of these two experiments yielded the same results – 76 of 108 responses were majority responses – while the control conditions were different – 61 of 108 responses in Experiment 2 were majority responses, compared with 65 of 108 in Experiment 3. This slight difference suggests small sample size could have played a role. Additionally, confusion or incredulity that actors were truly “guessing” could have led to a slightly higher number of majority responses in the control condition of Experiment 3.

Though these results suggest children consider different sources of information in a non-socially-constructed domain, it is unknown whether they would do so in a socially constructed
domain. In the causal conditions, the amount of feedback (i.e., hearing the toy play music) children received about actions’ effects was easily quantifiable; however, it is less straightforward what would demonstrate positive or negative feedback about informants’ endorsements in an object labeling condition. Future studies could explore how to convey feedback in an object labeling condition—perhaps showing successful or unsuccessful communication achieved through using the label—and the effect it would have on children’s inferences.

In Experiment 1, we found that children’s endorsement of majority testimony varies by domain, but the flexibility with which children incorporate conflicting information within each of these domains has yet to be determined. Accepting majority testimony as universally informative could potentially mislead a learner; individual members of the majority opinion could be mistaken, or the majority opinion as a whole could be flawed (see Esser, 1998 for a review on groupthink). Future work can identify the cues used to identify a reliable or unreliable majority. Children could discount informant testimony for rational reasons—for instance, if an informant is unreliable or unknowledgeable—or for less rational reasons—for instance, bias against out-group members (Kinzler & Spelke, 2011).

Additionally, the nature of the beliefs underlying children’s endorsements has yet to be examined. Children may endorse majority testimony in the moment in order to conform to societal norms, but do not truly believe this testimony correct. In social psychology, this is called compliance. It is also possible that children internalize this new social knowledge and believe it to be true. Future studies could explore whether children are merely complying with social norms in similar object labeling tasks, or whether they internalize the majority’s testimony. Children could be asked to teach others the names of objects, or to recall objects’ novel labels in sessions hours or days later.

Overall, these results suggest that young children are discerning when considering others’ testimony; the extent to which they prefer majority testimony is dependent on task type. This work also adds to the growing body of literature that suggests children consider information from multiple sources to make rational inferences.
Chapter 4

Tracking informants’ knowledge sources

4.1 Introduction

Social learning is the cornerstone of human society. It has been proposed that our propensity for learning from others, rather than our intelligence or ingenuity, is responsible for our success as a species (Boyd and Richerson 1985; Boyd, Richerson, & Henrich, 2011). For children, social learning is especially beneficial—with little expertise and few life experiences, children can quickly acquire large amounts of new information from other people without spending time and effort to learn through trial-and-error. However, not all information from others is equally dependable. People can be ignorant, make mistakes, or even intentionally mislead others, and children may receive information from multiple people whose testimony conflicts. When learning from others, children must be discerning about whom they learn from, or else they may accept inaccurate information. Here we examine whether children are sensitive to the source and quality of informants’ knowledge.

A large body of literature about children’s epistemic trust has found that children do evaluate their informants’ knowledge. In particular, starting at around three years, children begin to consider informants’ past accuracy in answering questions (e.g., Koenig, Clément, & Harris 2004; Koenig & Harris, 2005; Pasquini et al., 2007). Evaluating informants’ previous accuracy can help children assess informant reliability, but in everyday situations, this information is not always available. We routinely learn from people we do not know well, or even from strangers. Corriveau, Fusaro, and Harris (2009) argue that in the absence of information about informants’ accuracy, children could instead evaluate consensus among different informants’ endorsements. In one study, they found that preschoolers prefer informants whose opinions received support from others (Fusaro & Harris, 2008), and endorse object labels given by a majority of informants (Corriveau, Fusaro, & Harris, 2009).

However, groups, like individuals, can provide unreliable information. To ensure the reliability of information, children must go beyond endorsing whatever the majority of people endorse: they must also consider the source of each person’s information. Information from a person with direct experience should be more valuable than information from a person with indirect experience. For instance, you might trust a friend’s restaurant recommendation more if she has actually been to the restaurant than if she’d merely heard positive reviews of it. Previous research has shown that, beginning at three years of age, children believe that direct experience is valuable. Children understand that visual experience can provide informants with knowledge (O’Neill, Astington, & Flavell, 1992; Pillow, 1989; Sodian & Wimmer, 1987), and prefer information from people who have seen something directly over information from people who have not (Povinelli & deBlois, 1992; Robinson, Champion, and Mitchell, 1999; but see Palmquist & Jaswal, 2012).

In the absence of informants with direct experience, informants with indirect experience (e.g. hearsay) from several sources may be more reliable than informants with indirect experience from a single source. For instance, you may be more likely to trust a restaurant recommendation from several friends who have each talked with a different person who ate at the restaurant and liked it, versus a recommendation from several friends who all talked to the same person.
In three experiments, we explore how preschoolers assess testimony from multiple informants who have different sources of knowledge. Children watch informants give testimony about which of two boxes contains the better option. In Experiments 1 and 2, equal numbers of informants endorse each box, but some informants have direct knowledge (visual access) about the boxes’ contents, whereas others have indirect knowledge (hearsay). In Experiment 3, again, equal numbers of informants endorse each box, but some informants receive their information from several different individuals, whereas others receive their information from the same individual.

4.2 Experiment 1: Direct knowledge vs. hearsay

In Experiment 1, participants watched as informants gave opinions about which of two boxes contained the better option. Equal numbers of informants endorsed each box, but one box was endorsed by informants who had looked in the boxes and had direct knowledge of what was inside, whereas the other box was endorsed by only one informant with direct knowledge while the other three received hearsay about which box was better.

4.2.1 Methods

4.2.1.1 Participants. Participants were 22 preschoolers (mean age = 4 years 1 month; range = 43 – 66 months; 12 female, 10 male). Participants were recruited in a major metropolitan city by mail and phone calls or from local preschools. An additional child was tested, but was excluded due to fussiness.

4.2.1.2 Materials. Materials included two black boxes, each of which contained a toy (a plastic cement truck or stuffed leopard) or a snack (Goldfish cracker or Froot Loop). Informants were eight paper dolls (four male, four female) designed and made available online by illustrator Kyle Hinton. (Permission from Mr. Hinton to use his illustrations in this study was obtained prior to data collection.) Dolls were roughly 7” tall and were glued to a wood block base.

4.2.1.3 Procedure. Children participated in two trials: a snack trial and a toy trial. Which trial was presented first was counterbalanced across participants. To begin the first trial, the experimenter showed the participant the two boxes and explained that each box contained a snack (in snack trials) or toy (in toy trials), but that she did not know what was inside (see Appendix A for full script). Then, the child watched as dolls gave opinions about which box contained the better option (“I think this [toy/snack] is better!”). Dolls came on stage one by one, and left the stage after stating their opinion.

Of the eight dolls, four endorsed one box and four endorsed the other. These two groups differed in the type of information they received before giving their opinions. In the direct group, all four dolls walked over and looked inside the lid of each box before giving their opinion. In the indirect group, only the first doll in the group looked inside the boxes before giving their opinion. This doll then crossed paths with the second doll, and the experimenter made indiscriminate whispering sounds to convey that the two dolls were speaking with one another. The second doll gave their opinion, and passed on their hearsay to a third doll. This process was repeated with the third doll, who passed the hearsay on to a fourth doll. Each group included equal numbers of male and female dolls (two male, two female). Group order (direct or indirect
first) and side of box endorsed by direct group (left or right) were counterbalanced across participants.

After all dolls gave opinions, the experimenter brought all eight dolls back on stage and placed them in front of the box they endorsed, then recapped the demonstration by gesturing to the dolls that endorsed each box. The experimenter then asked the child, “Would you like to have a snack? Which snack would you like to try?” (in snack trials) or “Would you like to play with a toy? Which toy would you like to play with?” (in toy trials). Once children gestured at or touched a box, they were presented with the object inside, and were offered an opportunity to play with the toy (in toy trials), or to eat the snack (in snack trials). The experimenter cleared all materials from the table, and proceeded to the second trial. The procedure of the second trial was identical to the first.

4.2.2 Results

A 2x3 Fisher’s exact test found that gender had no effect on outcome (choosing box endorsed by direct group in zero, one, or both trials). Fisher’s exact tests revealed that there was no effect of group order (direct group first versus indirect group first) or side of box endorsed by direct group (left or right).

Overall, children chose the box endorsed by the direct group in 32 of 44 responses (72%; 2-tailed binomial test, \( p < .01 \), .95CI = 59.46%, 85.99%). In the snack trial, 18 of 22 children (82%) chose the box endorsed by the direct group, and in the toy trial, 14 of those same 22 children (64%) chose the box endorsed by the direct group in the toy trial. (See Table 4.1.) A Fisher’s exact test comparing outcomes from the two trial types (toy and snack) revealed that there was no significant difference between trial types in the rate of selecting the box endorsed by the direct group.

Table 4.1.

<table>
<thead>
<tr>
<th>Children’s choices in Experiments 1 and 2.</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of trials in which children chose box endorsed by indirect group</td>
<td>Number of trials in which children chose box endorsed by direct group</td>
</tr>
<tr>
<td>Experiment 1 (snack)</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Experiment 1 (toy)</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Experiment 1 (total)</td>
<td>12</td>
<td>32</td>
</tr>
<tr>
<td>Experiment 2 (sticker)</td>
<td>6</td>
<td>16</td>
</tr>
</tbody>
</table>

4.3 Experiment 2: Replication with new domain

In Experiment 1, we found a slight domain difference in children’s responses. Though this difference was not significant, we were interested in investigating children’s responses in other domains. Experiment 2 replicates our results from Experiment 1, with one change: rather than using snacks or toys, we used stickers.
4.3.1 Methods

4.3.1.1 Participants. Participants were 22 preschoolers (mean age = 4 years 6 months; range = 43 – 61 months; 10 female, 12 male). Participants were recruited in a major metropolitan city by mail and phone calls or from local preschools. Two additional children were tested; one was excluded due to fussiness and the other because she did not speak English.

4.3.1.2 Materials. Materials were the same as in Experiment 1, except 1” x 1” animal stickers were used in place of snacks or toys.

4.3.1.3 Procedure. The procedure of Experiment 2 was identical to Experiment 1, except that there was only one trial instead of two.

4.3.2 Results
A Fisher’s exact test found no effects of gender, group order (direct first versus indirect first), or side of box endorsed by direct group (left or right).

Table 4.1 shows the main results of Experiment 2. In 16 of 22 trials (72%), children chose the box endorsed by the direct group over the box endorsed by the indirect group (two-tailed binomial test, \( p < .05, 95\% \text{CI} = 53.96\%, 91.49\% \)).

4.4 Experiment 3: Hearsay from multiple sources vs. one source

In Experiments 1 and 2, informants with direct knowledge were contrasted with those with indirect knowledge. Children’s choice of the box endorsed by those with direct knowledge could have been due to two different characteristics of the information the direct group received. The direct group’s knowledge was not only direct, but also independently discovered, whereas the indirect group’s knowledge was dependent on all previous informants’ knowledge (the first informant passed information to the second, which was passed down to the third and fourth informants). Experiment 3 was designed to eliminate this confound, so that all informants received indirect information while the independence of the sources varied.

As in Experiments 1 and 2, participants in Experiment 3 watched as informants gave opinions about which of two boxes contained the better option. Again, equal numbers of informants endorsed each box. One box was endorsed by informants who each received hearsay from a different source, whereas the other box was endorsed by informants who each received hearsay from the same source.

4.4.1 Methods

4.4.1.1 Participants. Participants were 24 preschoolers (mean age = 4 years 2 months; range = 40 – 62 months; 14 female, 10 male). Participants were recruited in a major metropolitan city by mail and phone calls or from local preschools. An additional three children were tested, but were excluded due to fussiness.

4.4.1.2 Materials. Like Experiments 1 and 2, materials included two black rectangular boxes, each of which contained a snack (Goldfish cracker or Froot Loop) or an animal sticker. Informants were ten paper dolls (five male, five female) designed and made available online by illustrator Kyle Hinton.
4.4.1.3 Procedure. Children participated in two trials: a snack trial and a sticker trial. Which trial was presented first was counterbalanced across participants. To begin the first trial, the experimenter showed the participant the two boxes and explained that each box contained a snack (in snack trials) or sticker (in sticker trials), but that she did not know what was inside. She then put the two boxes down on opposite sides of the table (see Appendix B for full script). Then, the child watched as the experimenter introduced four paper dolls (the source dolls), who each looked inside both of the boxes. These four dolls were then put in a separate area on one side of the demonstration table, where they were still visible to the child.

Then, six informant dolls came on stage one by one. Each encountered a source doll who was “taking a walk” away from the source doll area towards the informant doll. The informant doll whispered with this source doll. Of the six informant dolls, three endorsed one box, and three endorsed the other. These two groups differed in which source doll(s) they whispered with before giving their opinions. In the independent group, the three informant dolls received information by each whispering with their own, independent source doll. In the dependent group, all three informant dolls whispered with the same source doll. Group order (independent or dependent first) and side of box endorsed by independent group (left or right) were counterbalanced across all participants.

After each informant doll talked with a source doll, (s)he endorsed a box by saying to the source doll: “Oh, you think this box is better? Well, then, I think this box is better, too.” Then, the informant doll remained in front of the box they endorsed, while the source doll returned to the source doll area of the table. Once all six informant dolls had given opinions, the experimenter removed the source dolls from the table, and asked the child to choose a snack or sticker. Once children gestured at or touched a box, they were presented with the object inside. In sticker trials, children had the option of keeping the sticker they chose, and in the snack trials, children had the option of eating the snack they chose. The experimenter cleared all materials from the table, and proceeded to the second trial. The procedure of the second trial was identical to the first. [Source dolls in trial 1 were always informant dolls in trial 2, and genders of dolls in independent and dependent groups (2 males, 1 female vs. 2 females, 1 male) were also changed between trials.]

4.4.2 Results

A 2x3 Fisher’s exact test found that gender had no effect on outcome (choosing box endorsed by direct group in zero, one, or two trials). Fisher’s exact tests revealed that there was no effect of group order (independent group first versus dependent group first), or side of box endorsed by direct group (left or right).

Overall, children chose the box endorsed by the independent group in 32 of 48 total trials (67%; two-tailed binomial test, p < .03, .95CI =53.36%, 79.97%). (See Table 4.2.) In the snack trial, 15 of 24 children (63%) chose the box endorsed by the independent group, and in the sticker trial, 17 of those same 24 children (71%) chose the box endorsed by the independent group. A Fisher’s exact test comparing outcomes from the two trial types (sticker and snack) revealed that there was no significant difference between trial types in the rate of selecting the box endorsed by the independent group.
Table 4.2

*Children’s choices in Experiment 3.*

<table>
<thead>
<tr>
<th></th>
<th>Number of trials in which children chose box endorsed by dependent group</th>
<th>Number of trials in which children chose box endorsed by independent group</th>
<th>Percentage of trials in which children chose box endorsed by independent group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 3 (snack)</td>
<td>9</td>
<td>15</td>
<td>63%</td>
</tr>
<tr>
<td>Experiment 3 (sticker)</td>
<td>7</td>
<td>17</td>
<td>71%</td>
</tr>
<tr>
<td>Experiment 3 (total)</td>
<td>16</td>
<td>32</td>
<td>67%</td>
</tr>
</tbody>
</table>

4.5 Discussion

These three studies provide the first empirical evidence that preschoolers weigh multiple informants’ opinions using the quality of their knowledge source to assess the reliability of their testimony. We find that children favored opinions from informants who received knowledge directly (visual access) over informants who had received knowledge indirectly (hearsay from other informants). Additionally, when children encountered informants who received only hearsay, they favored opinions from informants who received hearsay from several different sources over informants who received hearsay from the same source.

This complements previous work that suggests children understand that not all testimony is equal. To succeed in this task, children had to evaluate opinions from multiple informants at once, and to consider each informant’s source knowledge. Furthermore, while previous studies asked children to make factual judgments (e.g., what’s in a box) from testimony, children in this study were asked to make a preferential choice based on others’ opinions. This suggests that children look to others for social information to inform their preferences, as well as facts.

It is clear from these studies that with equal numbers of informant endorsements, children consider knowledge from visual access more valuable than knowledge from hearsay, and hearsay from multiple sources more valuable than hearsay from a single source. Future studies could investigate the relative values children place on the number of informants endorsing an option versus the quality of those informants’ knowledge sources.

For example, if there were four informants with visual access who endorse box A, versus five with hearsay who endorse box B, would children favor the box chosen by informants with the higher quality source of knowledge (box A), or the box endorsed by the most informants (box B)? From a knowledge-acquisition perspective, additional informants in the indirect group do not provide any new information, since they repeat the same information provided by the first informant in the group.

While adults show some sensitivity to source of knowledge (Whalen, Buchsbaum, & Griffiths, 2013), they often do not treat repetition of information rationally. Instead, adults mistakenly attribute more credibility to ideas that are repeated, even by the same person (Foster et al., 2010; Weaver et al., 2007). Investigating the tipping point at which children perceive hearsay as equivalent to direct knowledge could provide an even richer understanding of children’s social learning biases.
Chapter 5

Conclusions

Together, these three sets of empirical studies illustrate the richness of preschoolers’ social inference skills. Children not only use statistical information to learn about other people, but can also use it to decide from whom to learn. Moreover, they can use other social or knowledge cues to guide their learning. This body of work suggests preschoolers’ inference abilities are even more sophisticated than previously thought: children are able to track statistical information but recognize that they must use it in conjunction with other contextual cues. Children’s ability to make inferences by combining their prior knowledge with additional information, like informants’ reliability, suggests that children are Bayesian learners. Previous work suggests children use Bayesian strategies for efficient learning about many related domains, like language (Goldwater, Griffiths, & Johnson, 2009; Regier & Gahl, 2004; Xu & Tenenbaum, 2007), causality (Bonawitz et al., 2011; Goodman, Baker, & Tenenbaum, 2009; Kushnir & Gopnik, 2007; Schulz et al., 2007; Sobel et al., 2004), and others’ beliefs (Goodman & Shafto, 2008; Lucas et al., 2014).

In chapter 2, children inferred an agent’s graded preferences just from observing his choice actions. From a short demonstration, these preschoolers inferred that the proportions, not raw frequencies, of the agent’s actions indicated his preferences, and were able to make inferences about objects that were not directly compared to one another during the demonstrations. Children made different inferences about the agent’s preferences for objects chosen 100%, 70%, and 20% of the time. Even in situations where there were no objects that were chosen in every instance they appeared (items chosen 83%, 63%, and 29% of the time), children demonstrated this ability. Though previous studies have shown that children can use choice actions to infer preferences, this was the first evidence that children may be implicitly tracking the statistics behind choice actions to understand the strengths of an agent’s preferences for multiple objects.

In chapter 3, children learned about a socially-determined domain (word meaning) from a consensus of informants, but used their own observations to guide their learning in a non-socially-determined domain (causal relationships). When the ability to make personal observations was removed from the socially-determined domain (i.e., children were not able to see whether an action made a toy play music), children defaulted to learning about the toy by imitating the action endorsed by a consensus of informants. This set of studies showed children’s ability to go beyond mere statistics, and to integrate additional context by taking into account the domain type and any additional evidence available when considering consensus testimony.

In chapter 4, children were again able to go beyond mere statistics. In three studies, children preferred the box endorsed by informants who had superior sources of knowledge. In experiments 1 and 2, children preferred the box endorsed by informants with direct knowledge (sight) of a box’s contents over informants who had indirect knowledge (hearsay). Additionally, in the absence of informants with direct knowledge, children preferred informants who received indirect knowledge from multiple independent sources as opposed to informants who received their indirect knowledge from a single source.

This line of research has important implications for developmental psychology in general. First, it adds to a growing body of work that suggests that children are implicitly tracking statistical information in their environment. In none of these studies were children explicitly
asked to note the number of times an agent chose an object or the number of people who chose an option, and it’s unlikely that children have received explicit instruction in their daily lives that these metrics are important. Rather, children have likely learned through their personal experience that choice actions, consensus, and knowledge type are useful indicators for others’ preferences or reliable informants. The richness of children’s implicit statistical skills further reduce the explanatory power of nativist theories that rely on poverty of the stimulus argument, and nudge the field towards theories that acknowledge that a combination of innate and learned skills contribute to children’s development.

Secondly, children incorporate other types of information into their social inferences. Implicit statistics are just one of these, but results from chapters 3 and 4 indicate that this is not the only factor children consider. Even when two options – the words or actions in chapter 3, or the boxes in chapter 4 – have the same number of endorsements, children tend to choose the option associated with more reliable informants, rather than choosing randomly between the two options.

Future studies can investigate the trade-off between statistical information and social cues, like informant reliability. If both types of information are available, children have the task of deciding which cues are most valuable, and in which contexts. In chapter 4, I suggest that a future direction for the study is to give children where the best option based on number of endorsements conflicts with the best option based on quality of knowledge (4 direct, vs. 5 hearsay). It is unknown how children weigh one option’s desirability against the number of informants who endorsed it. One possible approach would be to examine children’s choices in a set of simple, real-life example, where children are presented with two options and shown people’s “ratings” (simply positive or negative). Children’s average responses could tell us about what strategies they use when evaluating such information: for starters, do they distinguish between the options? If so, do they prefer options with more positive ratings? fewer negative ratings? the best ratio of positive to negative ratings?

Other studies could explore individual differences in development across the early years in a more systematic way. Children from a wide range of ages were pooled together to represent a general preschool population, and there was no demographic data collected about participants besides their gender and what languages were spoken in the home. Receptiveness to statistical or social information could vary depending on a host of individual and experiential factors, including race, exact age, social exposure (e.g., whether children attend daycare), family income, and parents’ education. For instance, there is some evidence that Asian-American children may be more susceptible to consensus testimony (Corriveau & Harris, 2010) and that socioeconomic status affects language ability (Hart & Risley, 2003; Fernald, Marchman, & Weisleder, 2012), memory (Noble, 2014), and brain function (Noble, Norman, & Farah, 2005; Noble et al., 2006; Noble, McCandliss, & Farah, 2007), which could have downstream effects in social cognition tasks. It has yet to be determined how well the samples included in these empirical studies represent children’s general cognitive abilities, and what differences may exist between these samples and other samples that represent a more diverse demographic.

Follow-up studies could also compare this set of empirical data with data from other populations. A comparison to adults’ behavior in these tasks could be useful; though we may have intuitions or make educated guesses from previous research about how adults would likely respond in these tasks, a dedicated study comparing adult behavior to preschoolers’ behavior could highlight the ways in which children’s inferences are “adult-like”, and what aspects of children’s decision-making will likely change in their transition through adolescence and
adulthood. Children’s data could also be compared to model predictions to explore what assumptions children may have when making social inference. For instance, in looking at children’s choices between two games, comparing children’s responses to a model with a prior assumption most games are fun versus a model with a prior assumption that games are neutral could tell us what children’s default assumptions about how fun games generally are.

In conclusion, the empirical studies in this dissertation contribute new insights about preschoolers’ statistical and social inference skills. These children are discerning young learners who use a rich array of cues to gain new knowledge about the people and things in their world.
References


Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in


