Young Children's Ability to Integrate Social and Numerical Information:

The Origins of Base-rate Neglect

by

Samantha Gualtieri

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor in Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Lili Ma Associate Professor
Supervisor	Stephanie Denison Associate Professor
Internal Member	Heather Henderson Professor
Internal Member	Jonathan Fugelsang Professor
Internal-external Member	John Turri Professor

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The seminal work of Kahneman and Tversky (1973) sparked an interest in the biases that govern decision-making, notably due to their findings on adults' tendency to neglect base-rate information (i.e., prior probability) when it conflicts with social information (e.g., a personality description, or testimony information). Though research over the past 45 years uncovered the conditions that lead to base-rate neglect, very little has investigated the origins of these biases. Young children can use base-rate and social information on their own, but their ability to integrate this information remains poorly understood. Do children show a preference for social information as soon as they are able to use it? Or, does their preference for social information develop over time? The current thesis explores 4- to 6-year-old children's ability to integrate base-rate and social information, providing insight into the origins of base-rate neglect.

In three projects, I assessed young children's ability to integrate base-rate and social information. A first project investigated children's use of base-rate information when it conflicted with individuating information (i.e., a personality description). Typically, adults classify an individual by evaluating how well the individuating information matches a stereotypical member of each social group, underusing prior base-rates of the groups in their decision. Using stereotypes familiar to young children, I presented them with an 8:2 base-rate of characters (e.g., 8 nice, 2 mean). One character was randomly selected from the group, with its membership unknown, and children were given a short personality description. By age 6, children performed similarly to adults and over relied on individuating information. Notably, 4-year-olds preferred base-rates more than the older age groups. I further explored these age differences in a second project that manipulated the quality of the base-rate and individuating information given. Six-year-olds' use of base-rates varied with manipulations depending on the

strength of the available individuating information. However, 4-year-olds consistently used base-rates across manipulations, even in situations where it would be reasonable to rely on the individuating information. Thus, children seem to initially show a preference for base-rate information and develop a bias toward individuating information by the age of 6, though attempt to reconcile individuating information with base-rates. A third project extended my findings to another type of social information. I presented children with testimony information from a witness that conflicted with base-rates. Rather than integrating information, adults typically use the witness' accuracy alone, thus neglecting base-rates. Here, 4- and 5-year-olds relied exclusively on an accurate witness, but they integrated information when the witness was less accurate. For young children, testimony from a witness is a strong cue, even stronger than stereotypical information.

With findings from the youngest age group tested to date, my dissertation provides evidence that heuristic strategies strengthen with development and vary depending on the type of social information provided. These findings highlight the importance of research on the role of age-related changes in experience and cognitive abilities in information integration strategies.

Acknowledgments

Over the last five years, I have received support and encouragement from a number of people. First and foremost, I would like to extend my deepest appreciation to my supervisor, Stephanie Denison, who made this work possible. Without your support, expertise, and advice, I would not have been able to complete this work. Thank you for helping me grow as a researcher and as a person.

I am also grateful for my fellow lab mates, past and present, for their support, feedback, and friendship. I am extremely thankful for the support of Liz Attisano, who believed in me more than I ever did. Thanks for always being there for me whenever I needed someone. I am also very grateful to Reem Tawfik, who always encouraged me to advocate for myself. I could not have had a better testing partner. And special thanks to Tiffany Doan, who was there with me throughout this whole journey. I would also like to extend my gratitude to all the research assistants who spent hours recruiting, testing, and coding these data, including, but not limited to: Nicole Asztalos, Bethany Nightingale, Alex Rett, Ashley Blayney-Hoffer, Emily Stonehouse, and Annika Voss. Special thanks to Annika for the extra hours of moral support and writing motivation.

I am extremely grateful to the principals, teachers, museum staff, parents, and junior scientists who made this research possible.

Finally, I would like to thank my family and friends for supporting me throughout this journey. Special thanks to Bryce for the much-needed emotional support over the last year. I am extremely thankful for the support and encouragement provided by my parents and brothers, who made sure I saw this through to the end. I dedicate this work to my mom, whose love and enthusiasm for children must have sparked something in me.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

On a daily basis, human decision-makers effortlessly make complex inferences based on relatively limited or incomplete information. Probabilistic information about prior likelihoods can help us predict future events. Subtle social cues allow us to make general inferences about one's personality and behaviour. Remembering how accurate someone was in the past allows us to determine if we should trust their account of an event. Though these processes are quite complex in and of themselves, we are almost never in situations where only one piece of information is relevant to our judgment. That is, we are often faced with the challenge of integrating and properly weighing, or choosing between, multiple pieces of information, which complicates our decisions even further.

In their seminal work, Kahneman and Tversky (1973; 1981) presented adult participants with various problems that tasked them with integrating base-rate (i.e., prior probability) and social information (e.g., a personality description, or testimony). Instead of properly weighing base-rate information in their responses, adults tended to underuse base-rates and relied more heavily on the conflicting piece of social information. Though much research has investigated the features of these problems that lead adults to neglect base-rates, very little has explored the developmental origins of this bias. I address this gap in the literature and explore the origins of these biases by investigating young children's ability to integrate base-rates and social information. I attempt to address questions that elucidate the developmental origins of heuristic reasoning, which provides important theoretical implications for dual process theories of cognition. Notably, it is unknown if children show a preference for social information, over base-rates, as soon as they are able to use social information, or if experience plays a role in strengthening these biases over time.

This chapter provides an overview of children's requisite skills and their ability to integrate numerical and social information, which must be established before addressing questions about the origins of base-rate neglect. I begin by discussing young children's ability to use base-rate information and improvements in probabilistic reasoning with development. Next, I review the literature on children's ability to use social information in their inferences and age-related improvements in these abilities. Finally, I discuss the paucity of research on children's ability to integrate these types of information and highlight important remaining questions that I explore empirically in this dissertation.

Children's Use of Statistical and Probabilistic Information

Our environments are often stochastic and unpredictable. We may never predict, with certainty, the outcome of an event, though probabilistic information can help us make a best guess. Moreover, the statistical structures of our environments are dynamic in the sense that they can change. For instance, if you are foraging for berries, you may find that particular trees yield more berries than others. However, the abundance of fruit may dwindle based on other factors, such as an increase in competitors foraging from those more prolific sources. By considering these factors, you can update your representation of the availability of berries from each tree, and can make changes to your foraging behaviour. Remarkably, humans are sensitive to statistical, proportional, and probabilistic data and use this information to detect structures in their environments, even within the first year of life. Tracking the co-occurrences of linguistic and visual events helps infants make sense of the "blooming, buzzing confusion" (James, 1890) intrinsic in their environments, which are often complex, dynamic, and contain large amounts of data and accompanying noise. This powerful statistical learning mechanism supports infants' development of language, recognition of objects, and acquisition of social behaviours (Saffran & Kirkham, 2018).

Probabilistic Reasoning in Infancy

Intuitions about probabilistic events emerge early in development. Findings from the violation-of-expectation (VOE) looking-time paradigm have established that probabilistic reasoning emerges within the first year of life, as infants look longer at events that they find novel or unexpected. Remarkably, 12-month-old infants are able to predict a future event based on a probabilistic distribution (Téglás, Girotto, Gonzalez, & Bonatti, 2007). Infants are shown a lottery machine device that contains three yellow objects and one blue object. The objects bounce around the machine to familiarize the infants with the distribution, and then the machine is briefly occluded while an object exits the device. Infants' looking times reveal that they expect one of the three yellow objects to exit, since they look longer when a blue object exits. Thus, infants use the distribution of objects to make inferences regarding the most probable outcome, expecting a yellow object to exit since a majority object is most likely to be sampled (Téglás et al., 2007). Moreover, when shown a large array of objects, infants as young as 6 months of age expect a sample to resemble its population (Denison, Reed, & Xu, 2013; Xu & Garcia, 2008).

Infants use probabilistic information to predict the outcome of a single random draw and explicitly use this information to inform their choices (Denison & Xu, 2010; Denison & Xu, 2014). Two objects (i.e., a pink lollipop adorned with stickers, and an undecorated black lollipop) are placed in front of the infant and the experimenter encourages them to crawl or walk toward an object. By choosing an object, the infant confirms their preference, and this object becomes the target object on subsequent test trials. Next, participants are shown two populations containing different mixtures of target and non-target objects (e.g., Population A contains 75%

target objects, Population B contains 25% target objects). After the infant is familiarized to the populations, the experimenter covers the containers and draws an object from each population. The coloured portion of the lollipop is covered by the experimenter's hand while she places it in an opaque cup. Once in the cup, the lollipop stick remains in view as a reminder to the infant that there is an object in the cup. After the experimenter has sampled the objects, she places the cups closer to the infant and encourages them to choose a sample. By 10 months of age, infants select the sample taken from the population with the higher proportion of target objects, suggesting that they use the distribution of target to non-target objects to inform their choices.

Probabilistic Reasoning in Childhood

Infants consistently use base-rate information to inform their predictions about future events, demonstrating their intuitive understanding of probability. How do these abilities develop throughout childhood? Piaget and Inhelder's seminal work on children's explicit intuitions about chance was the first empirical investigation of 5- to 12-year-old children's understanding of probability. These findings suggested children were unable to use probabilistic information in their decisions until quite late in childhood (Piaget & Inhelder, 1975). Piaget and Inhelder first familiarized children with a population of coloured beads by displaying them clearly on a tray. The beads were then placed in an opaque bag and shaken. The experimenter asked the child which colour they were more likely to get if they drew an object from the bag without looking. The youngest children in the sample did not systematically base their responses on the proportion of colours. Children around eight years of age used the proportion of colours to determine which object was more likely to be sampled, though it was not until age 11 that children understood how to update the distribution once objects were sampled without replacement (Bryant & Nunes, 2012; Yost, Siegel, & Andrews, 1962).

However, subsequent research provides a different picture of young children's understanding of probability, suggesting that Piaget and Inhelder's methods hindered their performance. This method placed high verbal comprehension demands on the youngest children, who may have had difficulty interpreting the test question. Children's performance was also contingent on their productive verbal abilities, because their responses were coded as correct only if they explicitly mentioned the proportion of objects. More recent variations on this design, using age-appropriate test questions and less stringent scoring criteria (i.e., coding a response as correct if a child indicates the more probable colour), have shown that young children use base-rate information to predict outcomes, indicating that even 3- and 4-year-olds use basic proportional comparisons to inform their decisions (Denison, Konopczynski, Garcia, & Xu, 2006; Denison, Bonawitz, Gopnik, & Griffiths, 2013; Gualtieri & Denison, 2019; Yost et al., 1962).

In concordance with Piaget and Inhelder's findings, recent studies have shown that probabilistic reasoning abilities improve with age (Falk, Yudilevich-Assouline, & Elstein, 2012; Girotto & Gonzalez, 2008; Girotto, Fontanari, Gonzalez, Vallortigara, & Blaye, 2016; Gualtieri & Denison, 2019; O'Grady & Xu, 2019). Importantly, these age differences occur in the absence of formal schooling of probabilistic concepts, suggesting that children's ability to extract and compute proportions naturally becomes more precise over the course of development. For instance, children's use of shortcuts in their probabilistic reasoning decreases with development, with older children primarily opting to compute and compare proportions of objects in lieu of solely relying on absolute quantities, which does not always produce correct responses (Bryant & Nunes, 2012; Falk et al., 2012; Girotto et al., 2016; O'Grady & Xu, 2019). Children's ability to make more complex comparisons with closer relative likelihoods (i.e., comparing proportions that are closer together, which are more difficult to visually discriminate) becomes more precise with age (Falk et al., 2012; O'Grady & Xu, 2019). More sophisticated reasoning abilities, such as integrating base-rate information with additional case-specific evidence, also improves with development (Girotto & Gonzalez, 2008). In these problems, children are shown four squares and four circles that are then placed into an opaque bag. All of the squares are black, while one circle is black and the remaining three circles are white. When asked which colour is more likely to be sampled, children as young as five correctly report that black is the more probable colour. The researcher then reaches into the bag to sample an object and reports that they believe they are holding a circle. Children's ability to update their response to reflect the more probable colour of circles improves with age.

In sum, infants and young children are proficient in using probabilistic information in their decisions, fine-tuning this ability with age. Over the course of development, children are able to make more complex comparisons and even consider additional information when making these inferences, including preferences and goals (Kushnir, Xu, & Wellman, 2010; Ma & Xu, 2011). From this, one may wonder why adults neglect base-rate information in their judgments. If humans become more adept at using probabilistic information, integrating base-rate information into one's decisions should be effortless by adulthood. However, it is important to note that adults' base-rate neglect often occurs when base-rates are presented in tandem with competing social information. In the next section, I discuss how our proclivity to use different types of social information develops during childhood.

Children's Use of Social Information

Although statistical and probabilistic data from the environment underpins much of infants' and children's learning, social input is equally critical in children's acquisition of numerous concepts (Gelman, 2009; Vygotski, 1929). The types of social information children use in their inferences and the various ways in which information is socially transmitted to young learners contribute to the multifaceted nature of children's social cognition. In general, humans exhibit fairly stable traits and behaviours over time, which allows us to make general inferences about a social partner's personality, goals, and attitudes (Jones, 1979; Trope & Higgins, 1993). However, pre-existing knowledge and beliefs from our culture can affect these inferences. That is, instead of tracking the behaviour, traits, and proclivities of every person we meet, we tend to categorize others based on certain features they possess. This leads to stereotyped inferences constructed from our representation of the social category or group an individual belongs to (Macrae & Bodenhausen, 2001). These concepts are cultural in nature, and are thus learned via communication within our social groups (Over & McCall, 2018). Transmitting knowledge between individuals is a defining characteristic of human cognition and communication, crucial to our acquisition of various concepts that cannot be learned through direct observation, such as various scientific and cultural concepts (Csibra & Gergely, 2009; Harris, Koenig, Corriveau, & Jaswal, 2018; Harris & Koenig, 2006). For instance, children's own observations about the shape of the Earth provide them with data that suggests the Earth is flat. To correct their understanding, children incorporate testimony from others into their representation. Concepts that are important to a child's culture, such as religion and the afterlife, are also acquired via testimony information due to the abstract nature of these phenomena (Harris & Koenig, 2006).

Children's Social Inferences

Trait-based Inferences.

Much of our world is social in nature, and our interactions with others are at the crux of human social cognition. Trait attributions play an important role in adult social cognition. The internal, stable nature of traits allows adults to make inferences about others' thoughts, goals, and proclivities, which are crucial to their predictions and expectations about individuals (Jones, 1979; Trope & Higgins, 1993). The ease with which Western adults use trait information in their inferences is illustrated by the fundamental attribution error, wherein adults readily explain behaviours in terms of internal traits and characteristics, and underestimate the role of context and situational factors in behaviour (Jones & Harris, 1967).

Although adults readily use trait information in their inferences, the development of children's understanding of traits is quite nuanced. By around the age of 4, young children are proficient in using social information to generate trait-based inferences. Young children use trait labels to make inferences about one's behaviour, and, conversely, use behaviour to make inferences about one's traits (Boseovski & Lee, 2006; Boseovski, Chiu, & Marcovitch, 2013; Gelman & Heyman, 1999; Gonzalez, Zosuls, & Ruble, 2010; Heyman & Gelman, 2000; Liu, Gelman, & Wellman, 2007). With age, children generate trait-based inferences from fewer instances of trait-consistent behaviour. In particular, 3-year-olds often require multiple examples of behaviour before they can make an inference, though 5- and 6-year-olds can make an inference after one example (Boseovski & Lee, 2006; Boseovski et al., 2013).

Further, young children have trouble using behavioural information to predict future, trait-consistent, behaviours, and make stronger inferences when trait labels are provided instead of trait-relevant behaviours (Boseovski et al., 2013; Liu et al., 2007). While children are developing an understanding of traits and their role in predicting consistent behaviour, the use of labels highlights their stable nature (Gelman & Heyman, 1999; Gonzalez et al., 2010; Liu et al., 2007). That is, the language used to describe a characteristic can facilitate children's expectations about its stability over time. If a novel characteristic is described in terms of a noun (e.g., “Rose is a carrot-eater”), young children believe that the characteristic is more stable and consistent over time than if it was described using a verbal-predicate (e.g., “Rose eats carrots whenever she can”; Gelman & Heyman, 1999). Moreover, young children are more likely to make trait-based inferences if trait information is described categorically rather than continuously. Adults and older children show an appreciation for the continuous nature of traits, recognizing that different people possess varying degrees of a particular characteristic (e.g., some people are a little bit shy, while others are very shy). Young children are more adept at making trait-based inferences when the information is presented categorically (e.g., shy versus outgoing) rather than continuously (e.g., a little bit shy or very shy), while they are developing their understanding of the dimensional nature of traits (Gonzalez et al., 2010). Thus, children first recognize trait labels and apply them categorically when making inferences. With development, children’s inferences become more sophisticated as they appreciate the dimensional nature of traits (Gelman & Heyman, 1999; Gonzalez et al., 2010; Liu et al., 2007).

Finally, it should be noted that young children’s trait attributions are more advanced than simple evaluations of character. Although young children’s trait-based inferences correlate with their evaluative ratings of an individual as “good” or “bad” (e.g., a nice person is rated as “good”, a mean or selfish person is rated as “bad”), children are able to make trait-based inferences that are not mediated by their global evaluations. For instance, children tend to rate shyness as a “neutral” trait (i.e., people who are shy are not simply “good” or “bad”), but they can still make trait-based inferences about a shy character’s behaviour (Liu et al., 2007). Thus, although children are still developing their understanding of traits, they recognize that these are separate characteristics from evaluations of general goodness or badness.

In sum, young children are able to use trait information in their inferences, though this ability becomes more sophisticated over the course of development. Young children show an early appreciation for trait labels, readily drawing inferences about traits from relevant behaviours, and vice versa. With age and experience, children need less information to draw these inferences and they develop a more adult-like understanding of the dimensional nature of traits. The ability to draw inferences based on traits and behaviours is pivotal to navigating social environments and interacting with others. However, instead of tracking individual traits, abilities, and behaviours for every person they encounter, adults and children form representations of social groups that inform them about the "typical" features their culture expects each member to possess. Thus, we often make inferences about an individual's personality and behaviour based on our representation of a social group. Although this tendency to categorize others can facilitate more efficient navigation of social environments, it often leads to biased inferences that are based on very limited information about an individual. Next, I discuss the use of stereotype information and representations of social groups during early childhood.

Stereotype-based Inferences.

Our social cognition is largely shaped by our culture. Notably, we form culturally-based representations about social groups that allow us to quickly categorize others based on particular aspects of their identity, such as their gender, religion, ethnicity, or occupation. Via enculturation, we create a representation of the social category that includes the characteristics we believe, and expect, members of each group to possess. Our acquisition of social categories across a variety of contexts suggest that we are predisposed to categorizing others. These tendencies emerge early in development and play a significant role in our efficient navigation of social environments (Bigler & Liben, 2007; Over & McCall, 2018; Shutts, 2013; Shutts, 2015).

The trade-off of this efficiency is that these representations leave little room for recognizing the differences among individuals within a social group (Macrae & Bodenhausen, 2001; Pettigrew & Tropp, 2006).

Although attitudes and beliefs about various social groups are acquired early in childhood (Over & McCall, 2018), gender is a particularly salient social category for young children (Arthur, Bigler, Liben, Gelman, & Ruble, 2008; Bigler & Liben, 2007; Shutts, 2013; Shutts, 2015). Because gender can often be perceptually discriminated by certain aspects of one's appearance (e.g., hair, makeup, clothing), children's classification of category members becomes more practiced with age, which in turn strengthens their tendency to categorize members automatically. Moreover, through the use of gendered pronouns, gender is a social category that is often highlighted in speech. From referring to students as "boys and girls", to asking students to use different bathrooms based on gender, aspects of children's environments provide clues that gender is an important social grouping, which may lead children to track the distribution of gender in their social environments. For instance, young children in early child-care settings may notice that most, if not all, of their teachers are female. Children may also notice that certain activities, such as dance class, contain a gender imbalance among students, and that certain toys are preferred more by a particular gender, such as dolls or trucks. By noticing the imbalance within these contexts, children's attention is directed to the gender of individuals, thus making it a salient feature of these environments (Arthur et al., 2008; Bigler & Liben, 2007). Finally, children's entertainment (e.g., books, television shows, movies) reinforces cultural gender stereotypes in their depictions of male and female characters (Over & McCall, 2018).

Children use gender information to inform their behaviour and social preferences early in childhood. Gender information first affects preschoolers' social preferences, as three-year-olds often report favouritism toward members of their own gender (Martin & Ruble, 2004; Shutts, 2013; Shutts, Roben, & Spelke, 2013; Shutts, 2015). Notably, children's preference for members of their own gender emerges across a variety of cultures (e.g., North America, Africa, Europe, South America), whereas other preferences, such as those based on race, tend to emerge later in development and are more culturally determined (Shutts, 2013; Shutts, 2015). Following this initial bias in children's own preferences, gender information impacts children's inferences about social affiliations. Four- to six-year-old children believe other children are more likely to have friends of the same gender rather than the opposite gender, and their tendency to believe friendship affiliations are based on gender increases with age during this period (Arthur et al., 2008; Shutts et al., 2013).

Thus, young children use gender information to make inferences about social affiliations. However, inferences about social affiliation do not reveal when children acquire cultural beliefs related to gender roles and stereotypes, and how these beliefs impact their social cognition. Children possess beliefs about gender quite early in development, with the toddler and preschool years critical to their acquisition of gender stereotypes (Martin & Ruble, 2004; Trautner et al., 2005). By the preschool years, children possess stereotypes regarding gender-typical traits, behaviours, roles, and occupations, and they build on this knowledge during the early childhood years (Edelbrock & Sugawara, 1978; Kuhn, Nash, & Brucken, 1978; Martin & Little, 1990). By 3 years of age, children reference gender stereotypes when justifying other children's preferences, though they rarely mention their gender when explaining their own preferences (Arthur et al., 2008; Eisenberg, Murray, & Hite, 1982). Without any prior knowledge of a child's toy preferences, children use gender to infer that other children prefer to play with toys that accord with gender stereotypes. Throughout early childhood, children acquire more gender-related information and apply it to their inferences about others, rigidly adhering to these stereotypes up until the age of 7, when they begin to view gender as a more flexible feature of one's identity. That is, although children are increasingly exposed to gender stereotypes with development, they recognize that gender stereotypes are not fixed and do not perfectly predict behaviour (Martin & Ruble, 2004; Trautner et al., 2005).

Thus, gender serves as an important distinction in young children's social environments. Readily forming representations of social categories aids in children's navigation of their social world and affects the ease with which they draw inferences about others based on limited information. Children rely on different sources of information to form and update their beliefs about social categories. When forming their representations, children attend to statistical regularities that provide them with clues about how often a social group typically engages in a particular behaviour (Riggs, 2019). However, communication with others is an important source of information for young children, providing them with information about a wide range of topics, including social categories, cultural knowledge and traditions, and scientific concepts. Children readily form representations about social categories they have not encountered, using information gleaned from others' testimony as the basis of their representation (Lane, Conder, & Rottman, 2019; Over & McCall, 2018). In the next section, I outline research on children's ability to use testimony information in their learning.

Children's Social Learning

Together with perception and memory, testimony information provides human learners with crucial knowledge about their environments (Harris et al., 2018; Over & McCall, 2018). Children receive much social input in the form of testimony information, that is, via verbal assertions intended for the transmission of credible information (Gelman, 2009; Harris &

Koenig, 2006). Testimony information complements information gleaned through direct observations and is critical to the acquisition of scientific, cultural, and social concepts that cannot be acquired through first-hand experience (Harris et al., 2018; Harris & Koenig, 2006). In turn, this allows listeners to effortlessly update representations based on information that they otherwise did not encounter through their own percepts (Gelman, 2009).

From quite an early age, children glean explicit information from others and rely on their testimony as a source of information. Children often direct their questions to adults, recognizing that adults are a reliable source of information particularly when they encounter something that they cannot reconcile with their current beliefs (Harris & Koenig, 2006). However, young children exhibit healthy skepticism and do not simply accept every assertion, especially if the informant has provided them with inaccurate information in the past (Mills, 2013). One experimental paradigm demonstrates young children's proficiency in discerning prior accuracy to determine if they should learn from an informant (Koenig & Harris, 2005; Pasquini, Corriveau, Koenig, & Harris, 2007). Children are familiarized to two speakers who take turns naming objects. Notably, the objects during this phase are familiar to young children, allowing them to evaluate the speaker's accuracy using their own knowledge of object labels. The two speakers are seated at a table and an object is placed between them (e.g., a ball). One of the speakers correctly labels the object (e.g., labels the ball "a ball"), while the other provides an incorrect label (e.g., labels the ball "a shoe"). Children are presented with a number of these trials, in which they establish that one speaker is accurate at naming objects, while the other is inaccurate. After the familiarization trials, participants complete test trials, in which they are asked to judge the accuracy of the speakers and indicate which speaker they would seek information from when learning about a novel object. Children's ability to discern and generalize accuracy information improves with age due to advances in children's ability to track the proportion of correct and incorrect responses produced by a speaker (Pasquini et al., 2007). Three-year-old children have no trouble trusting a speaker who is always correct, though they distrust a speaker as soon as they make a single error, even if the contrasting speaker is consistently incorrect. For instance, 3-year-olds have trouble determining who to trust when they are presented a choice between a speaker who made one error and a speaker who made errors on all four trials. Children's ability to track the frequency of correct and incorrect responses made by a single speaker improves with age. By the age of 4, children track the proportion of correct and incorrect responses produced by a speaker, readily comparing this information between speakers to determine who to trust (Koenig & Harris, 2005; Pasquini et al., 2007).

In sum, young children exhibit a healthy skepticism when learning from others, appropriately considering the accuracy of an informant before accepting any assertions they have made. This ability becomes more fine-tuned with age, as children become more proficient in weighing past accuracy when deciding who to seek information from. In this section, I reviewed the ways in which young children navigate their social environments, from tracking information about individuals and social groups, to representing social groups via categories, and by learning from others. The complexity of children's social cognition highlights the ecological importance of considering multiple pieces of information in their decisions. In the next section, I turn to this important aspect of children's decision-making: their ability to weigh and integrate information.

Children's Ability to Integrate Information

In our daily lives, we are seldom in situations where only one piece of information is relevant to a decision, so we are accustomed to weighing and integrating multiple pieces of information. For instance, we readily consult different sources before we decide to try a new restaurant. We track the proportion of five-star to one-star reviews to get an overall estimate of others' experiences. We can also read the content of each review to get a better sense of specific factors that may play a part in our decision, such as the quality of food, service, and price. That is, a restaurant might have a high overall rating, but we may notice a number of negative reviews mentioning very slow service. By integrating these pieces of information according to our perception of their relative importance, we can make more informed decisions.

But how does our ability to integrate information develop? Children exhibit quite sophisticated reasoning abilities early in development, yet in much of the aforementioned literature, children are presented with one piece of information or one source of knowledge and are not required to consider multiple pieces of information in concert. Thus, many questions remain regarding children's ability integrate different sources of evidence. In their seminal work, Kahneman and Tversky documented adults' tendency to overweigh social information in lieu of base-rates (Kahneman & Tversky, 1973; Kahneman, 2011; Tversky & Kahneman, 1981). Although adults' underuse of base-rates has been well-documented over the last 40 years, very little research explores the development of this bias. Do children integrate numerical and social information in their decisions, or do they prefer to rely on a specific type of information? And how does this weighing change over the course of development? In this section, I review the relatively small body of research investigating young children's ability to integrate information.

Remarkably, 4-year-olds integrate testimony information with causal frequency information observed first-hand (Bridgers, Buchsbaum, Seiver, Griffiths, & Gopnik, 2016). Children are shown a blicket detector and two different shaped blocks, and are told that one of the blocks activates the machine. Children are either given testimony from a confident informant who believes that they know which block activates the machine, or they are given testimony from an informant who claims to be guessing. After receiving the testimony information, children observe a sequence of trials in which the *other* block more successfully activates the machine probabilistically (e.g., 66% of trials) or deterministically (e.g., 100% of trials). Thus, children observe evidence that conflicts with the informant's testimony. Depending on the informant's confidence and the statistical sequence observed, children weigh the testimony and base-rate information in their judgments. Children who observe a probabilistic sequence of events rely more on the confident informant's testimony, though children who observe a deterministic sequence rely more on the causal frequency information observed first-hand.

Though this provides evidence that young children can integrate numerical and social information, this paradigm was not designed to be analogous to the classic base-rate neglect problems presented to adults. Children's first-hand observations of causal evidence in a statistical sequence may be more salient to their decisions than a base-rate depicting a population of objects. Although some researchers have examined children's use of base-rate and social information in adaptations of the classic problems, issues with their designs make it difficult to draw conclusions from this work. In these experiments, 6- to 12-year-olds were presented with a set of child-friendly problems inspired by the classic adult paradigm. Surprisingly, the 6- and 7-year-old children in the sample relied more on base-rate information, leading to the somewhat counterintuitive conclusion that young children were "more rational" than older children and adults, as they experienced less base-rate neglect (Davidson, 1995; Jacobs & Potenza, 1991).

However, recent work has suggested that these findings were due to children's lack of familiarity with the stereotype information presented (De Neys & Vanderputte, 2011; Stanovich, West, & Toplak, 2011). That is, 6- and 7-year-old children were less familiar with the group stereotypes used in the problems than the older children, and thus relied on the base-rate information due to their inexperience with these social groups (e.g., children were presented with cheerleader and band member stereotypes, which are likely unfamiliar to early school-aged children). To investigate this claim, 5- and 8-year-olds were presented with a set of problems that manipulated the familiarity of the stereotype information (De Neys & Vanderputte, 2011). Prior to data collection, group stereotypes were presented to a sample of children to establish two sets of problems that contained stereotypes that 5-year-olds were familiar or unfamiliar with. Using this problem set, the classic task was adapted for children as a card game. Cards contained a drawing of a person on one side (e.g., a girl), and a drawing of an object associated with a stereotype on the other side (e.g., a doll). In each problem, children were presented with a set of ten cards that depicted a 9:1 base-rate of individuals (e.g., nine boys and one girl). The experimenter shuffled the cards and placed them into an opaque bag. The experimenter then drew one card from the bag and showed the child the side with the object. Children were asked to indicate which character they thought was on the other side of the card. In general, 5-year-olds relied more on base-rates when presented with unfamiliar stereotypes, though they relied more on individuating information when presented with familiar stereotypes. From this, the authors concluded that children will use stereotypes and neglect base-rates if they are familiar with the presented stereotypes. I probe the validity of this interpretation closely in this dissertation.

The Current Experiments

The aim of this dissertation is to explore children's ability to integrate social and numerical information in problems adapted from the classic adult work on base-rate neglect. I developed versions of the adult problems to ensure they were suitable for use with preschool age children. Because children's ability to use base-rate information was historically thought to emerge later in development, researchers have yet to present child-friendly versions of these problems to this age group. Thus, I explored children's ability to integrate information in the youngest age group tested to date. To ensure that children were able to follow my tasks, I tested their baseline ability to use each piece of information separately before presenting them with problems that included both pieces of information together. This design feature allowed me to explore how aspects of various problems affect children's base-rate use at different ages, and how salient different types of social information are to young children's judgments.

Specifically, Chapter 2 explores 3- to 6-year-old children's ability to integrate information in a child-friendly version of the lawyer-engineer problem. In this problem, adults often over rely on case-specific, individuating information when it conflicts with base-rate information about a population (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). Chapter 3 further explores these age differences between 4- and 6-year-olds using additional manipulations of the lawyer-engineer problem. Chapter 4 examines 4- and 5-year-old children's ability to integrate information in a child-friendly version of the taxi-cab problem. In this problem, adults often over rely on testimony information when it conflicts with base-rate information (Bar-Hillel, 1980; Tversky & Kahneman, 1981).

Chapter 2: The Development of the Representativeness Heuristic in Young Children

A version of this chapter appears in:

Gualtieri, S., & Denison, S. (2018). The development of the representativeness heuristic in young children. *Journal of Experimental Child Psychology*, *174*, 60-76.

Cognitive psychologists have a long-standing interest in the biases that affect decision-making. In their seminal work, Kahneman and Tversky established numerous cases of base-rate neglect. For example, in the lawyer-engineer problem, participants read a personality description randomly selected from a sample of lawyers and engineers. The description was of a conservative man who enjoyed puzzles and did not care for social issues. Participants were then asked how likely it was that the man was an engineer, as opposed to a lawyer. Importantly, participants were either told that the group from which he was sampled included 70 lawyers and 30 engineers, or that it included 30 lawyers and 70 engineers. Despite these differences in base-rates, participants in both conditions estimated that the man was an engineer at nearly identical levels. That is, people neglected base-rate information (i.e., the number of lawyers and engineers in the sample) and relied more heavily on individuating information (i.e., the personality description, which fit their representation of a typical engineer). This is termed the representativeness heuristic (Kahneman & Tversky, 1973) and it can lead to base-rate neglect and biased judgments.

Since the initial publication of this work, many have investigated why adults neglect base-rates (see Kahneman, 2011, for a review), but few have investigated when this bias develops. Most research on the development of base-rate neglect has focused on children over the age of five, as this was when basic probabilistic reasoning was thought to emerge. However, we now know that preschoolers, infants, and even non-human primates can use proportional information in their inductive inferences (Denison et al., 2006; Denison & Xu, 2014; Kushnir et al., 2010; Rakoczy et al., 2014; Téglás et al., 2007; Xu & Garcia, 2008). This early emergence of probabilistic reasoning provides an opportunity to investigate the development of heuristics and reasoning biases, particularly those involving base-rates, in children younger than those previously tested. Therefore, one main goal of the current experiments is to examine the age at which children begin to neglect base-rates in favor of applying representativeness by examining the youngest age group tested to date, 3- to 6-year-old children.

Because I am testing children younger than those previously studied, it is important to establish whether children at each age can use each piece of information on its own to make predictive inferences. Some previous investigations of the representativeness heuristic in young children did not provide baseline measures of children's ability to use base-rate and individuating information separately. This limits the conclusions that can be drawn from this work. In particular, previous work often overlooked young children's knowledge of characteristic information, which is essential to generating a response based on the representativeness heuristic. For instance, although Jacobs and Potenza (1991) and Davidson (1995) found that 6- and 7-year-old children provided responses in line with base-rate information more often than older children and adults, these younger participants were presented with characteristic group information they were likely unfamiliar with (e.g., cheerleader and band member stereotypes). Thus, their increased use of base-rates likely arose from a lack of category information, rather than an ability to override a heuristic response with a normative one (Stanovich et al., 2011). This important problem was highlighted in an investigation of 5- and 8-year-old children's responses to base-rate problems that used both familiar and unfamiliar group information (De Neys & Vanderputte, 2011). The younger children provided normative responses only when they were unfamiliar with the presented stereotypes (and thus only had base-rates to rely on). This demonstrates the importance of establishing whether children have the relevant social information to provide a normative or a heuristic response in a particular paradigm. Therefore, in Experiment 1, I examine preschoolers' use of base-rate and individuating information separately, to provide a foundation for interpreting their responses when this information is provided in tandem in Experiment 2.

Important variants of the lawyer-engineer problem, which manipulate the relevance of the social information provided, have revealed how readily adults use individuating information in inferences (Kahneman & Tversky, 1973). In a typical experiment, participants are given individuating information relevant to their classification. Relying on individuating information is reasonable in this context to an extent, because the information fits one of the group's characteristics. Further, the pragmatics of the experimental situation might implore people to use individuating information more strongly than in every day reasoning (Schwarz, Strack, Hilton, & Naderer, 1991). Investigating participants' use of base-rates when given individuating information *irrelevant* to classification sheds light on how strongly they are committed to relying on it and how sensitive they are to the pragmatics of the experimental situation. In a variant of the lawyer-engineer problem, adults were once again given a description of an individual randomly selected from 70 lawyers and 30 engineers (Kahneman & Tversky, 1973). However, this individual was described as motivated, successful, and well-liked by colleagues. That is, participants were given individuating information that could describe either a lawyer or an engineer, in addition to the skewed base-rate. In this case, participants should certainly rely on base-rate information in their estimates and state that there is a 70% chance the person is a lawyer, because the individuating information is uninformative. However, adults rate the likelihood of this person being an engineer at approximately 50%, neglecting base-rates and attempting to rely on uninformative individuating information in their response. This reveals a particularly strong bias in adult cognition.

Previous developmental studies have not manipulated the relevance of individuating information, leaving unknown whether children's biases are as extensive as adults, and whether this changes throughout the early childhood years. Therefore, a second major goal of this work is to examine the pervasiveness of base-rate neglect and the representativeness heuristic in development by manipulating the relevance of individuating information. In Experiment 2, I present some children with individuating information that lacks relevance, as in the case described above. If children show sensitivity to base-rates in this context and only use *relevant* individuating information, this would indicate that early in development the bias is less pervasive and is deployed more conservatively than in adult reasoning. However, if they perform similarly to adults and neglect base-rates even when presented with irrelevant individuating information, then children's bias toward individuating information would be more extensive than previously known. In another variant of the lawyer-engineer problem, adults are not given any individuating information (Kahneman & Tversky, 1973). Experiment 2 will also include a condition in which children are provided no additional information about the individual case. Again, such a condition has not been explored in children, but adults typically default to base-rates in these cases. This condition is informative because it allows us to see whether simply placing children in a context in which characteristic group information is highlighted throughout a narrative affects their classifications and results in any base-rate neglect.

The Experimental Approach

To facilitate comparison between children's responses and the typical adult findings, children in the present experiments were given problems in conditions based on Kahneman and Tversky's (1973) original studies. Across problems, children aged 3 to 6 were presented with traits (i.e., nice and mean) and stereotypes (i.e., gendered toy preferences) that children at these ages are likely to be familiar with (Arthur et al., 2008; Edelbrock & Sugawara, 1978; Eisenberg et al., 1982; Heyman & Gelman, 2000; Liu et al., 2007; Serbin & Sprafkin, 1986; Shutts, 2013; Trautner et al., 2005). I included a trait-based problem to mimic the adult trait-based portions of the occupational stereotypes, such as a person being described as "careful, conservative, and ambitious" for an engineer (Kahneman & Tversky, 1973). I included trait labels and examples of behaviours in our experiments to ensure that children were given enough information to make inferences (Liu et al., 2007). I also included a gender-based problem because it evokes a more culturally embedded stereotype (which are more commonly studied than traits in the developmental literature on representativeness) and mimics parts of the adult problem that reference the kinds of activities people in different occupations might enjoy (e.g., "carpentry and math puzzles" for an engineer; Kahneman & Tversky, 1973). I reference activities that children view as being preferred by different genders, such as playing dress up or playing with trucks. Including one problem that relies on gender stereotypes and one problem that relies on simple trait-based inferences provides broad insight into the role of heuristic use in cognitive development. Previous research has focused heavily on stereotypes and has not examined traits. I chose an example of each kind to determine whether children's judgments and decisions vary depending on whether the problem posed involves these different types of groups or categories.

In the current experiments, young children's use of base-rate and individuating information in their inferences was examined using a novel paradigm. Consistent with previous developmental approaches to exploring base rate neglect, I use a forced-choice paradigm (Davidson, 1995; De Neys & Vanderputte, 2011; Jacobs & Potenza, 1991).

Experiment 1

To provide an independent assessment of whether children can use the current base-rate and individuating information in their inferences, Experiment 1 investigated 3- to 5-year-old children's use of base-rate and individuating information when presented alone.

Methods

Participants.

In both experiments, children were tested individually in lab, at their school or daycare, or at a local children's museum. The final analyses included 96 children, with 48 children in the base-rate condition, including 16 three-year-olds, 16 four-year-olds, and 16 five-year-olds (M_{age} = 4;4 [year;months], *range* = 3;0-5;11; 20 females). An additional child was tested and was excluded due to failing to correctly identify the majority group on both problems (see Procedure for details). Forty-eight children participated in the individuating condition, including 16 three-year-olds, 16 four-year-olds, and 16 five-year-olds (M_{age} = 4;6, *range* = 3;0-5;11; 24 females). An additional child was tested and was excluded due to non-compliance.

Procedure.

Participants in both conditions completed two problems. The experimental session was presented to the children on a laptop, using a PowerPoint presentation, which was narrated live by an experimenter. Children were told that they were going to hear about robots on another planet that is just like earth, and that they would answer some questions after.

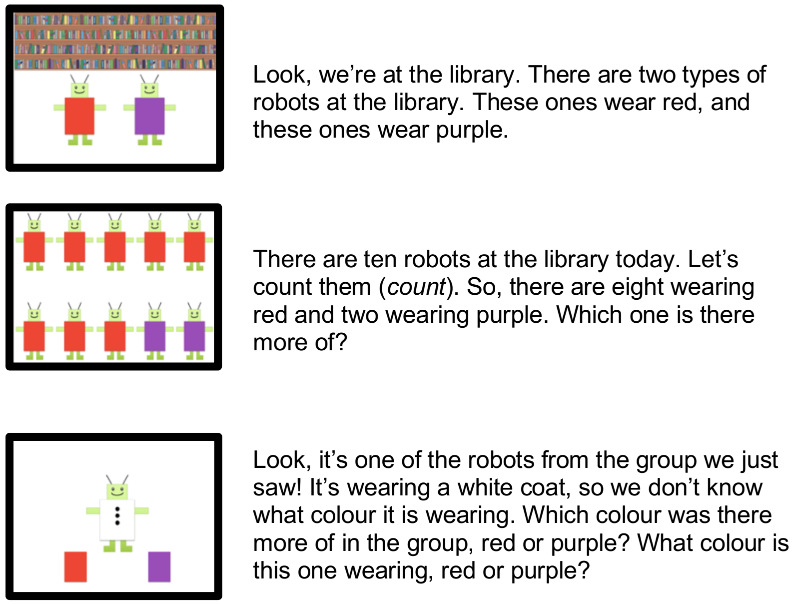

Figure 1. Colour problem in base-rate condition of Experiment 1.

Participants in the base-rate condition completed a colour and a shape problem, in counterbalanced order (see Figure 1 for a diagram of the base-rate condition procedure). Two problem-types (shape and colour) were used. In the colour problem, participants were introduced to two robots standing side by side at a library, one robot wearing red and another robot wearing purple. After establishing that the child could correctly point to a robot wearing red and one wearing purple, participants saw a group of ten robots at the library, eight wearing red, and two wearing purple. As a comprehension check, children were asked to indicate which type of robot there was more of. The experimenter either agreed or disagreed with the child's answer, stating that there were lots of robots wearing red and few wearing purple. Following this, children saw a single robot wearing a white coat, making its type unclear. Children were told that this robot was from the group of robots they just saw. Children were asked to recall which type of robot there was more of in the group. Participants were then asked to indicate which type of robot they thought the one wearing the white coat was. This procedure was also followed for the shape problem, including the comprehension check; however, participants saw one type of robot wearing hearts and another type of robot wearing stars. To ensure participants did not confuse the robots with those from the previous problem, the shape problem took place at a grocery store. Participants were excluded only if they did not correctly identify the majority group in the comprehension check for both problems.

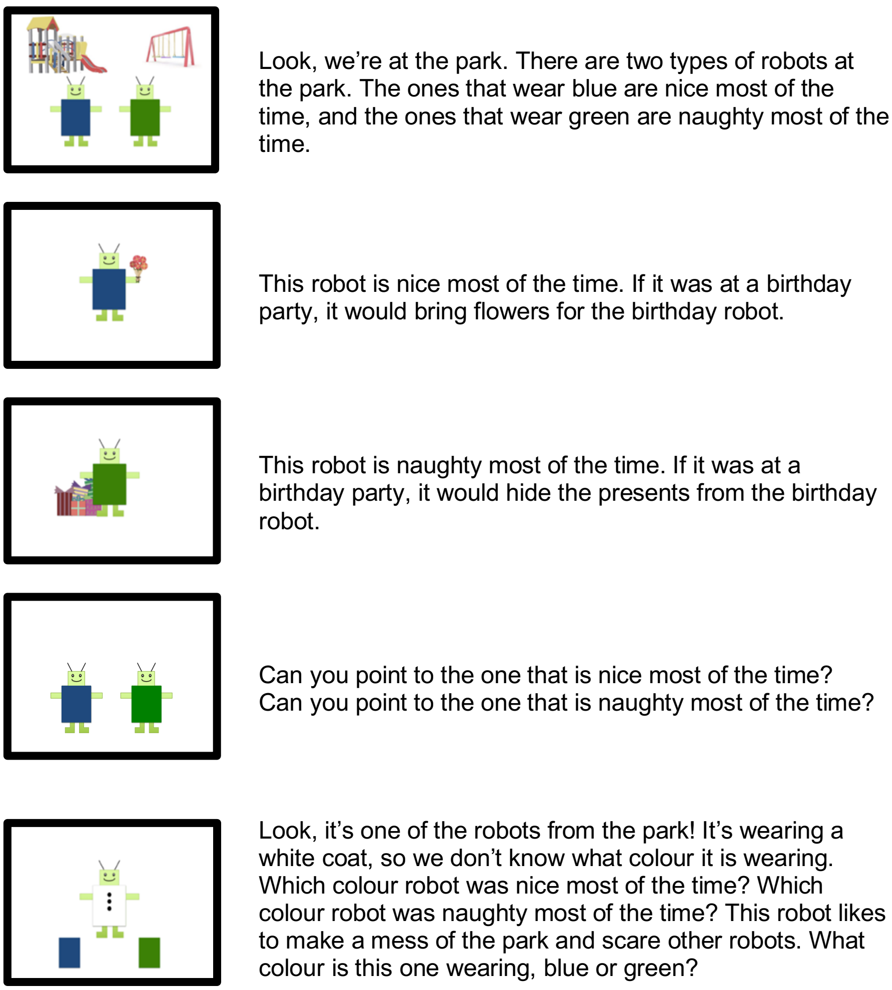

Figure 2. Trait problem in individuating condition of Experiment 1.

In the individuating condition, each participant completed a trait and a gender problem in counterbalanced order (see Figure 2 for a diagram of the individuating condition procedure, and Table 1 for group and individuating information). In the trait problem, following the initial introduction, participants then saw two robots standing side by side at a park, one robot wearing blue and another wearing green. They were told that most of the time the robots wearing blue were nice, and the ones wearing green were naughty. A visual manipulation, in which the characteristic group information mapped on to particular colours, was used so that the base-rates could be presented visually to the children in Experiment 2 when base-rate and group information are presented together[1].

Children were given characteristic group information in the form of examples of the robots' behaviours. These were included to ensure the salience of this information was equated with the salience (e.g., amount of time it takes to convey the information) of the base-rate information in the base-rate condition. Participants were asked to indicate which type of robot was nice most of the time, and which type of robot was naughty most of the time. Following this, children saw a single robot wearing a white coat, making its type unclear. Children were asked to recall which traits were associated with each colour. The experimenter then provided some individuating information about the robot wearing the white coat, which matched one of the traits. Half of the children heard a description of nice behaviour, while the other half of the children heard a description of naughty behaviour. Unlike the group information, the individuating information did not describe the robot as either "naughty" or "nice"; children were only told about the mystery robot's behaviours which were characteristic of a person who was either nice or naughty. This mimics the classic adult problems, in which participants are told the relevant categories leading up to the final classification (lawyers and engineers) but the category names are not mentioned in the individuating information. Participants were asked to indicate which type of robot they thought the one wearing the white coat was. This procedure was also followed for the gender problem. To ensure participants did not confuse the robots with those from the previous problem, the gender problem took place at a school.

[1] Presenting base-rates visually is standard practice in developmental studies of heuristics (De Neys & Vanderputte, 2011).

Table 1. Group and individuating information for each type in Experiments 1 and 2.

	Girl	*Boy*	*Nice*	*Mean*
Group information	This robot likes to play with toys girls like most of the time. See, it has a unicorn, a teddy bear, and a doll.	This robot likes to play with toys boys like most of the time. See, it has a dinosaur, a puppy, and a helicopter.	This robot is nice most of the time. If it was at a birthday party, it would bring flowers for the birthday robot.	This robot is naughty most of the time. If it was at a birthday party, it would hide the presents from the birthday robot.
Individuating information	This robot likes to play dress up and house.	This robot likes to play with trucks and trainsets.	This robot likes to clean up the park and help its friend play on the swing.	This robot likes to make a mess of the park and scare other robots.

In both conditions, the position of the two types of robots (e.g., green on the left and blue on the right) was counterbalanced in each problem. The experimenter always began by telling the child about the robot on the left. In the base-rate condition, the colour of the majority group was counterbalanced across both problems. The position of the majority group was also counterbalanced in both problems, as the majority was either presented as the first eight robots or the last eight robots in the group of ten. In the individuating condition, the individuating information given about the robot wearing the white coat was counterbalanced as representative of each type.

Results

Only sessions in the lab were video recorded. All videos from in-lab participants were secondary coded, with a 100% agreement rate between the experimenter and secondary coder (N=65). For children in the community, transcription of the data was reliability coded for half of the children to ensure agreement between the hand coding and electronic data. There was 100% agreement between these coders.

Base-rate condition.

Children received a score of 1 on each trial for selecting the group that corresponded to the majority. Preliminary analyses indicated there were no effects of problem order (i.e., first trial, second trial) on children's responses. Because each child completed two trials, 96 trials were included in the analyses. A Generalized Estimating Equation (GEE) binary logistic regression with age (3-year-olds, 4-year-olds, 5-year-olds) and problem type (shape, colour) as between-subjects factors revealed a significant age by problem type interaction, *Wald* $\chi^2(df = 2) = 9.57$, $p = .008$. The main effects of age, *Wald* $\chi^2(df = 2) = 3.48$, $p = .18$, and problem type, *Wald* $\chi^2(df = 1) = .05$, $p = .82$, were not significant. Pairwise comparisons provided additional insight on the age by problem type interaction. The performance of 3-year-olds on shape problems ($M = .44$, $SD = .51$) significantly differed from 3-year-olds on the colour problems ($M = .75$, $SD = 0.45$; $Mean_{Difference} = .31$, $p = .007$), 4-year-olds on the shape problems ($M = .88$, $SD = 0.34$; $Mean_{Difference} = .44$, $p = .003$), 5-year-olds on the shape problems ($M = .88$, $SD = 0.34$; $Mean_{Difference} = .44$, $p = .003$), and 5-year-olds on the colour problems ($M = .88$, $SD = 0.44$; $Mean_{Difference} = .44$, $p = .003$).

Recall that the goal of Experiment 1 was to establish baseline use of each type of information at each age. Therefore, I examined each age separately using additional GEE binary logistic regressions (see Figure 3 for overall means). Because three-year-old children's performance differed between colour and shape problems, I examined their performance on each problem type separately. Three-year-olds' performance on shape problems did not differ from chance, *Wald* $\chi^2(df = 1) = .25$, $p = .62$, and their performance on colour problems was marginally different from chance, *Wald* $\chi^2(df = 1) = 3.62$, $p = .06$. Next, I examined four- and five-year-olds' overall performance. Four-year-old, *Wald* $\chi^2(df = 1) = 6.12$, $p = .01$, and five-year-old children, *Wald* $\chi^2(df = 2) = 6.63$, $p = .01$, exceeded chance performance.

Individuating condition.

Children received a score of 1 on each trial for selecting the group that corresponded to the individuating information. Preliminary analyses indicated there were no effects of problem order (i.e., first trial, second trial) on children's responses. Because each child completed two trials, 96 trials were included in the analyses. A GEE binary logistic regression with age (3-year-olds, 4-year-olds, 5-year-olds) and problem type (gender, trait) as between-subjects factors revealed a significant main effect of age, *Wald* $\chi^2(df = 2) = 11.89$, $p = .003$. The main effect of problem type, *Wald* $\chi^2(df = 1) = .002$, $p = .96$, and the interaction between problem type and age, *Wald* $\chi^2(df = 2) = 1.44$, $p = .49$, were not significant. Pairwise comparisons revealed that the performance of 5-year-olds significantly differed from the performance of 3-year-olds ($Mean_{Difference} = -.44$, $p < .001$), and 4-year-olds ($Mean_{Difference} = -.25$, $p = .02$).

To further investigate development across age, children's performance at each age was examined separately using GEE binary logistic regressions (see Figure 3 for overall scores). Five-year-old children performed above chance, *Wald* $\chi^2(df = 1) = 14.73$, $p < .001$. However, 4-year-olds marginally exceeded chance, *Wald* $\chi^2(df = 1) = 3.01$, $p = .08$, and 3-year-olds did not exceed chance, *Wald* $\chi^2(df = 1) = 0$, $p = 1$.

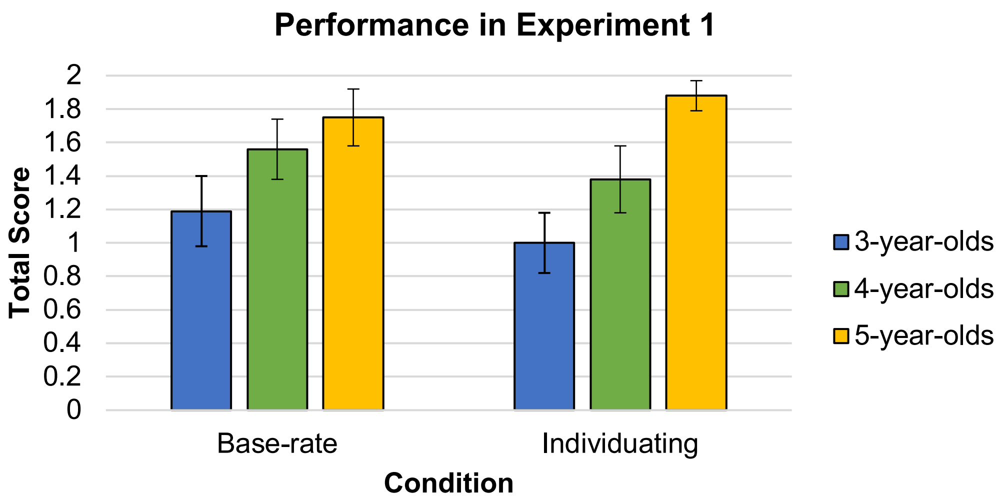

Figure 3. *Mean total scores for each age in Experiment 1. Error bars indicate standard error.*

Discussion

Experiment 1 established that children's abilities to use base-rate and individuating information develop between 3 and 5 years. When presented with base-rate information alone, 4- and 5-year-old children predicted group membership at above chance levels when colour and shape were used to depict base-rates, while 3-year-olds predicted group membership at above chance levels only when different colours were used. When presented with individuating information alone, 5-year-old children predicted group membership at above chance levels. Four-year-old children were marginally different from chance, using individuating information to predict group membership around 70% of the time. Because 16 children were tested in each age group, 4-year-olds' marginal performance may be due in part to a power issue. Although their performance was marginally significant, 4-year-olds seem to use individuating information in their inferences. Three-year-old children were unable to consistently use individuating information to make a predictive inference as they performed at chance in this task.

In sum, if a simple dimension like colour is used, children can produce a response based on base-rate information by the age of three. By the age of four, children can use individuating information in their inferences but this ability appears to be more fragile than in five-year-olds. This is pivotal to producing responses based on the representativeness heuristic, as well as those from integrating base-rates and individuating information.

Experiment 2

Having established children's developing abilities to use base-rate and individuating information separately, children were given both base-rate and individuating information together. The type of individuating information provided to participants was varied in four between-subjects conditions. Analogous to the adult literature, one group of children was presented with individuating information that conflicted with base-rate information, as in the classic problem. Another group of children was given problems that cued the same response (i.e., the two pieces of information did not conflict). These easier problems establish whether children at each age can hold both types of information in mind while following the story and whether presenting evidence that points in the same direction will affect inferences in an additive way (see De Neys & Vanderputte, 2011).

I included two additional experimental conditions that have been crucial in adult examinations of base-rate neglect but have not been presented to children. In the irrelevant condition, children were presented with base-rate information along with individuating information that was not useful to their classification. In the adult version of this problem, participants were given a description of an individual who was highly skilled, motivated, and well-liked by colleagues (Kahneman & Tversky, 1973). As this information could be indicative of either a lawyer or an engineer, it is rational for participants to default to base-rates in their estimates (i.e., produce a mean as a group that is statistically equal to 70%). However, participants in both base-rate conditions (30/70 and 70/30) believed there was about a 50%

chance that the individual was an engineer. Thus, underuse of base-rates in this condition suggests that participants are pulled away from base-rates by the presence of individuating information. Following this rationale, the irrelevant condition was included to gauge the pervasiveness of the bias in children. Therefore, children's performance in this condition will be compared to a prediction based on base-rates to see whether their classifications differ from this predicted level (i.e., 80%). If they fall significantly below the base-rate prediction (towards chance levels, as adults have done in previous work), then this suggests that they experience base-rate neglect and employ representativeness very strongly.

In the no individuating condition, children were taken through the same narrative and base-rate information as in the other conditions but were given no individuating information when it came time to make a classification. In the classic adult problem, participants were told that the individual was drawn from a base-rate of lawyers and engineers (70/30 or 30/70), though they were not given any specific individuating information about the person they were asked to classify. Adults correctly use the numerical information in their estimates and provide responses that reflect the base-rates; that is, they produced a mean as a group that was statistically equal to 70% (Kahneman & Tversky, 1973). Though similar to the base-rate condition in Experiment 1, this condition places children in a very different context. In Experiment 1, the robots wore different colours in an 8:2 ratio, but this information did not signify any characteristic information (i.e., gender or trait) associated with those groups. Unlike the base-rate condition of Experiment 1, characteristic group information was associated with each colour of robot throughout the narrative. However, at the time of classification, no individuating information was given. If children default to base-rates (as adults tend to do in this type of condition), this suggests that contexts in which social characteristics are present do not disrupt their base-rate use. If children do not default to base-rates, this suggests that the context alone is enough to make children neglect base-rates in their judgements, which would be indicative of an even stronger bias than in the irrelevant condition. Thus, children's responses in the no individuating condition will be compared to a predicted value corresponding to the base-rate of 80%.

To simplify the structure of the problems, only colours were used to depict base-rates. The trait and gender stereotypes from Experiment 1 were used because they provide an interesting test case: 4-year-olds appeared to be developing an ability to use these stereotypes in classification and 5-year-olds were very competent at making these inferences. To further investigate the development of base-rate neglect, 6-year-olds and adults were also tested in Experiment 2. In all problems, the normatively correct answer is to produce a response that takes base-rates into account.

Methods

Participants.

One-hundred and ninety-two children were included in the final analyses of Experiment 2. Forty-eight children participated in the conflict condition, including 16 four-year-olds, 16 five-year-olds, and 16 six-year-olds (M_{age} = 5;6 [year;months], *range* = 4;1-6;11; 21 females). An additional five children were tested and were excluded due to parental report of very low English language exposure (i.e., hearing English less than 50% of the time, a criterion that was set before data collection commenced, *n*=3) or experimenter error (*n*=2). Forty-eight children participated in the no conflict condition, 16 at each age (M_{age} = 5;6, *range* = 4;2-6;11; 20 females). An additional two children were tested and were excluded due to non-compliance. Forty-eight children participated in the irrelevant condition, 16 at each age (M_{age} = 5;6, *range* = 4;0-6;10; 31 females). An additional two children were tested and were excluded due to non-compliance.

Forty-eight children participated in the no individuating condition, 16 at each age (M_{age} = 5;5, *range* = 4;0-6;11; 26 females). An additional three children were tested and were excluded due to non-compliance (*n*=2) or failing to correctly identify the majority group in both problems (*n*=1).

Procedure.

Participants completed two problems, presented consecutively. As in Experiment 1, children completed a trait and a gender problem, using the same colour, group, and individuating information as before. Each child saw one colour pair and the associated group information on the first trial, and another colour pair and associated group information on the second trial. The instructions were the same as in Experiment 1; however, participants were specifically told that they would answer a question about one of the robots at the end.

Using the trait problem as an example, children were first introduced to the two types of robots and were told about the nice and mean traits that corresponded to each colour (see Figure 4 for an example of a trait conflict problem). Once again, they heard characteristic group information about the robots' behaviour. Next, participants were presented with the base-rate information: a group of ten robots, with eight wearing blue and two wearing green. Children were asked to indicate which type of robot there was more of at the park. Following this, children saw a single robot from the group wearing a white coat, making its type unclear. To ensure both pieces of information were equally salient to the child, the experimenter reminded the child of the base-rate and group information in a counterbalanced order. In each problem, base-rate and group information were explicitly mentioned twice to the child prior to hearing the individuating information about the robot. The experimenter then provided individuating information about the robot wearing the white coat. Participants were asked to classify the colour of the robot wearing the white coat.

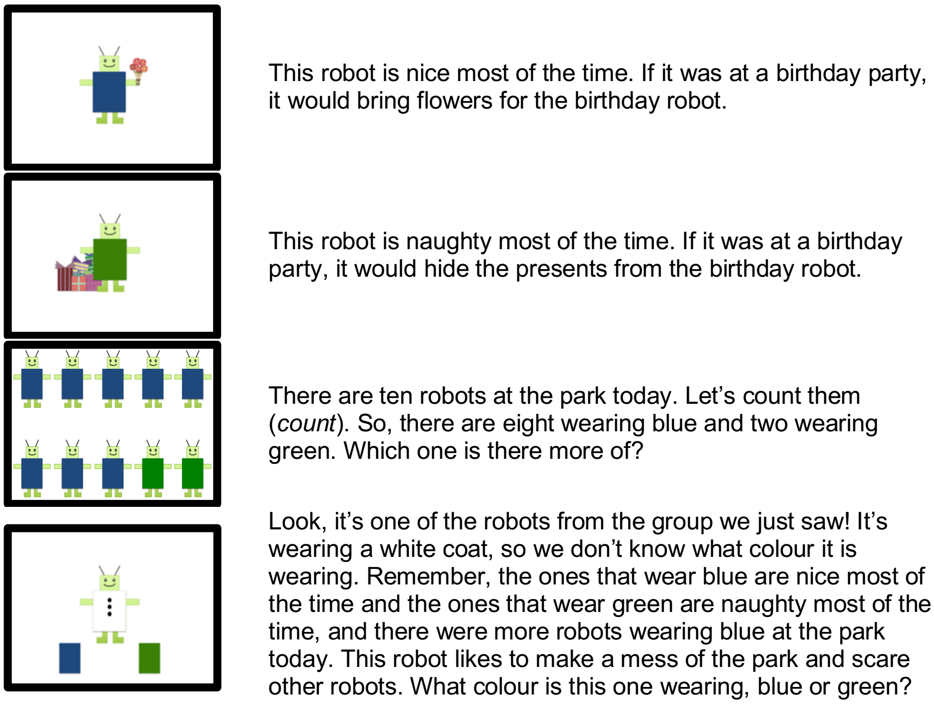

Figure 4. Sample of trait conflict problem in Experiment 2.

Children participated in one of four conditions, which determined the type of individuating information they received (see Table 2 for individuating information used in each condition). In the conflict condition, participants were told individuating information that conflicted with the base-rate. In the no conflict condition, participants were told individuating information that aligned with the base-rate. In the irrelevant condition, participants were told individuating information that was not useful to group classification. In the no individuating condition, participants were not given additional information about the robot. All participants were asked to classify the colour of the robot wearing the white coat.

All factors that were counterbalanced in Experiment 1 were counterbalanced here. The order of reminding about base-rate and group information was also counterbalanced. The group information associated with the robots in the conflict and no conflict conditions was counterbalanced (i.e., the type of robot the individuating information represented).

Table 2. Individuating information in Experiment 2.

Condition	*Girl*	*Boy*	*Nice*	*Mean*
Conflict and No conflict	This robot likes to play dress up and house.	This robot likes to play with trucks and trainsets.	This robot likes to clean up the park and help its friend play on the swing.	This robot likes to make a mess of the park and scare other robots.
Irrelevant	This robot likes to have a snack before recess and plays with its friends.		This robot likes to play on the slide and in the sand.	

Adult Participants.

To ensure that the stimuli elicited typical use of the representativeness heuristic in adults, undergraduate participants were given the same problems presented to children (*N*=32; *range* = 18- to 23-years-old; *female* = 20). Five additional participants were tested and replaced. One participant was replaced for indicating that they were familiar with the representativeness heuristic upon a manipulation check. That is, they stated that they were familiar with the representativeness heuristic and provided a reasonable description of it when they were explicitly asked at the end of the study. Participants were also replaced for changing their answers during the session, which they were instructed against doing (*n*=4). Four between-subjects conditions (conflict, no conflict, irrelevant, and no individuating), in which participants completed four problems of the same type, were included. For instance, a participant in the conflict condition completed four conflict problems. One story was used for the first two problems (e.g., trait), with one of the two robot types as the majority for each problem. The second story was used for the last two problems (e.g., gender). The experimenter provided the participant with the same group and individuating information told to children, except two additional pieces of irrelevant individuating information were added for adults. Only two problems were presented to children, one gender and one trait problem, so the same irrelevant individuating information was used regardless of the base-rate. Because adults completed two problems of each type (four in total), I created an additional piece of irrelevant individuating information for each problem. This ensured they did not hear the same information twice. For the second gender irrelevant problem, participants were told that the robot liked to have a sandwich at lunch and coloured with its friends. For the second trait irrelevant problem, participants were told that the robot liked to play on the monkey bars and on the swing. To facilitate comparison with child participants, the stories were presented using a PowerPoint presentation that was narrated live by an experimenter. Participants selected their choices to the experimenter's questions on a response sheet. Following the session, adults were asked whether they were familiar with the representativeness heuristic.

Results

All videos from in-lab participants were secondary coded (*N*=27), with a 100% agreement rate between the experimenter and secondary coder. Half of the transcription data for the community participants was also reliability coded. There was one disagreement, which was resolved via discussion. For each problem, responses were coded such that a score of 1 corresponded to a choice of the majority group and 0 corresponded to a choice of the minority group on each problem (i.e., problems were scored according to base-rate use). Because adults completed four problems and children completed two problems, I computed the percentage of responses in which each participant used the base-rate, after confirming via preliminary analyses that there were no effects of problem type (gender or trait) and no effects of counterbalancing.

Child Participants.

384 trials were included in the analyses, as each child completed two problems. A GEE binary logistic regression including age (4-year-olds, 5-year-olds, 6-year-olds) and condition (conflict, no conflict, irrelevant, no individuating), as between-subjects factors revealed a

significant interaction of age and condition, *Wald* $\chi^2(df = 6) = 14.1$, $p = .03$, and a main effect of condition, *Wald* $\chi^2(df = 3) = 48.73$, $p < .001$. The main effect of age was not significant, *Wald* $\chi^2(df = 2) = 4.45$, $p = .11$.

To explore the relationship between age and condition, I conducted follow-up GEE binary logistic regressions for each condition, including age as a between-subjects factor (see Figure 5). There was no main effect of age for the irrelevant, no individuating, or no conflict conditions. The age by condition interaction was driven by performance in the conflict condition, *Wald* $\chi^2(df = 2) = 13.70$, $p = .001$. Pairwise comparisons revealed that the performance of 6-year-olds differed from that of 4-year-olds ($Mean_{Difference} = .50$, $p < .001$), and 5-year-olds ($Mean_{Difference} = .25$, $p = .02$).

Moreover, I conducted binomial tests to explore children's performance in the irrelevant and no individuating conditions. Recall that these analyses are critical to examining the pervasiveness of base-rate neglect in children; adults typically do not default to base-rates in irrelevant conditions, though they correctly use base-rates in no individuating conditions. I compared children's mean responses at each age to the predicted value of 80%, which reflects correct use of the base-rate in both conditions. In the irrelevant condition, children at all ages produced responses that were significantly below the predicted value of 80% (4-year-olds: M = 50%, $SD = .5$, $p < .001$; 5-year-olds: M = 56%, $SD = .50$, $p = .002$; 6-year-olds: M = 53%, $SD = .51$, $p < .001$). In the no individuating condition, 4-year-olds' responses did not differ from the predicted value of 80% (M = 78%, $SD = .42$, $p = .47$), but 5- and 6-year-olds' responses marginally differed from it (5-year-olds: M = 56%, $SD = .50$, $p = .002$; 6-year-olds: M = 56%, $SD = .50$, $p = .002$).

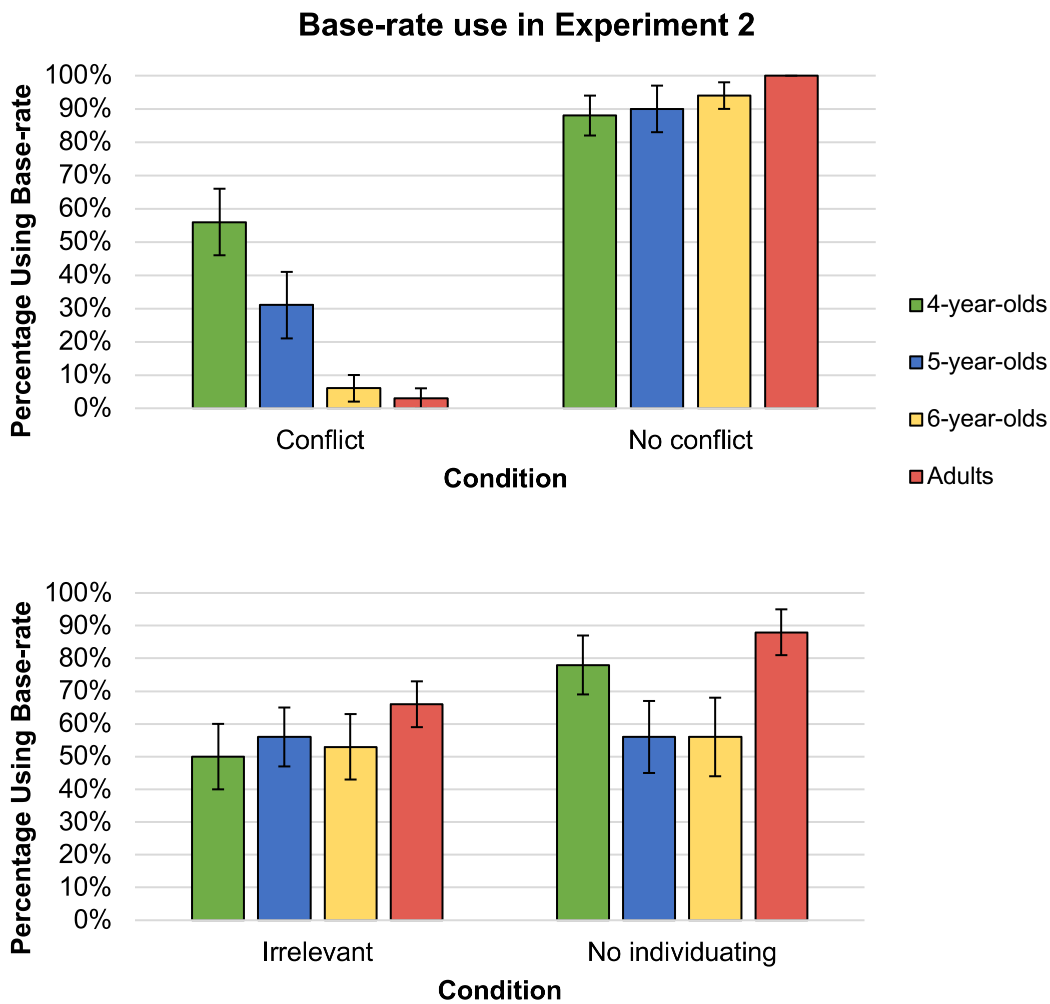

Figure 5. Mean percent choosing base-rate, averaged across trials for each age by condition in Experiment 2. Error bars indicate standard errors.

Adult Participants.

128 trials were included in the analyses because each adult participant completed four trials. A GEE linear regression, including condition as a between-subjects factor (conflict, no conflict, irrelevant, no individuating) revealed a significant effect of condition, *Wald* $\chi^2(df = 3) = 1133.51$, $p < .001$. Pairwise comparisons revealed that the conflict condition (*M*=3%, *SD*=.09) differed significantly from the no conflict (*M* = 100%, *SD* = 0; $Mean_{Difference}$= -.97, $p < .001$), irrelevant (*M* = 66%, *SD* = .19; $Mean_{Difference}$ = .62 $p < .001$), and no individuating (*M* = 88%, *SD* = .19; $Mean_{Difference}$ = -.84, $p < .001$) conditions. The irrelevant condition differed

significantly from the no conflict ($Mean_{Difference}$ = -.34, p < .001) and no individuating ($Mean_{Difference}$ = -.22, p = .01) conditions. The no conflict and no individuating conditions also significantly differed ($Mean_{Difference}$ = -.12, p = .05).

I conducted additional binomial tests to compare performance in the irrelevant and no individuating conditions to the predicted value of 80%. Adults' performance in the irrelevant condition was significantly different from the base-rate predicted value of 80% (M = 66%, SD = .19, p=.04), though performance in the no individuating condition was not significantly different from the predicted value (M = 88%, SD =.19, p = .20).

Discussion

Across conditions, I varied the relevance of the individuating information to examine its impact on children's use of base-rate information. Children's responses on the conflict problems revealed an interesting age trend: 4-year-olds gave significantly more base-rate-consistent responses, while 6-year-olds relied heavily on individuating information, suggesting greater base-rate neglect with age. Responses from 5-year-olds were in between those from the 4- and 6-year-olds, as they were more likely to neglect base-rates but not as often as the older children. In the no conflict condition, all age groups produced responses consistent with both base-rate and individuating information at high levels. Children's responses in the no conflict condition were more extreme than those in either of the conditions in Experiment 1. Thus, children may integrate base-rate and individuating information, as their responses appear to consider both pieces of information additively.

The conditions that manipulated relevance were particularly enlightening. At every age (including adulthood) participants did not default to base-rates with irrelevant individuating information, suggesting that the bias to rely on such information even when it is unwarranted is strong as soon as children begin to use the heuristic. That is, the presence of even irrelevant individuating information draws participants away from base-rate use. This suggests that the presentation of individuating information may impel participants to incorporate this information into their decision. This could be largely due to the role of pragmatics in these problems (Schwarz et al., 1991) or to a particularly strong tendency toward base-rate neglect that would also be present in real-world judgments. Future work examining the relationship between children's pragmatic reasoning abilities and base-rate neglect on this task could illuminate these roles.

In the no individuating condition, 4-year-old children and adults used base-rate information to predict group membership, providing group level responses in line with the base-rate predicted value of 80%. However, 5- and 6-year-old children did not use base-rates to the same extent, as they produced judgments that were marginally below the predicted value. These two conditions, tested with children for the first time, provide additional insight into the development of base-rate neglect. Four-year-olds (like adults) correctly used base-rates when given no additional information, but the presence of irrelevant individuating information pulled their responses away from base-rates (also like adults). It is surprising that older children in the sample did not produce responses consistent with base-rates in the no individuating condition. It is possible that merely presenting 5- and 6-year-old children with group information in the narrative led to base-rate neglect, as they may be particularly sensitive to group characteristics, or pragmatics, in their predictive inferences at these ages.

General Discussion

In two experiments, I explored children's use of base-rate and individuating information. The results provide a more complete picture of the development of heuristics and reasoning biases in children. Experiment 1 established the ages at which children can use base-rate and individuating information on their own to make inferences. This provided a foundation from which to interpret children's responses when given both types of information together in Experiment 2.

The current experiments make two broad novel contributions to our understanding of the use of the representativeness heuristic in young children. First, these experiments examine the representativeness heuristic and its prerequisite abilities in the youngest children tested to date, 3- and 4-year-olds. This revealed that, while 3-year-olds are capable of making inferences using only base-rate information, they do not consistently make inferences based on stereotype or trait information alone in this task. Thus, the current data suggest that further examinations of the representativeness heuristic might best begin with 4-year-old participants, or a more simplified design may be necessary with younger children. The experiments also revealed interesting insights into 4-year-olds' reasoning abilities; they could use both base-rate and individuating information alone, and they made stronger inferences when this information cued the same response (i.e., in the Experiment 2 no conflict condition). When given conflicting individuating information, 4-year-olds produced responses that were closer to the base-rates than those of older children and adults. This suggests that they potentially integrate this information in a way that produces more rational responses than do older decision-makers. This is consistent with other recent findings, which suggest that younger children may *benefit* from their lack of experience, allowing them to engage in less biased reasoning that is more sensitive to objective data, such as base-rates and other statistical information (Lucas, Bridgers, Griffiths, & Gopnik, 2014).

Second, manipulating the relevance of the individuating information, along with including children of three ages, gives novel insight into the pervasiveness of representativeness use and base-rate neglect in early childhood. When presented with irrelevant individuating information, 4-year-olds' responses were closer to base-rates than were older children's responses. Though this may suggest that younger children are less likely to neglect base-rates than are older children, even 4-year-olds did not recognize that irrelevant individuating information should be entirely ignored. Thus, by the age of four, children are headed toward weighing individuating information too heavily. Interestingly, unlike 5- and 6-year-olds, they did default to base-rates when no individuating information was provided, showing that the mere presence of trait or stereotype information in the narrative does not impede their base-rate use. Furthermore, 5-year-olds' responses in the conflict condition were not as extreme as 6-year-olds' responses. This suggests that 5-year-olds may weigh base-rates in their predictive inferences, even though they still show a bias, but 6-year-olds may almost entirely ignore them.

In fact, 6-year-olds' inferences were nearly identical to adults'. The one notable difference between these groups was their performance in the no individuating information condition, where adults, but not 5- or 6-year-olds, appropriately defaulted to base-rates. This performance in the no individuating condition was striking; children of all ages were skilled at using base-rate information alone in Experiment 1 when the task was presented in a context entirely devoid of social group information. One possibility is that for 5- and 6-year-olds, the mere presence of characteristic group information could trigger base-rate neglect in favor of attempts to rely on social information. Thus, it is possible that as heuristic use is becoming engrained in early development, children are particularly sensitive to conditions that typically signify its use. With additional practice in making predictions, decision-makers could start to apply heuristics in more appropriate contexts, though notably, even adults are pulled away from base-rates when presented with irrelevant individuating information.

In light of these age trends, the development of the representativeness heuristic is more nuanced than previously thought. Predictions based on previous work would have suggested that as soon as a child is familiar with a stereotype, they will immediately favor it nearly exclusively and neglect base-rates (De Neys & Vanderputte, 2011), but the current results suggest that this bias strengthens between 4 and 6 years of age, as children gain experience making social inferences based on characteristic information.

Limitations

One potential limitation of the current design was using the robots' shirt colours to depict the different groups. An alternative way of visually depicting the groups is to use a more fixed physical characteristic such as a biological marking, implying a more intrinsic and enduring difference across groups. Although depicting group differences visually through shirt colours may seem arbitrary, this design feature has some benefits over depicting the groups using a more fixed physical characteristic. Importantly, a fixed feature could imply to children that the traits and stereotypes associated with the feature are completely inflexible. Depicting traits and stereotypes in this manner could cause participants to think that these characteristics apply deterministically and this might subsequently cause them to think that the individuating information is perfectly diagnostic, making the base-rate information entirely irrelevant. Future studies could use designs that imply higher or lower levels of fixed correspondence between the group markings and the characteristics to assess whether these design choices impact children's use of base-rates and individuating information.

A second limitation of the current studies is the forced-choice method of responding. To ensure children of all ages could provide a response, I used a binary dependent variable. This differs from the more graded judgments provided by adults, which can provide insight on their ability to integrate social and numerical information. I return to this important point in Chapter 3.

Implications and Conclusions

In sum, quite soon after children reliably use base-rate and individuating information separately in their inferences, they favor individuating information when both pieces of information are presented together. Thus, when children have the available knowledge and cognitive skills required for the task, their decision-making is often governed by the more efficient process. As many previous studies have used individuating information without first examining the use of the individuating information at the relevant ages on its own, these results are the first to suggest that the representativeness heuristic is readily used by 5-year-olds and at similar levels to adults by age 6.

This chapter explored the development of the representativeness heuristic in young children. For the first time in a developmental investigation, 4- to 6-year-old children were presented with pivotal manipulations that establish the extent of their bias. These conditions illuminated interesting age differences between the youngest and oldest children in the sample. In light of these age differences, I explore how varying the strength of the base-rate and individuating information affects children's inferences across ages in Chapter 3.

Chapter 3: Examining Developmental Change in Young Children's Information Integration

The findings in Chapter 2 are the first to examine the pervasiveness of the representativeness heuristic in 4- to 6-year-old children. Within a period of two years, children's sensitivity toward base-rate information is overtaken by a preference for case-specific individuating information. That is, by the age of 6, young children show a bias for individuating information at similar rates as adults, though 4-year-olds tend to use base-rate information in their inferences. This is quite striking; 4-year-olds relied on individuating information when it was the only available information in the problem, though they did not show a bias toward this information when it conflicted with base-rates. These findings are in contrast to previous work suggesting that children show a preference for individuating information as soon as they possess the requisite trait and stereotype information (De Neys & Vanderputte, 2011). Due to the small body of work on the representativeness heuristic in young children, it is unclear what mechanisms contribute to the emergence and strengthening of a bias to neglect base-rates between the ages of 4 and 6.

It is possible that the emergence of the representativeness heuristic is due in part to a general feature of children's cognition: the development of children's intuitive theories about the role of personal characteristics in predicting others' thoughts, preferences, and behaviours. Like adults, children may have a strong prior belief about the importance of personal characteristics in predicting others' actions. However, this theory may not be fully developed until the age of 6. This is consistent with findings on the development of the fundamental attribution error in 4- and 6-year-old American children (Seiver, Gopnik, & Goodman, 2013). In Western cultures, adults often favor person-specific causes of behaviour over explanations that consider situational factors, which is known as the fundamental attribution error. Seiver et al. (2013) explored if children would correctly use covariation evidence to explain and predict a character's behaviour or if they would be biased toward more person-specific reasoning. While 4-year-olds relied on the covariation data in their predictions and did not show a strong a bias toward person-specific explanations, 6-year-old children showed a preference for person-specific explanations of behaviour over situational explanations, similar to adults in Western societies. Thus, as they gain experience navigating complex social environments situated within a Western society, it is possible that children are developing overarching beliefs about what types of information are useful for predicting others' actions, biasing them toward personal characteristics in general, as this is an efficient shortcut. As their intuitive theory strengthens, children weigh personal characteristics more highly and thus downgrade the importance of other information, such as base-rates or covariation data. This might give rise to the somewhat counter-intuitive data pattern revealed in Chapter 2, where younger children produced responses that were more in line with mathematically normative judgments, in which the base-rates influenced their responses. That is, younger children, whom are less entrenched in using traits and stereotypes to make inferences about behaviour, provide less biased and more rational judgments that are more aligned with the observed data. In the previous chapter, 6-year-olds showed a bias toward individuating information. Moreover, the inclusion of group information affected their inferences even when the problem did not contain individuating information, suggesting that the mere presence of a social scenario caused them to neglect base-rates. Conversely, 4-year-old children, who may not hold a strong prior belief regarding how individual characteristics should be used to make inferences about people, trended more toward the base-rate information in their responses. Thus, children use the more reliable base-rate information in their inferences while they are gaining strength in their theory about the role of personal characteristics in predicting behaviours.

Alternatively, the age differences between 4- and 6-year-old children may be due to 6-year-olds' increased experience in using *particular* stereotypes and traits in their inferences. Between 4 and 6 years of age, children gain more experience using information about traits to predict behaviours and require less information to make trait-based inferences. For instance, after observing a character take toys that belonged to other children, 3- and 4-year-olds predicted that the character would continue to take toys after observing four instances of taking, although 5- and 6-year-old children made this inference after observing one instance of taking (Boseovski & Lee, 2006; Boseovski et al., 2013). Moreover, children acquire more information about gender stereotypes as they attend to gender information in their social environments and are exposed to cultural ideals (Arthur et al., 2008; Bigler & Liben, 2007; Martin & Ruble, 2004; Shutts, 2013; Trautner et al., 2005). This is not to say that 4-year-olds' use of base-rates was due to a lack of familiarity with the specific trait information and gender stereotypes tested in Chapter 2; recall that 4-year-olds relied on the individuating information in their inferences when it was the only available information in Experiment 1. Rather, more experience and knowledge with these *specific* traits and cultural stereotypes may have led 6-year-olds to rely on the individuating information due to the ease with which they have come to use this information.

In Experiments 3 and 4, I presented 4- and 6-year-old children with another child-friendly version of the lawyer-engineer problem. This problem was similar to the problem presented to children in Chapter 2, however, it was modified to achieve two main theoretical goals and one methodological goal. First, I created two new problems that contained group information that signified a preference for an activity. This allowed for a conceptual replication of the key findings in Chapter 2. In lieu of the previous trait and stereotype information, which children likely conceptualize as dichotomous, I chose preferences that are not dichotomous. In the activity problem, participants were told a story about children in a class that could play baseball or make crafts during free time. In a second problem, participants were told a story about children in a class that could learn about space or wild animals during a trip to a museum. The preferences used in these problems are not as dichotomous as the trait and gender information used in Chapter 2; nice and mean are often viewed as opposite traits, and gender roles and stereotypes are categorized in a dichotomous fashion in Western cultures. The current problems, however, contained preferences that were not mutually exclusive. Though some individuals prefer sports over arts, or learning about a particular topic, it is not uncommon to partake in both activities even if one has a preference. Including problems that contain less stringent features allowed for a broader view on children's conceptual development.

Second, I aimed to further examine the age differences in children's information integration abilities. I manipulated the strength of the base-rate and individuating information to examine if its use was moderated by the quality and strength of the information it provided. Along with the manipulations of relevance in the previous experiments, presenting children with individuating information that is equally relevant, though varied in strength, provides further insight into the mechanisms underlying children's base-rate neglect. It is unclear from the previous experiments if 6-year-olds are attending to the quality of the individuating information, or if they are categorizing the individual based on the content of the individuating information without evaluating the quality of the information. Likewise, I also manipulated the base-rate information to examine children's ability to weigh this information in their judgements. From the previous experiments, it is unclear if 6-year-olds would ever factor base-rate information into their judgements, because they did not default to base-rates when no individuating information was available. This manipulation will also provide additional insight on the performance of 4-year-old children, who tend to closely represent the observed data in their inferences (Gualtieri & Denison, 2018; Seiver et al., 2013). Manipulating the strength (i.e., the skew) of the base-rate in this problem will allow me to examine if children at both ages are neglecting this information all together, or if they are attempting to reconcile the information provided by the base-rate and individuating information.

A final goal of the current experiment was methodological and exploratory in nature. In adult versions of the task, participants are asked to indicate the likelihood that the individual belongs to an occupation (e.g., they are asked to provide an estimate using a percent value that Jack is an engineer). In developing a child-friendly version of this task, I employed a forced-choice procedure suitable for children in this age group, which is consistent with all previous approaches (De Neys & Vanderputte, 2011; Gualtieri & Denison, 2018). Young children cannot provide estimates using percentage values and seldom provide relevant explanations for their behaviour and thought processes. However, using a binary dependent measure is limiting, and may be masking nuances in children's weighing of base-rates and individuating information. To combat this problem, researchers have created inference scores from children's binary responses and their estimates of confidence in their response and have found that this can provide a more complete picture of performance (Kalish, Kim, & Young, 2012). In Experiments 3 and 4, I continued to ask children to choose between two options when identifying an unknown individual, though, after children made their choice, I also asked them to estimate their confidence in their response using a 3-point scale. Creating inference scores from the product of these measures might provide insight the strength of children's beliefs about their choice.

Experiment 3

Experiment 3 tested 4- and 6-year-old children's use of individuating information when it conflicted with base-rate information. To explore the effects of individuating information quality on children's performance, children participated in one of two between-subjects conditions that varied the strength of individuating information provided.

Methods

Participants.

Children were tested individually at their schools or at a local museum. Sixty-four children were included in the final analyses, with 16 four- and 16 six-year-olds in each of the two conditions (see Table 3 for age and gender breakdown). An additional five children were tested and excluded for not finishing the task (*n* = 3), failing both comprehension checks (*n* = 1; see Procedure for details) and parental report of atypical development (*n* = 1).

Table 3. Age and gender breakdown per condition in Experiment 3.

	Mean age	Min age	Max age	Female
More individuating				
4-year-olds	54.89 months	51.69 months	58.76 months	6
6-year-olds	77.19 months	72.33 months	83.33 months	11
Less individuating				
4-year-olds	53.69 months	48.62 months	59.62 months	10
6-year-olds	78.86 months	74.07 months	83.76 months	11

Note: *n* = 16 per age, in each condition.

Materials and Procedure.

Participants heard two stories about children in a class at a school, narrated live by the experimenter using a PowerPoint presentation (see Figure 6 for an overview). In the activity story, participants were told that children in a class could make a craft or play baseball during free time. Participants were shown a base-rate of children who completed each activity that the experimenter counted aloud. The base-rate for each problem consisted of eight children completing one activity and two completing the other activity (e.g., eight children making a craft and two children playing baseball). After counting, participants were asked which activity was completed by more children. Depending on the participant's response, the experimenter agreed or disagreed with the child's answer and stated that there were more children who made a craft and less who played baseball. Following this, participants were told that one child in the class went home for lunch, and they were given case-specific individuating information about the child's traits and preferences that were more stereotypical of a child who, for example, is more likely to enjoy sports versus arts and crafts. The base-rate information always conflicted with this description, because the individuating information was representative of the minority group. More specifically, if a participant was shown a base-rate of eight children making a craft and two playing baseball, they were given information that corresponded to an interest in sports.

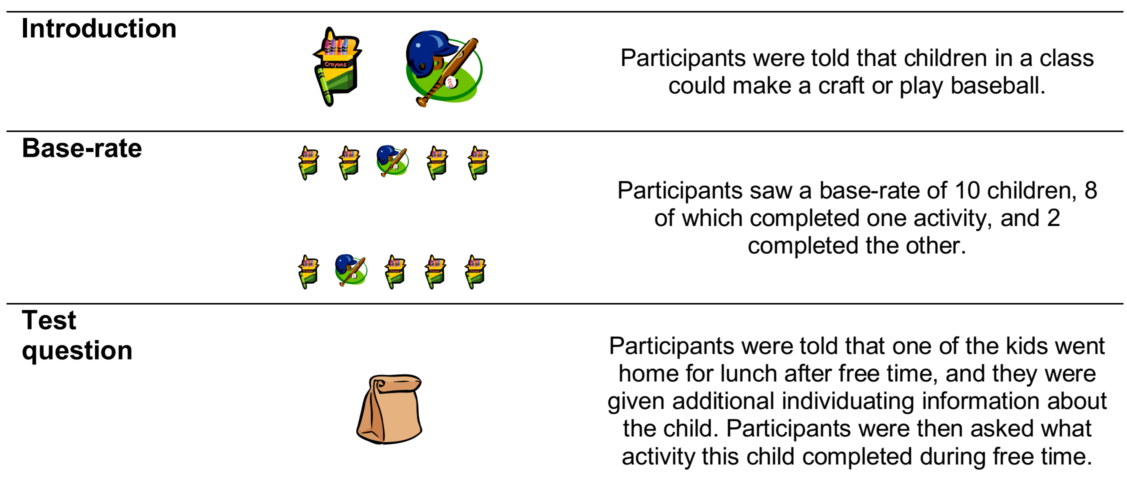

Figure 6. Overview of the procedure in Experiments 3 and 4.

I varied the amount of individuating information given to participants across two conditions. Children in the *more individuating* condition were given a lengthier description that was typical of the descriptions used in the classic studies with adults and the investigation in Chapter 2. Children in the *less individuating* condition were given a brief description of the child's preferences, which was still representative of one of the groups (see Table 4 for individuating information for each condition). After they were told the individuating information, the experimenter asked the participant to indicate which activity they thought the child completed earlier in the day. Once the child made their choice, the experimenter asked them to rate their confidence in their response (i.e., "Is it for sure that one, maybe that one, or are you guessing?") In the museum story, participants were told a story about a class on a field trip where the children chose between a wild animal exhibit and a space exhibit. Similar to the activity story, I used an 8:2 base-rate and participants were given individuating information that corresponded with preferences that seemed more in line with the minority group. The order of the stories presented (i.e., activity story first, museum story first), the majority group (and, thus, the individuating information used), the placement of the base-rate objects, and the group introduced first were counterbalanced across participants.

Table 4. Individuating information provided to children in Experiments 3 and 4.

	Activity story		Museum story	
Condition	Craft	Baseball	Wild animal	Space
More individuating	This kid likes to paint and play with play-doh. They like to make cool things and use their imagination.	This kid likes to play soccer and tag. They like to run around at recess and ride their bike to school.	This kid wanted a cat on their cake. They want to take care of animals when they grow up, and they read books about sharks.	This kid wanted a rocket on their cake. They want to be an astronaut when they grow up, and they read books about aliens.
Less individuating	This kid likes to paint and play with play-doh.	This kid likes to play soccer and tag.	This kid wanted a cat on their cake.	This kid wanted a rocket on their cake.

Results

Children received a score of 1 if they selected the group that corresponded to the individuating information. Though selecting the majority group indicated by the base-rate would be a closer approximation to a normative response, I coded the data in this way because I manipulated the strength of the individuating information (which also facilitates comparison to the results of Experiment 4, when the base-rate is 50/50). 128 trials were included in the final analyses, because each child completed two trials (see Figure 7 for overall means). Preliminary analyses indicated there were no effects of counterbalancing or story type (i.e., problem order: first, second; story: activity, museum) on children's responses. A Generalized Estimating Equation (GEE) binary logistic regression with age (4-year-olds, 6-year-olds), condition (more individuating, less individuating) and the interaction between age and condition revealed a significant main effect of condition, *Wald* $\chi^2(df = 1) = 9.99$, $p = .002$, and age, *Wald* $\chi^2(df = 1) = 5.18$, $p = .02$. A significant age by condition interaction was also observed, *Wald* $\chi^2(df = 1) = 5.18$, $p = .02$.

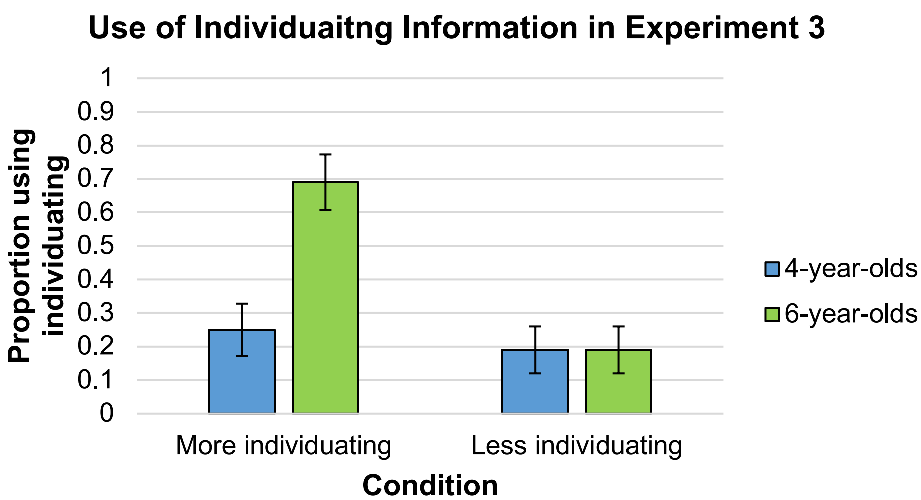

Figure 7. Proportion of children's responses that used individuating information in Experiment 3 (lower scores indicate use of base-rate information).

Pairwise comparisons provided additional insight on the age by condition interaction. There was a significant difference between 6-year-olds in the more condition ($M = .69$, $SD = 0.47$) and 4-year-olds in the more condition ($M = .25$, $SD = .44$; $Mean_{Difference} = -.44$, $p < .001$) and 4-year-olds in the less condition ($M = .19$, $SD = .40$; $Mean_{Difference} = -.50$, $p < .001$). There was also a significant difference between the performance of 6-year-olds in the more condition and those in the less condition ($M = .19$, $SD = .40$; $Mean_{Difference} = -.50$, $p < .001$).

I created inference scores from children's responses and confidence scores. Children who used the individuating information and were confident received a score of 6, whereas those who used the base-rate and were confident received a score of 1. This created a 6-point scale. A GEE linear regression with age (4-year-olds, 6-year-olds), condition (more individuating, less individuating), and their interaction revealed significant main effects of child's age, $Wald\ \chi^2(df = 1) = 4.60$, $p = .03$, and condition, $Wald\ \chi^2(df = 1) = 16.36$, $p < .001$. The interaction between age and condition was also significant, $Wald\ \chi^2(df = 1) = 8.84$, $p = .003$. These findings replicate those found with the forced-choice dependent measure, validating that these trends hold when a more sensitive dependent variable is employed (see Figure 8 for overall scores).

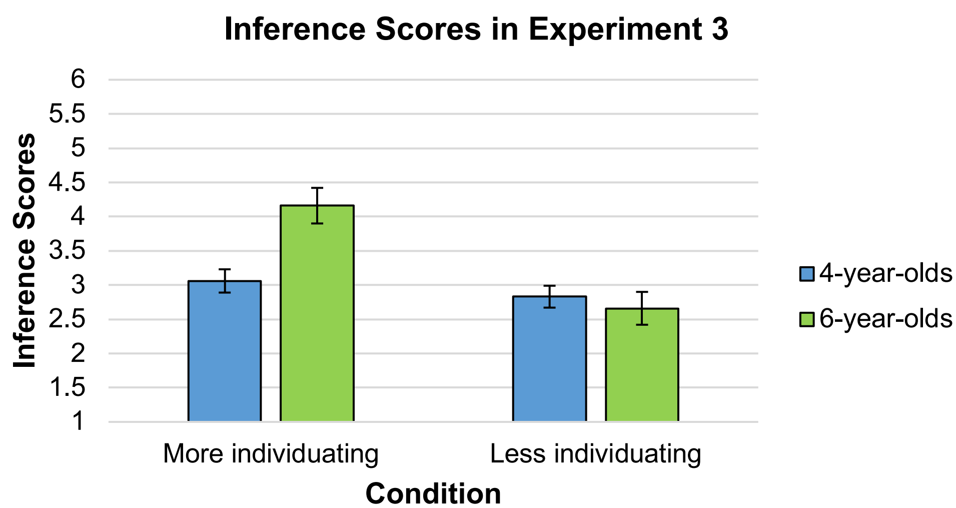

Figure 8. Inference scores for Experiment 3.

Discussion

In Experiment 3, I presented 4- and 6-year-old children with an adapted version of the lawyer-engineer problem in which base-rate and individuating information conflicted. When children were given a more descriptive piece of individuating information, 6-year-olds relied on the individuating information in their inferences while 4-year-olds opted to use the base-rate information. This replicated the findings in the conflict condition of Experiment 2, in which 4-year-olds' inferences trended toward base-rate use, while 6-year-olds over-relied on individuating information. That is, 6-year-olds continued to neglect base-rates in these problems, which contained personal preference information that is less practice and culturally engrained than the trait and gender stereotypes used in Chapter 2. Thus, it seems that their use of individuating information may be due in part to a stronger prior regarding the importance of personal characteristics in behavioural predictions. However, 6-year-olds did not arbitrarily rely on the individuating information, weighing its quality with the information provided by the base-rate. When presented with the less descriptive individuating information, 4- and 6-year-olds used the base-rate information in their decision. Thus, although 6-year-olds show a general tendency to rely on individuating information when it conflicts with base-rates, they seem to monitor the quality of the information provided and opt to rely more on base-rates when the individuating information is less informative.

The general trends observed with the binary choice dependent measure were replicated using the more sensitive inference scores (i.e., the product of a child's binary choice and their confidence in their response). This suggests that the forced-choice procedure is not masking children's responses, and affirms that the procedure used in Chapter 2 provided a valid assessment of children's estimates.

The results from Experiment 3 suggest that by age 6, although children are generally biased toward individuating information, they evaluate the quality of the individuating information in their inferences and weigh this information with the base-rate. Conversely, 4-year-olds seem to align their responses with the base-rate, regardless of the informativeness of the individuating information. From these findings, it is unclear if 6-year-olds are truly attending to the base-rate information or if their responses are simply based on the quality of the individuating information. That is, it is possible that 6-year-olds may be selecting the opposite of the individuating information because they find it uninformative. It is also unclear if 4-year-olds are able to use the individuating information in their responses at all. With no baseline data for Experiment 3, I cannot say whether 4-year-olds are prioritizing the base-rates or are instead completely unable to use this new group information in their inferences. In Experiment 4, I present children with an equal base-rate to examine their use of individuating information when the base-rate is uninformative. If children are attending to the base-rate and individuating information, they should rely on the individuating information in their inferences, though their use of this information should reflect the quality of the description it provides.

Experiment 4

Experiment 4 tested children's use of individuating information with an equal base-rate (i.e., 5:5). As in Experiment 3, children participated in one of two between-subjects conditions that varied the amount of individuating information provided.

Methods

Participants.

Children were tested individually at their schools or at a local museum. Sixty-four children were included in the final analyses, with 16 four- and 16 six-year-olds in each of the two conditions (see Table 5 for age and gender breakdown). An additional five children were tested and excluded for not finishing the task ($n = 2$), or failing both comprehension checks ($n = 3$; see Procedure for details).

Table 5. Age and gender breakdown per condition in Experiment 4.

	Mean age	Min age	Max age	Female
More individuating				
4-year-olds	55.87 months	48.91 months	59.13 months	6
6-year-olds	76.67 months	72.16 months	83.42 months	5
Less individuating				
4-year-olds	55.79 months	49.29 months	59.62 months	7
6-year-olds	77.09 months	73.88 months	83.91 months	9

Note: n = 16 per age, in each condition.

Materials and Procedure.

The procedure was identical to Experiment 3, with the exception of the base-rate information. In this experiment, children were presented with a 5:5 base-rate in both problems. Thus, the base-rate information should indicate a 50/50 chance that the child completed either activity. As in Experiment 3, the individuating information was varied between conditions, with half of the participants completing the *more individuating* condition, and half completing the *less individuating* condition. The order of the stories presented (i.e., activity story first, museum story first), the individuating information used, the placement of the base-rate objects, and the group introduced first were counterbalanced across participants.

Results

Children received a score of 1 if they selected the group that corresponded to the individuating information. 128 trials were included in the final analyses, because each child completed two trials (see Figure 9 for overall means). Preliminary analyses indicated there were no effects of counterbalancing (i.e., problem order: first, second; story: activity, museum) on children's responses. A GEE binary logistic regression with age (4-year-olds, 6-year-olds), condition (more individuating, less individuating), and their interaction revealed a significant main effect of and age, *Wald* $\chi^2(df = 1) = 4.92$, $p = .03$. Condition, *Wald* $\chi^2(df = 1) = 1.50$, $p = .22$, and the age by condition interaction, *Wald* $\chi^2(df = 1) = .41$, $p = .52$, did not have a significant effect on children's scores. Six-year-olds (more individuating *M* = .88, *SD* = .34; less individuating *M* = .75, *SD* = .44) tended to rely more on the individuating information than 4-year-olds (more individuating *M* = .66, *SD* = .48; less individuating *M* = .59, *SD* = .50).

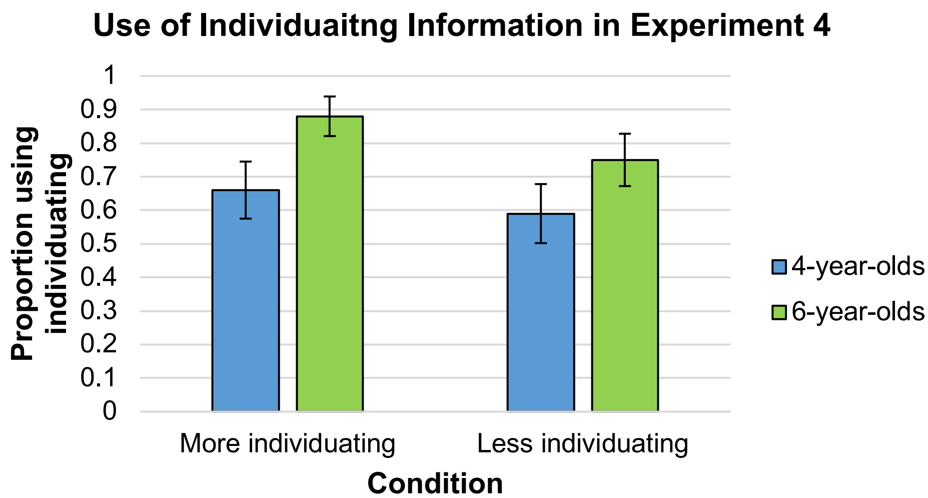

Figure 9. Proportion of children's responses that used individuating information in Experiment 4 (performance around 50% indicates use of base-rate information).

I created inference scores from children's responses and confidence scores. That is, children who used the individuating information and were confident received a score of 6, though those who did not and were confident received a score of 1. This created a 6-point scale. A GEE linear regression with age (4-year-olds, 6-year-olds), condition (more individuating, less individuating), and their interaction revealed a significant main effect of child's age, *Wald* $\chi^2(df = 1) = 7.17$, $p = .007$. Condition, *Wald* $\chi^2(df = 1) = 2.66$, $p = .10$, and the age by condition interaction, *Wald* $\chi^2(df = 1) = 1.54$, $p = .26$, did not have a significant effect on children's scores (see Figure 10 for overall scores).

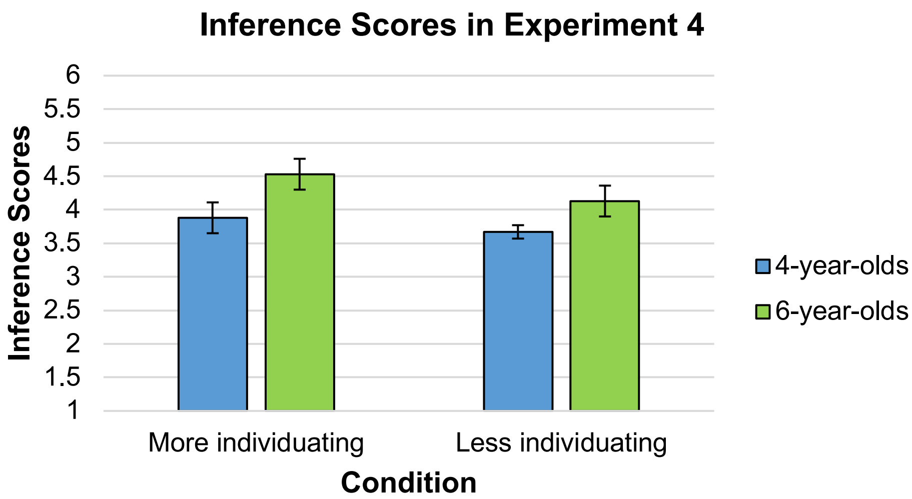

Figure 10. Inference scores for Experiment 4.

From these findings, it is unclear if 4-year-olds are favouring the base-rate information or simply did not find the individuating information representative of a child who might like the activity. When 4-year-olds' performance was collapsed across both conditions (as there was no effect of condition), it was marginally different from chance, *Wald* $\chi^2(df = 1) = 2.94$, $p = .087$. Although this suggests that they could, to some extent, use the individuating information in their decisions but were more pulled toward the base-rate, I presented a group of 4-year-old children ($N = 32$, $M_{age} = 55.48$ months) with the individuating information used in the less individuating condition. This was to ensure that they were able to provide a response based on individuating information in that condition, as children's performance was quite close to chance in this condition (59%). Similar to the baseline condition in Chapter 2, children were given the group information (i.e., some kids like to play baseball, and others like to make crafts) and were given individuating information about a child that corresponded to the descriptions used in less individuating conditions of Experiments 3 and 4. After they were given this information, participants were asked which activity the child preferred. Thus, children were not given any base-rate information. A GEE binomial logistic regression indicated that children used the individuating information in their responses at above chance levels ($M = .73$, $SD = .45$), *Wald* $\chi^2(df = 1) = 13.47$, $p < .001$, suggesting that they found the description representative of each group and could use this information in their inferences.

Discussion

In Experiment 4, I presented children with an equal base-rate and varied the quality of the individuating information. In this context, 6-year-olds used the individuating information in their responses. They also showed signs of sensitivity to the quality of the information, relying more on the individuating information when it provided them with a stronger description (88% vs. 75% of the time, though this difference was not statistically significant). Four-year-olds were less sensitive to this information and did not rely on the individuating information in their responses. In fact, they seemed pulled to the base-rate information, with their mean use of the individuating information close to chance performance. To rule out the possibility that 4-year-olds did not find the individuating information indicative of either activity, I established that they were able to use this information in their inferences when no base-rate information was available. From these findings, it seems that 4-year-olds are biased toward the base-rate information, though 6-year-olds weigh the quality of the information in their inferences.

The general trends observed with the binary choice dependent measure were once again replicated using the more sensitive inference scores, suggesting that the forced-choice procedure is not masking children's responses.

General Discussion

In Experiments 3 and 4, I presented 4- and 6-year-old children with base-rate and individuating information in a child-friendly version of the lawyer-engineer problem. To follow-up the experiments presented in Chapter 2, I manipulated the strength of the base-rate and individuating information provided. Across manipulations, 4-year-olds seemed biased to attend to the base-rate information. Although they were able to use the individuating information when presented on its own, 4-year-olds continued to neglect the individuating information even when it would have been reasonable to rely on it in their judgements. Conversely, 6-year-olds opted to rely on the individuating information when they were given a strong description that conflicted with the base-rates. However, when the individuating information was not as strong, 6-year-olds used the base-rates in their judgments. From these findings, it seems that 6-year-olds are generally biased toward the individuating information in their judgments, though they attend to the quality of the information when deciding whether to rely on it over the base-rate.

Between the ages of 4 and 6, children's approach to reconciling base-rate and individuating information undergoes drastic developmental change. Four-year-olds opted to use base-rate information in their judgements, even when this information was not particularly informative to their final decision in the equal base-rate condition. These findings are in line with previous work that has found that 4-year-olds tend to stick with the observed data, suggesting that they prefer a more data-driven approach than older children and adults (Gualtieri & Denison, 2018; Lucas et al., 2014; Seiver et al., 2013). Children use statistical data to learn complex relationships in their environments, and thus may consider base-rate data more informative than socio-cultural information. Given that even infants use base-rates and statistical information in their reasoning (Denison et al., 2013; Saffran & Kirkham, 2018), it is possible that 4-year-olds are more practiced in using statistical data. I return to this important point in Chapter 4.

Within a period of two years, children develop a preference for individuating information, as evidenced by 6-year-olds' reliance on a strong piece of individuating information that conflicted with a base-rate. Thus, this suggests that 6-year-olds have a more general preference for individuating information, which is not specific to the familiar traits and stereotypes that were used in Chapter 2. Rather, 6-year-olds seem to have a strong prior belief regarding the importance of personality descriptions in predicting behaviour, which is also evident in their preference for this information over statistical data in prior work (Gualtieri & Denison, 2018; Seiver et al., 2013). Moreover, 6-year-olds' preference for individuating information was moderated by the quality of the information. Whereas they relied on the strong individuating information when it conflicted with the base-rate, they used the base-rates in their inferences when it conflicted with a weaker piece of individuating information that contained representative information. Though this suggests that 6-year-olds may be integrating information, because their use of individuating information is moderated by its quality, it is somewhat puzzling that they were sensitive to this information here but were not discriminate in their use of individuating information in Chapter 2. Recall that when given irrelevant and no individuating information in Chapter 2, 6-year-olds did not default to base-rates. Though, when given lower quality, yet relevant, individuating information in Chapter 3, 6-year-olds used the base-rate in their inferences. Why might we see a difference across these manipulations? One possibility is that children in the irrelevant condition of Experiment 2, much like adults, believed that the description was relevant to their decision and attempted to reconcile this information with the base-rate, which led them to not fully default to base-rates in their judgments. However, it is unclear why 6-year-olds did not default to base-rates when given no individuating information, though relied on base-rates when weaker individuating information was available. It is possible that the no individuating problem in Chapter 2 contained garden-path features, misleading 6-year-olds due to the structure of the problem and lack of individuating information. Including lower quality individuating information in the less individuating conditions may have facilitated a comparison of the base-rate and individuating information, as both were present, whereas this was not highlighted when the problem lacked individuating information. Because these findings are still open to interpretation, it would be beneficial to systematically study the conditions in which 6-year-olds neglect base-rates. Determining which factors affect their base-rate use would aid in the design of interventions that target potential biases.

Limitations and Future Directions

In Experiments 3 and 4, I presented children with social group information that is less practiced in Western societies. This allowed me to examine the extent of children's preference for individuating information in general, apart from their prior experience with the specific information. Together, the findings in Chapters 2 and 3 suggest that 6-year-olds' preference for individuating information is not specific to traits and stereotypes that they are highly practiced in, but a general feature of their cognition. However, it would be beneficial to systematically examine the effect of prior experience on children's use of individuating information with relevant and irrelevant individuating information that vary in strength. Previous findings have established that familiarity with relevant social stereotypes is pivotal to children's use of individuating information (De Neys & Vanderputte, 2011). Because these studies did not manipulate the relevance and quality of the individuating information, systematically manipulating children's familiarity with social group information along with features of the individuating information would shed light on nuances in their base-rate neglect.

A methodological goal of the current experiments involved examining the validity of inference scores as a dependent measure. These scores were created using the product of children's response to the forced-choice question and their confidence in their response. The findings from the inference scores confirmed the general trends captured by the binary dependent measure, which aided in establishing the validity of the forced-choice scores. However, it is unclear if, and to what extent, this measure illuminated any nuances in children's ability to integrate information. Although the inference scores replicated the general findings of the binary dependent measure, they provided no further insight on children's decision-making. In Chapter 5, I revisit the issue of using a binary dependent measure and discuss alternative solutions for creating a continuous dependent measure suitable for this age group.

Implications and Conclusions

In Chapter 3, I presented findings that suggest 4- and 6-year-old children employ different strategies when reconciling base-rate and individuating information. Four-year-olds tend to align their responses with the base-rate data, even in situations where it would be reasonable to rely on individuating information. Conversely, 6-year-olds show a preference for strong individuating information, although weigh the quality of the individuating information and rely on base-rates when the individuating information is less descriptive. Together with the findings in Chapter 2, these four experiments provide pivotal insight on the development of the representativeness heuristic, illuminating nuances in the emergence and strengthening of children's biases that were previously unexplored.

From these findings, one may conclude that 4-year-old children rely heavily on base-rate information when making a decision, leading to what looks like mathematically normative responding from a relatively young group of children. However, it is plausible that 4-year-old children's preference for base-rate information may also be due in part to the social information at hand, just as it is for 6-year-olds. Although 4-year-olds prefer base-rates over individuating information, it is important to establish the parameters of this preference and examine their use of base-rates when they conflict with socio-cultural information that may be more salient in their early learning environments. Thus, in Chapter 4, I examine 4- and 5-year-olds' use of base-rates when they conflict with testimony information.

Chapter 4: The Development of a Testimony Bias in Young Children

In our daily lives, we are frequently in situations where multiple pieces of information should factor into our judgments and decisions. However, we sometimes forgo more comprehensive computations that involve integrating information and instead make decisions using simpler strategies that trade off accuracy for speed and computational efficiency. In a classic test of this phenomenon, adults were tasked with identifying the colour of a cab involved in a traffic accident (Bar-Hillel, 1980; Tversky & Kahneman, 1981). They were told that 85% of all cabs in the city were green and the other 15% were blue. A witness identified the cab as blue, and it was noted that they were accurate 80% of the time when identifying colours under viewing conditions similar to those during the accident. In their subsequent estimates, most participants reported that there was an 80% chance that the cab was blue. However, this estimation grossly neglects the base-rate of cabs in the city. According to Bayes' theorem, if base-rate and testimony information are properly considered, there is only a 41% chance of the cab being blue[2]. That is, there is a 41% chance that the cab is blue once you consider the relatively low base-rate of blue cabs, and the chance that the witness accidentally misidentified one of the more common green cabs as blue. Instead of integrating base-rate and testimony information, people complete a much easier computation and use the witness' accuracy as a shortcut for the full computation.

In Experiments 5 and 6, I examined the origins of the adult tendency to rely on testimony and neglect base-rates. I presented 4- and 5-year-olds with a child-friendly version of the taxi-

---

[2] Bayes Theorem: $\Pr(B|t=B) = \frac{\Pr(t=B|B)\Pr(B)}{\Pr(\mathrm{t}=\mathrm{B}|\mathrm{B})\Pr(B)+\Pr(t=b|G)\Pr(G)}$, where t is the witness' testimony, and B and G indicate blue and green respectively. We can compute $\Pr(B|t=B)$, the probability that it is really a blue car, given that the witness said it was blue, by substituting in the accuracy and base-rate information given in the classic problem: $\Pr(B|t=B) = \frac{(.8)(.15)}{(.8)(.15)+(.2)(.85)} \approx 0.41$

cab problem, in which they must decide whether or not to agree with a witness when her testimony conflicts with base-rate information. Researchers have begun to examine when heuristic reasoning, coupled with base-rate neglect, emerges in development. Young children are presented with age-appropriate versions of classic tasks, in which base-rate information conflicts with case-specific individuating information (e.g., the "lawyer-engineer" problem to examine use of the representativeness heuristic). For instance, participants may see a base-rate that contains more boy characters, though they may be told about a particular individual that likes to play with toys that, based on cultural gender stereotypes, girls typically like. Use of this heuristic seems to strengthen with development; 4-year-olds trend more toward base-rate use, while 5-year-olds begin to show a preference for the representativeness heuristic, which is further strengthened to nearly adult levels by 6 years of age (Gualtieri & Denison, 2018). A similar developmental difference has been observed in American children's proclivity towards the fundamental attribution error. This error is indicated by a bias toward person-specific explanations of others' behaviour that overlook the role of situational factors. By the age of 6, children endorse person-specific explanations of others' behaviour, similar to adults in Western societies. However, 4-year-olds are not biased towards these explanations and instead stick more closely to the observed behavioural covariations (Seiver et al., 2013). These experiments suggest that heuristic reasoning strengthens during early childhood.

At first glance, it might seem surprising that younger children would stick more closely to statistical data in their decisions, particularly when older children and adults use heuristics in lieu of this objective, reliable data. However, from as early as infancy, children are quite adept at using statistical data in their reasoning (Aslin, Saffran, & Newport, 1998; Denison et al., 2006; Denison et al., 2013; Girotto et al., 2016; Kirkham, Slemmer, & Johnson, 2002; Téglás et al., 2007; Xu & Garcia, 2008). Thus, by 4 years of age, children are highly practiced at reasoning about base-rates and covariation data and this may be what allows them to effortlessly use this information in their judgements and decisions (e.g., Gualtieri & Denison, 2018). However, they are not as practiced at using particular kinds of socio-cultural information in their reasoning, such as trait and stereotype-based information (Boseovski & Lee, 2006; Boseovski et al., 2013; Gonzalez et al., 2010; Liu et al., 2007; Martin & Ruble, 2004; Trautner et al., 2005), and thus they do not apply heuristics based on them when making judgments and decisions.

Although 4-year-olds are still developing their reasoning about traits and stereotypes, they are very adept at using testimony information in their inferences. One of the most important sources of knowledge for very young children is the social transmission of facts and norms (Harris et al., 2018). By the preschool years, children can judge whether a particular speaker is a good source of knowledge by considering factors like their past accuracy, confidence, and expertise (Koenig & Harris, 2005; Koenig & Sabbagh, 2013; Mills, 2013; Pasquini et al., 2007; Poulin-Dubois & Brosseau-Liard, 2016; Sobel & Kushnir, 2013).

Given that young children are so adept at identifying and learning from accurate testimony, it seems plausible that a bias to over-rely on such information might emerge very early in human ontogeny. On the other hand, children’s early emerging ability to skeptically evaluate social information (Harris et al., 2018; Mills, 2013) and their ability to integrate testimony with causal frequency information (Bridgers et al., 2016), suggest that they could succeed at integrating testimony with strong base-rate data. I tested these possibilities with 4- and 5-year-old children for two main reasons. First, 4-year-olds, but not 3-year-olds, have the ability to make rational inferences with probabilistic testimony data (Koenig & Harris, 2005; Pasquini et al., 2007). Thus, this is the youngest age group that possesses the requisite abilities to reason about testimony information in a taxi-cab problem. It is critical that the witnesses in the current problems are *probabilistically* accurate. If the witness is perfectly accurate (i.e., 100% correct) or inaccurate (i.e., 0% correct) at identifying colours, then there is no rational reason to integrate testimony and base-rate information, because children should always trust a perfectly accurate witness, or mistrust a perfectly inaccurate witness[3]. Second, including 5-year-olds allowed me to examine whether integrating testimony and base-rates or use of a testimony heuristic changes with age over this period or remains mostly stable.

Connecting the heuristics and biases and selective trust literatures has important implications for dual-process theories of cognition. Because heuristic strategies can result from learned associations (Stanovich et al., 2011) and 4-year-olds are very adept at learning from testimony, investigating children's ability to consider this particular social information in conjunction with base-rates in early development sheds light on children's reasoning more broadly. Using heuristics can be valuable, as they improve efficiency and are often effective, with the trade-off of introducing some systematic bias. However, applying a heuristic in inappropriate circumstances would be entirely ineffective. For instance, it would not be particularly useful to rely on a testimony heuristic (i.e., trust the information that a person provides rather than considering all available information) in cases where a person has proven unreliable in the past, particularly if other high-quality information is available. Thus, examining the circumstances in which children might over-rely on testimony is pivotal in establishing the pervasiveness of their bias, if it exists. In two experiments, I explored how 4- and 5-year-old children use testimony and base-rate information in tandem.

---

[3] For a 100% accurate witness, $\Pr(B|t = B) = \frac{(1)(.15)}{(1)(.15)+(0)(.85)} = 1$, so the correct behaviour is always to disregard the base-rate.

Experiment 5

Experiment 5 explored children's use of base-rate and testimony information separately in three between-subjects conditions (the base-rate condition, the accurate testimony condition, and the inaccurate testimony condition) to assess baseline use of this information for later comparisons to Experiment 6. The base-rate condition presented children with a group of ten dogs, eight wearing one colour collar and two wearing another colour. I was interested in children's use of this numerical information when guessing the collar colour of an unknown dog that was randomly sampled from the group. Based on previous work using similar types of paradigms, I predict that most children will choose the majority colour (Denison et al., 2013; Gualtieri & Denison, 2018).

There were also two accuracy conditions. Children in both accuracy conditions were introduced to a girl who liked to watch dogs in the park and identified the colours of six dogs' collars as they caught a ball. Her accuracy at identifying colours differed across conditions: in the accurate condition, she was correct 5/6 times on the previous day, while in the inaccurate condition, she was correct 3/6 times on the previous day. Following the accuracy sequence, children were introduced to a dog whose collar colour was unknown, and the girl provided testimony regarding which colour she thought she saw. Children were then asked to make an inference about the colour of the collar.

I developed these novel accuracy conditions to facilitate comparisons with the adult taxi-cab problem. I predict that most children should endorse the witness' testimony in the accurate condition, though it is unclear if they will discredit her testimony in the inaccurate condition. Children have opted to rely on information provided by an inaccurate informant when it is the only available information, and thus there was no conflicting information from another informant to rely on (Vanderbilt, Heyman, & Liu, 2014). Because the inaccurate testimony condition provides children with inaccurate information on its own, it was unclear if children would disregard the inaccurate witness.

Methods

Participants.

Prior to data collection, I established the criteria that I would stop testing children after I had obtained a full sample of 40 in each condition (see Table 6 for age and gender breakdown of participants in each condition). 120 children were included in the final analyses, with 20 four-year-olds and 20 five-year-olds in each of three conditions. Six additional children were tested and excluded due to parental report of low English language exposure (n=3) or non-compliance (n=3).

Table 6. Age and gender breakdown per condition in Experiment 5.

	Mean age	Female
Base-rate	60.58 months	18
Accurate condition	60.98 months	21
Inaccurate condition	60.75 months	20

Materials and Procedure

In three between-subjects conditions, children were told a story about a girl at a dog park via a PowerPoint presentation that was narrated live by an experimenter (see Figure 11 for an overview of the procedure).

In the base-rate condition, participants saw that there were ten dogs at the park wearing blue or yellow collars. Of the ten dogs, eight wore one colour (e.g., blue), and two wore the other colour (e.g., yellow). The experimenter counted the dogs and pointed out that more dogs were wearing one of the colours. Children were then asked to indicate which colour there was more of, and, depending on the child's response, the experimenter agreed or disagreed with their choice and stated that there were lots of dogs wearing blue and less wearing yellow. Children were then introduced to a dog at the park that day who was running away with a blanket covering its collar. Thus, the dog's group membership was unknown. Children were asked to recall which colour there was more of, and, depending on the child's response, the experimenter agreed or disagreed with their choice. The experimenter asked the child, "What colour is this one wearing?" The colour introduced first, the colour of majority collar, and the placement of the dogs in the base-rate array were counterbalanced.

In the accuracy conditions, participants were told that a girl at the park liked to identify what colour each dog was wearing while the dog chased a ball. During the history phase, participants saw what colour the girl thought she saw, followed by the actual colour of each dog, for six dogs. The witness was accurate 5/6 times in the accurate condition, and 3/6 times in the inaccurate condition. Children were asked if the witness was good or not good at identifying colours. Depending on the child's response, the experimenter agreed or disagreed with their choice: the experimenter stated the girl was good because she got five right and only one wrong (accurate), or stated that she was not very good because she got three right and three wrong, and was guessing (inaccurate). Children were then introduced to a dog at the park who was running away with a blanket covering its collar. Children were told what colour the girl thought the dog was wearing (i.e., "She saw it, so she says it's wearing yellow"). After this, participants were asked to recall what colour the girl thought the dog was wearing and if she was good or not very good at identifying the colours before. Children were corrected if they misremembered this information. The experimenter then asked the child, "What colour is this one wearing?" The colour introduced first, the order of collar colours during the accuracy portion, the order of the witness' correct responses during the history phase, and the colour of the witness' testimony were counterbalanced.

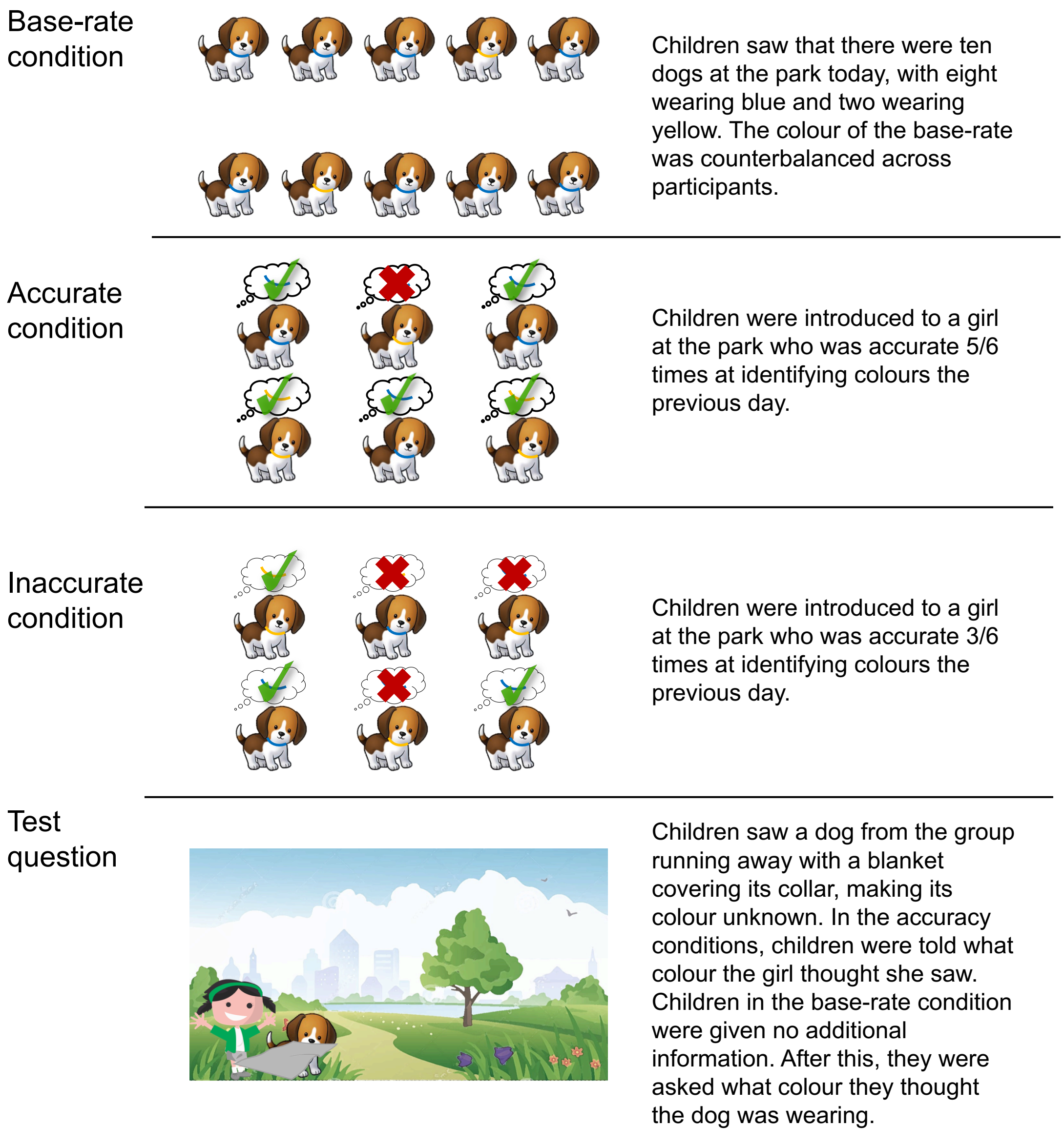

Figure 11. Overview of procedure in Experiment 5.

Results

Children were given a score of 1 if they chose the group that was indicated by the information they were given. That is, in the base-rate condition, children were given a score of 1 if they chose the majority group, and children in the accuracy conditions were given a score of 1 if they chose the colour indicated by the witness.

Table 7. Children's use of base-rate and testimony information in Experiments 5 and 6.

	Base-rate		Testimony	
	n	%	*n*	%
Experiment 5				
Base-rate condition	31	78%		
Accurate condition			29	73%
Inaccurate condition			24	60%
Experiment 6				
Accurate no conflict	38	95%	38	95%
Accurate conflict	6	15%	34	85%
Inaccurate no conflict	36	90%	36	90%
Inaccurate conflict	26	65%	14	35%

Note: n = 40 per condition.

I examined the base-rate condition separately from the accuracy conditions, given that children in this condition were responding to the question based on different information (see Table 7 for the means per condition). To explore any potential effect of age on responses, I conducted a logistic regression with children's age (4-year-olds, 5-year-olds) which indicated no significant effects of age on performance[4], *Wald* $\chi^2(df = 1) = .143$, $p = .705$. Children chose the majority colour at a rate higher than chance ($M = .78$, $SD = .42$, $p = .001$, binomial)[5].

I examined performance in the two accuracy conditions together to explore any potential effect of age. A logistic regression with accuracy condition (accurate, inaccurate) and children's age (4-year-olds, 5-year-olds) in the model revealed no significant effects of condition, *Wald* $\chi^2(df = 1) = 1.394$, $p = .238$, or age, *Wald* $\chi^2(df = 1) = .510$, $p = .475$. Despite the lack of

[4] For all regression analyses across both experiments, I found similar effects (no changes in significance cut-offs) when age was treated continuously.

[5] I also explored children's performance when they misremembered the information before the test question. Importantly, all children were corrected before moving on. In the base-rate condition, 6/40 participants misremembered the base-rate. In the accurate condition, 4/40 kids misremembered the witness' accuracy and 5/40 misremembered her testimony. In the inaccurate condition, 3/40 kids misremembered the witness' accuracy and 2/40 misremembered her testimony. Given that these numbers are so small, I did not perform any statistics, but it appears that children's data were very similar to the rest of the group when this information was misremembered but then corrected.

condition effect, I explored children's responses in each condition to establish the extent to which they relied on the testimony when it was the only available information. In the accurate condition, children chose the group indicated by the witness at a rate higher than chance (M = .73, SD = .45, p = .006, binomial), while performance in the inaccurate condition was not statistically different from chance (M = .60, SD = .50, p = .268, binomial).

Discussion

To establish children's baseline behaviour in our paradigm, Experiment 5 presented 4- and 5-year-old children with base-rate and testimony information separately. I observed no differences in performance as a function of children's age. Children in the base-rate condition relied on the 8:2 base-rate information and selected the majority group in their inferences at rates higher than chance. In the testimony conditions, children's responses did not significantly differ based on the witness' accuracy. Children who were presented with testimony from an accurate witness used this information in their inferences at rates higher than chance. However, children who were given testimony from an inaccurate witness did not use this information at rates significantly higher than chance, though they did not completely discredit it, given mean use at 60%. This finding is not particularly surprising because young children in previous studies were also willing to rely on testimony from an inaccurate informant in the absence of conflicting information (Bridgers et al., 2016; Vanderbilt et al., 2014).

Experiment 6

The results of Experiment 5 provide context for interpreting children's responses when they are presented with base-rate and testimony information together in the same problem. I manipulated the witness' accuracy at identifying colours (accurate: correct 5/6 times; inaccurate: correct 3/6 times) and whether this aligned or conflicted with the base-rate of dogs (no conflict:

her testimony aligns with the majority; conflict: her testimony conflicts with the majority, as she states it is the minority colour) in a 2x2 between-subjects design. This design results in four between-subjects conditions: the accurate conflict condition, the inaccurate conflict condition, the accurate no conflict condition, and the inaccurate no conflict condition.

The accurate conflict condition corresponds with the classic adult problem, which is why it is critical for examining whether children demonstrate base-rate neglect or, unlike adults, respond mathematically normatively and integrate the information. The witness, who is approximately 83% accurate, thinks that the collar of the missing dog is, for example, yellow, although 80% of the dogs are wearing blue. According to Bayes' Theorem, if children integrate the base-rate information with the witness' accuracy, they should say that the dog is wearing blue 45% of the time, as a group. If they rely exclusively on the witness' accuracy, as adults do, they should say the dog is wearing yellow approximately 83% of the time.

In the inaccurate conflict condition, the witness, who has been correct just 50% of the time, believes that the collar is yellow, and 80% of the dogs are wearing blue. This condition examines whether children more heavily weigh the reliable information (i.e., the base-rate information) and override testimony when a witness has proven to be unreliable. If children entirely neglect base-rates in favor of testimony, even when the witness has a history of inaccuracy, then it is possible they will use her testimony at a rate similar to Experiment 5 (i.e., approximately 60% of the time).

The two no conflict conditions serve as reference points for children's performance in this more complicated task. In the inaccurate no conflict condition, the witness is only 50% accurate and states that the collar is blue when 80% of the dogs are also wearing blue. This condition is included to rule out the possibility that children may reflexively provide the *opposite* response to an inaccurate witness' testimony, regardless of base-rates, when the problem becomes more complex and potentially harder to follow. Employing a shortcut to simply give the opposite response to the inaccurate witness would be irrational in this situation because the base-rate information points in the same direction. The accurate no conflict condition should be entirely uncomplicated. The witness, who is correct 83% of the time, thinks that the collar is blue and 80% of the dogs are also wearing blue. In sum, children in both no conflict conditions should choose the colour endorsed by the witness and the base-rate information. These conditions also allowed me to assess whether having two converging pieces of information have an additive effect on children's decisions.

Methods

Participants.

I again tested 40 children in each condition. 160 children were included in the final analyses, with 20 four-year-olds and 20 five-year-olds in each of the four conditions (see Table 8 for age and gender breakdown). Five additional children were tested and excluded because of interruption in the testing environment (contractors entered the room during testing; n=1) or non-compliance (n=4).

Table 8. Age and gender breakdown per condition in Experiment 6.

	Mean age	Female
Accurate no conflict	60.90 months	24
Accurate conflict	60.37 months	19
Inaccurate no conflict	60.65 months	21
Inaccurate conflict	60.80 months	24

Materials and Procedure.

Participants were told that a girl at the park liked to identify what colour collar each dog was wearing while they chased a ball (see Figure 12 for an overview of the procedure). During the history phase, participants were told about the witness' accuracy when identifying the colours of 6 dogs on the previous day, using the same 5/6 or 3/6 accuracy rates as in the accuracy conditions for Experiment 5. Participants then saw a group of ten new dogs and were told that these dogs were at the park on the current day. The experimenter counted the dogs and established the majority (8:2) as in the base-rate condition of Experiment 5. Because children were presented with two pieces of information in Experiment 6, I included a recap slide where the experimenter reminded participants what colour there was more of at the park on the current day, and how accurate the witness was at identifying colours the previous day. Information was always recapped in this order, mimicking the structure of the typical adult taxi-cab problem. This recap reduced the memory demands of the task and replaced the questions that the experimenter previously asked the children (and corrected if they provided incorrect responses). Because children's performance did not differ based on whether they misremembered this information or remembered correctly in Experiment 5, these questions were replaced with this recap slide to shorten the procedure while still ensuring that all children were reminded of the correct information. Children were then introduced to a dog at the park that day who was running away with a blanket covering its collar, making its group membership unknown. Children were told what colour the girl thought the dog was wearing (i.e., "She saw it, so she says it's wearing yellow"). The experimenter then asked the child, "What colour is this one wearing?". The colour introduced first, the order of collar colours during the accuracy portion, the order of the witness' correct responses during the history phase, the colour of the majority collar, the placement of the dogs in the base-rate array, and the colour of the witness' testimony were counterbalanced.

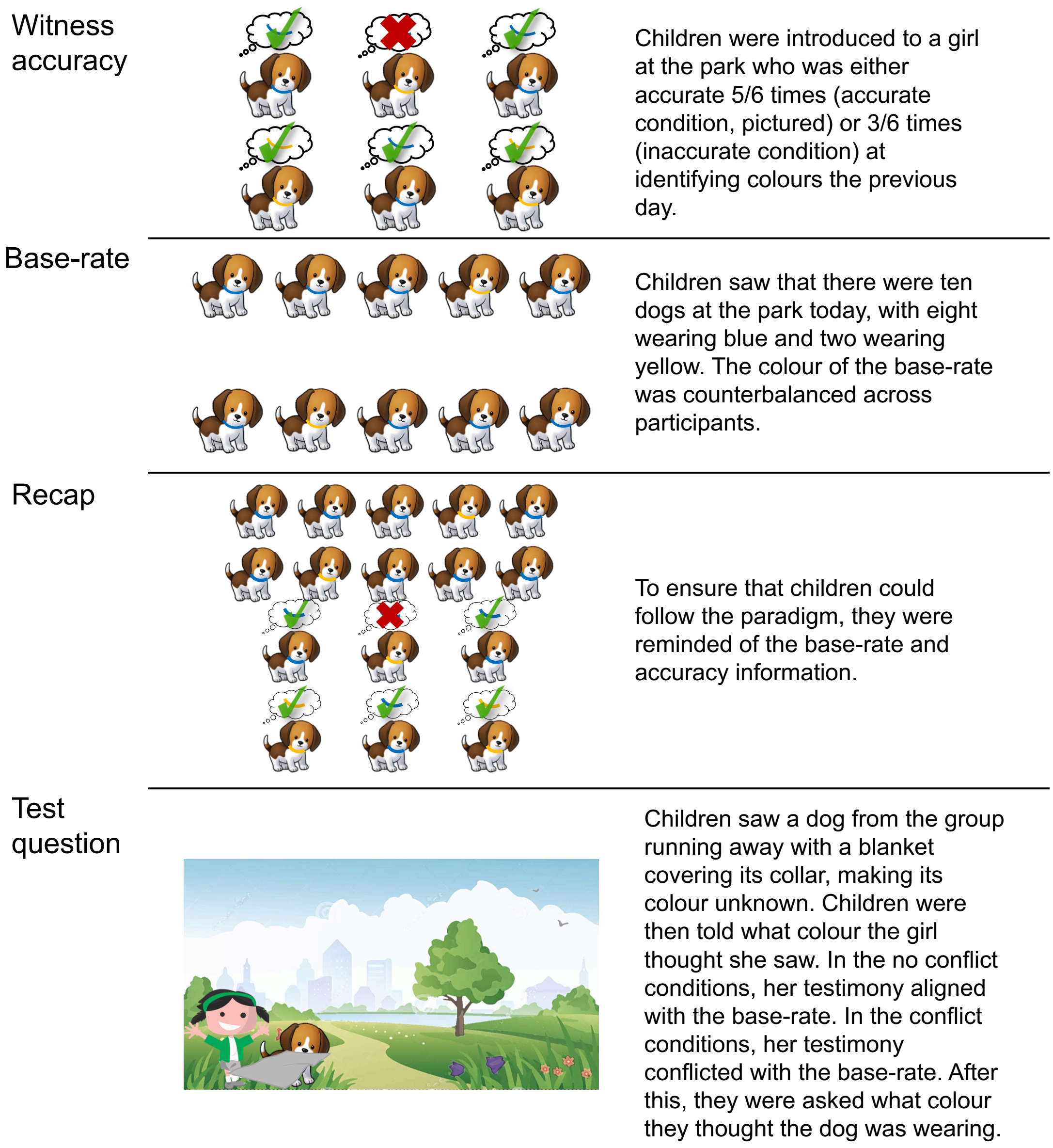

Figure 12. Overview of procedure in Experiment 6.

Results

Children received a score of 1 if they selected the group indicated by the base-rate (in no conflict cases, this cues the same response as when coded by testimony). See Figure 13 for a graph of the means per condition.

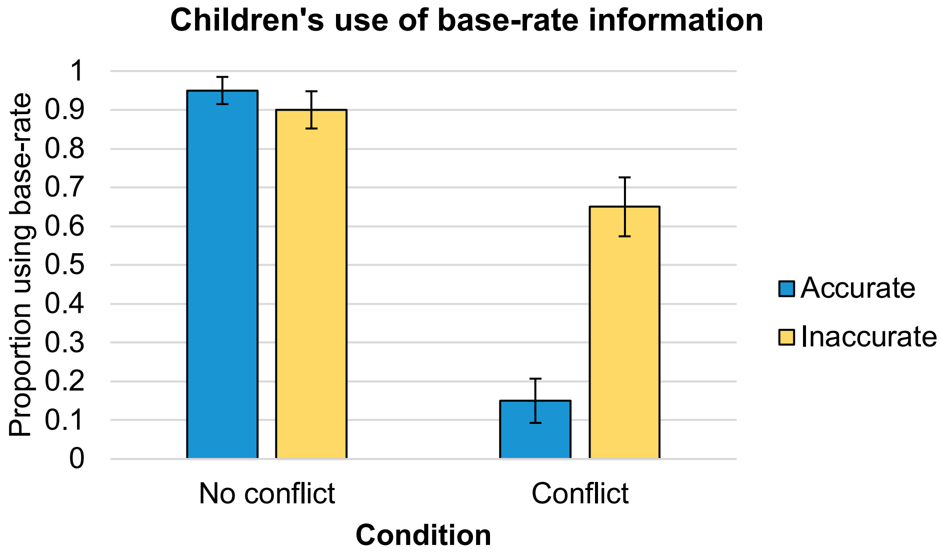

Figure 13. Proportion of children choosing the higher base-rate option in Experiment 6.

To explore children's responses across conditions and any effects of age, I conducted a logistic regression with conflict condition (conflict, no conflict), accuracy condition (accurate, inaccurate), children's age (4-year-olds, 5-year-olds), and the interaction between conflict condition and accuracy condition included in the model. This revealed a significant main effect of conflict condition, *Wald* $\chi^2(df = 1) = 35.273$, $p < .001$, and an interaction between conflict and accuracy condition, *Wald* $\chi^2(df = 1) = 8.662$, $p = .003$, no main effect of accuracy condition, *Wald* $\chi^2(df = 1) = 2.325$, $p = .127$, and no main effect of age, *Wald* $\chi^2(df = 1) = 0$, $p = 1$. A Fisher's exact test indicated that children's use of base-rate information on conflict problems significantly differed based on the witness' accuracy ($p < .001$).

To further examine children's use of base-rate and testimony information, I compared children's performance in Experiment 6 to the baseline conditions in Experiment 5 (see Table 7 for a comparison of performance). I first explored children's performance in the no conflict conditions. Testimony and base-rate information cued the same group in no conflict conditions, and thus higher scores reflect a tendency to respond based on both types of information. These responses were then compared to the baseline base-rate and testimony performance in Experiment 5 using Chi square analyses. In the accurate no conflict condition, children's responses (*M* = .95, *SD* = .22) differed significantly from their base-rate use in Experiment 5 (*M* = .78, *SD* = .42; *p* = .048), and their use of testimony in the accurate condition in Experiment 5 (*M* = .73, *SD* = .45; *p* = .013). This suggests that when the information converges and all information is reliable and relevant, there is an additive effect on children's judgments.

In the inaccurate no conflict condition, children's responses (*M* = .90, *SD* = .30) did not differ significantly from their base-rate use in Experiment 5 (*M* = .78, *SD* = .42; *p* = .225). However, children's responses differed significantly from their use of testimony in the inaccurate condition of Experiment 5 (*M* = .60, *SD* = .50; *p* = .004). In this case, having the reliable base-rate information coupled with the unreliable testimony led children to make stronger inferences than with unreliable testimony alone. The results from these conditions confirm that participants could follow the narrative in both accuracy conditions, and that they do not automatically disagree with an inaccurate witness.

I then examined children's performance in the conflict conditions, in which testimony and base-rate information cued different groups. I first examined performance in the accurate conflict condition, which maps onto the classic adult taxi-cab problem. First, I examined if children's responses were in line with an integration strategy that weighed both base-rate and accurate testimony information. If children were using this strategy, approximately 45% of participants should choose the group cued by the base-rate. However, I found that their performance significantly differed from this value (*M* = .15, *SD* = .36; binomial $p < .001$). I then examined whether children might be relying only on testimony using a Chi square: they relied on the testimony information in Experiment 6 (*M* = .85, *SD* = .36, coding reversed for comparison) at similar rates to Experiment 5 (*M* = .73, *SD* = .45; *p* = .274), suggesting that they were focusing on this information in Experiment 6. Altogether, these analyses are most consistent with the interpretation that children did not integrate base-rate and testimony information, neglected base-rates, and defaulted to testimony use. This is very similar to adult behaviour in the classic problem; adult judgments differ from the value that would be predicted if base-rates and testimony were integrated, and do not differ from the value that would be predicted if only witness accuracy was considered.

Next, I explored children's performance in the inaccurate conflict condition to examine the pervasiveness of children's reliance on testimony. In this condition, the inaccurate testimony conflicted with more reliable base-rate information, so reliance on testimony in this context would be ineffective. Children's use of testimony information (*M* = .35 *SD* = .48) differed significantly from their use of testimony in Experiment 5, as they relied on the witness significantly more in their inferences in Experiment 5 (*M* = .60, *SD* = .49; *p* = .043). Thus, children were selective in their use of testimony when more reliable base-rate information was available.

Discussion

In Experiment 6, I presented children with base-rate and testimony information in the same problem. I observed no effects of age. When both pieces of information aligned, children performed at near-ceiling levels, selecting the colour indicated by the base-rate and the witness. This effect appeared to be additive when both pieces of information were reliable. However, children over-relied on the accurate witness' testimony when it conflicted with the base-rate, opting to use the testimony in their inferences in a similar manner as adults in the classic taxi-cab problem. Notably, this bias for testimony was not extended to the inaccurate witness. Whereas relying on the accurate witness is reasonable, a bias toward the inaccurate witness' testimony would be irrational due to her prior inaccuracy. When the inaccurate witness' testimony conflicted with the base-rates, children were pulled more toward the base-rate information and did not reflexively rely on the testimony information.

General Discussion

The current experiments explored how children reconcile information from witness testimony with base-rates and found that children selectively exhibited a testimony bias. At baseline, children relied on base-rate and accurate testimony information when they were presented separately in Experiment 5. However, children did not rely on the inaccurate witness' testimony at rates higher than chance, though they did not completely discredit the information she provided. In Experiment 6, children were presented with base-rates and testimony information together. When both pieces of information cued the same response, children performed at near ceiling levels and selected the group that was suggested by both the base-rates and testimony. The pivotal conflict conditions presented children with conflicting base-rate and testimony information. When the witness was accurate, children were more likely to use her testimony in their inferences than the base-rates, and group level responses indicated no signs of integrating the evidence. This suggests that the same bias seen in adults is also present early in development. However, children's preference for testimony information over base-rates was appropriately selective. Compared to the accurate conflict condition, children in the inaccurate conflict condition were pulled toward the group cued by the base-rates. Although they relied on the inaccurate witness' testimony slightly more often than not when it was the only information available in Experiment 5, they downgraded this significantly when it conflicted with a meaningful base-rate in Experiment 6. Taken together, these findings suggest that children may use a testimony shortcut at very young ages, but they use it selectively.

Much previous work has established that young children reliably use base-rate (Denison et al., 2006; Kushnir et al., 2010; Ma & Xu, 2011) and accurate testimony (Harris et al., 2018; Koenig & Harris, 2005; Pasquini et al., 2007) information in their inferences. The current findings suggest that a bias in favor of testimony over base-rates emerges early in development, with young children, much like adults, over-relying on the information provided by an accurate but imperfect witness. On the surface, these findings appear to contrast with research on heuristic use in development that has suggested 4-year-olds rely on base-rate and covariation data in their inferences and that biases in favor of socio-cultural information strengthen with age (Gualtieri & Denison, 2018; Seiver et al., 2013). In the current experiment, a bias for testimony (and its resultant base-rate neglect) emerges earlier at 4-years of age, as opposed to 5- or 6-years. Why might we see such an early emergence of this adult-like behaviour? The tendency to neglect statistical base-rate information in favor of socio-cultural information may emerge at different ages depending on the socio-cultural information, reflecting differences in children's dexterity with the social information at hand. Specifically, 4-year-olds are fine-tuning their understanding about the role of traits in predicting behaviour, though 5- and 6-year-old children are quite proficient in using this information (Boseovski & Lee, 2006; Boseovski et al., 2013; Gonzalez et al., 2010; Liu et al., 2007). However, evaluating the accuracy of testimony becomes practiced much earlier in development, as evidenced in 4-year-olds' sophisticated inferences about probabilistically accurate informants (Koenig & Harris, 2005; Pasquini et al., 2007). Thus, the current findings provide additional insight into the nuances of young children's abilities to reason about multiple sources of information.

The findings of the current experiments are helpful in understanding the development of heuristic use and base-rate neglect. Notably, young children did not rely on the testimony of the inaccurate witness to the same extent as the accurate witness when her testimony conflicted with the base-rates. Children in the inaccurate testimony condition were pulled more toward base-rates than those in the accurate testimony condition. Recent findings, in which children were presented with a single, inaccurate informant, have also found that children's use of inaccurate testimony is contingent on the presence of conflicting information, which may facilitate their ability to weigh and contrast the information they are given (Bridgers et al., 2016; Vanderbilt et al., 2014). Children opted to rely on information provided by an inaccurate informant when it was the only available piece of information. However, children relied on a neutral informant, with no prior history of accuracy, who provided information that conflicted with the inaccurate informant (Vanderbilt et al., 2014). In the current experiments, children did not completely discredit the inaccurate witness, though they did not completely rely on her testimony when it was the only available information in Experiment 5. Similar to previous findings, when the inaccurate witness was paired with more reliable base-rate information, children trusted the inaccurate witness less when her testimony conflicted with the base-rate. Together with other recent findings on children's ability to integrate testimony and causal frequency information (Bridgers et al., 2016), these findings suggest that young children can effectively weigh testimony information with other pieces of information.

Conclusion

Chapter 4 presented findings that are the first to explore 4- and 5-year-old children's use of base-rates in the presence of conflicting testimony information by presenting them with a child-friendly version of the taxi-cab problem. Though young children used a heuristic shortcut to rely on the testimony of an accurate witness when it conflicted with base-rates, they were more selective in their use of inaccurate testimony. Because young children are quite sophisticated in their use of testimony and base-rate information early in development, the current findings have important implications for the development of heuristic thinking in children.

Chapter 5: General Discussion

In six experiments, I sought to examine the development of information integration and base-rate neglect in 4- to 6-year-old children. In Chapter 2, I examined the development of the representativeness heuristic. When base-rate and individuating information conflicted, 4-year-olds trended more toward relying on base-rates in their inferences, though 5-year-olds trended toward individuating information, which was used at adult-like rates by the age of 6. In Chapter 3, I further explored age differences in the performance of 4- and 6-year-olds. These results suggested that 4-year-olds focus more on the base-rate information, even in cases where it would be reasonable to rely on the individuating information. However, 6-year-olds seem to weigh the quality of the base-rate and individuating information, though they show a general bias toward the individuating information in their inferences. In Chapter 4, I explored 4- and 5-year-old children's use of base-rates when it conflicted with a witness' testimony. Children relied on the accurate witness' testimony when it conflicted with the base-rates, but trended more toward base-rate use when the witness was inaccurate. Thus, even 4-year-olds neglected base-rates that conflicted with social information, suggesting that nuances in children's base-rate neglect may lie in their experience with the socio-cultural information that conflicts with the base-rate.

Implications for Dual Process Theories of Development

The current experiments present findings on the development of base-rate neglect from the youngest age group tested to date, offering important implications for dual process theories of cognition, which have had little data to work with regarding the emergence of heuristics and biases in development. Dual-process theories posit that decision-making relies on two different types of processing (Stanovich et al., 2011). Type 1 processing refers to relatively quick and computationally-efficient processes automatically activated by stimuli as the result of innate modules or learned associations. In the lawyer-engineer problem, occupational stereotypes cued by the group information trigger Type 1 processing that may lead to a non-normative (often also referred to as heuristic) response. Type 2 processing refers to relatively slow and computationally-expensive processes that are deliberately employed. For instance, Type 2 processes are used when properly integrating base-rate and individuating information to derive a normative (often also referred to as analytic) response. Both Type 1 and Type 2 processes rely on different mindware (e.g., strategies, rules, and procedures) relevant to different responses. In order to provide a non-normative response, the decision-maker must be aware of occupational stereotypes and how they are used to infer behaviour. In order to provide a normative response, the decision-maker must be aware of these social stereotypes but also use the correct strategy to integrate information.

Considering the mindware young children have at a particular age is vital to understanding the emergence and strengthening of heuristics and biases. Heuristic and analytic strategies are fine-tuned as children learn new strategies and experience more opportunities to use them (Stanovich et al., 2011). Thus, depending on the Type 1 and 2 mindware required, both types of processing can improve with age. For instance, children acquire Type 1 strategies as they learn to rely on socio-cultural information, such as personal characteristics and testimony from others. Strategies based on this information strengthen with experience as children come to use this information in their inferences. Moreover, Type 2 analytic strategies also improve with age. Children acquire new strategies for integrating information, which can lead to analytical responses that appropriately weigh social and numerical information.

Initial work on the representativeness heuristic in children found an interesting developmental trend with young children producing more analytical responses than older children and adults (Davidson, 1995; Jacobs & Potenza, 1991). However, recent findings have suggested that children use heuristics as soon as they are familiar with the social information at hand, as young children's use of base-rates was contingent on their familiarity with the social stereotype (De Neys & Vanderputte, 2011). That is, 5-year-olds used base-rate information when they were presented with a problem that involved stereotypes about principals and construction workers, though they were less likely to use the base-rate when the problem contained stereotypes about boys and girls. From this work, it was speculated that children employed Type 1 strategies as soon as they were familiar with the stereotypes tested.

However, from the current findings, it is evident that familiarity with group information is not the sole determinant of its use. Recall that 4-year-olds relied on traits and gender stereotypes in their inferences when it was the only available information in the problem. Although children of this age can make inferences based on individuating information, I found that 4-year-olds produced responses that were more aligned with the base-rate information when both pieces of information were presented. Children's use of individuating information is quite pronounced by the age of 6, with children relying on the individuating information at similar rates as adults in the conflict problems in Chapter 2. Moreover, in Chapter 3, 6-year-olds continued to rely on individuating information in their inferences when it characterized social information that was less practiced. Thus, heuristic responses based on individuating information were strengthened within a period of two years, suggesting that a mechanism of children's general cognition might contribute to this preference apart from their familiarity with the social information at hand. Between the ages of 4 and 6, children form a strong theory about the role of personal characteristics in predicting behaviour. This is supported by findings on the fundamental attribution error in young children, which have uncovered a similar age difference between 4 and 6-year-olds, further establishing that by the age of 6, young children have formed a strong theory about what information is useful in predicting and explaining behaviours (Seiver et al., 2013).

The performance of 4-year-olds is quite interesting: although they are familiar with the individuating information, they opted to rely on the base-rates. Moreover, findings from Chapter 3 suggest that 4-year-olds may be in fact *over-relying* on base-rate information. When given an equal base-rate, 4-year-olds were pulled toward chance performance, even though it would have been reasonable to rely exclusively on the individuating information in this case. Thus, at this age, children may have not formed a strong theory about the importance of personal characteristics in predicting and explaining behaviour (Seiver et al., 2013). While they are developing this theory, 4-year-olds are more open to the observed data than older children and adults who hold strong prior beliefs that pull them away from the observations (Lucas et al., 2014). Because even infants are sensitive to numerical and statistical data (Denison & Xu, 2014; Saffran & Kirkham, 2018), children's proficiency in using base-rate information may lead to its overuse in some contexts. Four-year-olds' base-rate use also provides further insight on the distinction between Type 1 and Type 2 processes. The ease with which 4-year-olds rely on base-rate information provides support for the recent argument that base-rate reasoning can be automatically activated, and as quick and computationally efficient, as a Type 1 process (Pennycook, Trippas, Handley, & Thompson, 2014).

However, children's use of a particular strategy depends on the information presented, highlighting the differential emergence of base-rate neglect across problems. That is, although 4-year-olds relied on base-rates in variants of the lawyer-engineer problem, they neglected base-rates when it conflicted with accurate testimony information in Chapter 4. Similar to how 6-year-olds seem to have acquired a general preference for personal characteristics, this suggests that 4-year-olds are quite practiced in using accurate testimony information in their inferences and see it as more diagnostic than base-rates. Because 4-year-olds quite skillfully compare the accuracy of speakers to determine who to learn from (Koenig & Harris, 2005; Pasquini et al., 2007), the emergence of a biased Type 1 process seems contingent on children's acuity with the socio-cultural information available in the problem.

Children's Ability to Integrate Information

Finally, the current findings offer insight on the development of a more sophisticated ability to integrate social and numerical information. It should be noted that although base-rate use closely resembles normative responses, the correct response is to weigh the base-rate and social information using a computation that would require Type 2 processing. Due to the nature of a forced-choice dependent variable, the current studies could not directly investigate children's information integration abilities. This point is made clear when interpreting the performance of 4-year-olds in Chapter 3, who performed around chance in the equal base-rate condition. It is unclear if 4-year-olds were integrating base-rate and individuating information, over relying on base-rate information, or if this trend was related to noise in their responses.

Although the dependent measure is limiting, comparing children's responses across different conditions can provide initial insight on the development of children's ability to weigh information. For instance, 4- and 5-year-old children in Chapter 4 chose to rely on the accurate witness when her testimony conflicted with the base-rate, though they were selective in their use of testimony and did not rely on the inaccurate witness when she conflicted with base-rates. Though children used the inaccurate witness' testimony in their inferences more often than not when it was the only information available to them in the problem, the fact that children opted to rely on the base-rate information suggests that they are weighing the accuracy of the witness with the reliable base-rates (Bridgers et al., 2016). Moreover, 6-year-old children seemed to weigh the strength of the individuating information in Chapter 3, opting to rely on base-rates when the individuating information was less informative. Although this suggests 6-year-olds are attempting to integrate base-rate and individuating information, recall that they neglected base-rates in the irrelevant individuating condition and no individuating condition in Chapter 2, which highlights an issue in their ability to integrate information in those conditions. These questions are open to investigation, where perhaps a more continuous dependent measure could uncover nuances in children's information integration abilities that may have been overlooked.

Limitations

One limitation of the current design is the forced-choice response method. In the adult paradigm, participants are asked to estimate the likelihood that an individual is, for example, an engineer, or to rate the likelihood that the taxi-cab is blue, using a percent value (Bar-Hillel, 1980; Kahneman & Tversky, 1973; Kahneman, 2011; Tversky & Kahneman, 1974; Tversky & Kahneman, 1981). I used a binary choice paradigm to ensure that young children were able to provide a response; children this age cannot estimate likelihoods using percent values and face issues providing relevant explanations for their thought processes. Previous work with young children has indicated that an individual child's responses over repeated trials tend to reflect the group distribution as a whole, suggesting that aggregating responses across a group of children in a forced-choice paradigm reliably represents an individual child's beliefs (Denison et al., 2013). A binary response is also desirable from an ecological validity perspective because regardless of certainty, people ultimately make categorical decisions. Although this forced choice approach is typical of developmental studies of judgment and decision-making, a dichotomous dependent variable is limiting in terms of sensitivity. In Chapter 3, I attempted to develop a continuous measure from children's binary responses and their confidence in their response. Although I found that these scores replicated the general trends captured by the binary measure, it did not offer any additional insight on children's ability to integrate information, possibly due to issues with children's interpretation of the confidence question in which most children indicated they were guessing. In fact, it is reasonable that children indicated they were guessing on this question, since they could not be sure that their answer was correct (as they were not shown what activity the child in the story completed). Future studies could employ a rating scale to obtain more sensitive and graded judgments, providing additional insight into children's degree of belief in a particular choice. Developing a measure that better captures the degree of children's responses will aid future investigations, perhaps uncovering trends that were masked by the use of a binary dependent measure.

The final two limitations acknowledge plausible alternative accounts for the current age differences. First, it is possible that age differences in children's sensitivity to pragmatics may have accounted for some of the developmental trends observed between 4- and 6-year-olds. Adults that are presented with the lawyer-engineer problem have argued that they felt inclined to rely on the individuating information because of demand characteristics. That is, when an adult is presented with background base-rate information and then a lengthy description of the individual, they feel that the experimenter is imploring them to use that individuating information in their inferences. In fact, altering pragmatic aspects of the problem affects adults' strategy use (Schwarz et al., 1991; Zukier & Pepitone, 1984). I modeled my problems after the original work, which contains the pragmatic features that adults sometimes find problematic. Because of this design, children's sensitivity to pragmatics may have prompted use of individuating information in 6-year-olds, while an insensitivity to these features could have facilitated base-rate use in 4-year-olds. Between the ages of 4 and 6, children undergo drastic changes in their sensitivity to conversational pragmatics (Matthews, 2014). For example, 6-year-olds', but not 4-year-olds', sensitivity to the quality manipulation in Chapter 3 was interpreted as a development in sensitivity to information quality. However, these data could also be viewed as evidence for a pragmatic account. Because the lower quality individuating information was also much shorter, it might be seen as less heavy-handed, which could reduce its use if children are attuned to this pragmatic feature. To tease apart this alterative account, future work could present children with these problems and a measure of their pragmatic development to see whether better-developed pragmatic reasoning predicts heuristic use. One could also examine if experimentally manipulating pragmatic features of the problem affects 6-year-olds' use of individuating information.

Second, age differences in children's cognitive abilities could account for developmental trends related to their strategy use. Children's executive functions (e.g., working memory, inhibitory control, and cognitive flexibility) undergo drastic developments during the early childhood years (Diamond, 2013; Garon, Bryson, & Smith, 2008). When adapting adult tasks for children, one must capture the essence of the adult task while ensuring it is age-appropriate for a developmental population (Stanovich et al., 2011). In order to preserve the structure of the original problems, I presented children with problems that were somewhat complicated; children had to track the narrative, along with the numerical and social information, and needed to keep this information in mind in order to respond to the test question. Somewhat counterintuitively, less sophisticated cognitive abilities of 3- and 4-year-olds may *facilitate* their ability to learn from observed data, as evidenced in their tendency to maximize probabilistic information

(Thompson-Schill, Ramscar, & Chrysikou, 2009). Moreover, executive functions play a role in determining the strategies children use to solve physical reasoning (Baker, Gjersoe, Sibielska-Woch, Leslie, & Hood, 2011; Bascandziev, Powell, Harris, & Carey, 2016) and theory of mind (Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002) problems. Future work could formally test this relationship, uncovering how features of children's cognition, such as their executive functions, affect their approach to integrating base-rate and social information.

Future Directions

The current findings emphasize that the development of heuristics and biases is considerably more nuanced than previously thought. Much future work is needed to understand the emergence and strengthening of these strategies over the course of development. Uncovering the conditions that facilitate base-rate use may aid in the design of interventions that aim to improve sensitivity to numerical information.

Much work in the adult literature has focused on the conditions that *elicit* base-rate use. One feature of the problem that affects adults' use of base-rates is emphasizing the role of random sampling. That is, when random sampling is highlighted in the problem, adults tend to incorporate base-rates in their judgments (Koehler, 1996). Notably, infants and young children are sensitive to sampling information in their probabilistic inferences (Denison et al., 2006; Kushnir et al., 2010; Ma & Xu, 2011; Xu & Denison, 2009). Because human decision-makers are sensitive to sampling information quite early in development, it is possible that manipulating this feature of the problem would affect children's base-rate use, notably in cases where children may opt to rely on social information. Follow-up work could present children with a manipulation that specifies how the case was sampled. For instance, in an adaptation of the lawyer-engineer problem, participants could be told that the teacher drew the student's name out of a hat (random sampling), or that the teacher selected her favourite student (non-random sampling).

Moreover, adults utilize base-rate information if the base-rates can be incorporated into their causal schema of the problem space (Ajzen, 1977; Bar-Hillel, 1980; Krynski & Tenenbaum, 2007). For instance, in the taxi-cab problem, adults incorporate the base-rate information into their judgments if the base-rate contains causal information, such as the rate of accidents that each company has been involved in. Sensitivity to causal information also emerges early in childhood as children actively construct causal representations of the world around them (Gopnik et al., 2004). Exploring the ways in which causal models facilitate children's ability to integrate social and numerical information would provide insight fruitful to the design of interventions. An adapted version of the taxi-cab problem from Chapter 4 that may better fit children's causal model would be particularly illuminating. Instead of a population base-rate, children could be shown an 8:2 base-rate of the fastest dogs at the park, indicating that one colour tends to signify the faster dogs. If this information better suits children's causal model, they should rely more heavily on the base-rate in their inferences when it conflicts with the accurate witness.

Developing a paradigm that more closely resembles real-world problems would provide additional insight on children's ability to use social and numerical information on a daily basis. Recent work has also found that adults make accurate predictions about everyday phenomena, suggesting that they have intuitive theories about prior probabilities (Griffiths & Tenenbaum, 2006). It would be interesting to explore if children similarly hold accurate beliefs about prior probabilities and how they integrate this information in their judgements. For instance, children may hold prior beliefs about the ratio of nice to mean people in the world, as a higher proportion of people are consistently nice. In Chapter 2, participants were presented with a base-rate of nice and mean robots, which was heavily skewed toward mean characters for half of the participants. Presenting children with a problem that explores their use of a natural population base-rate would elucidate how they come to use this information in the everyday inferences. Moreover, base-rates that contain information about a population may be less salient in children's early learning environments than those that track the behaviour of an individual (e.g., a base-rate of how many times a particular child plays with gender-stereotypical girl toys and boy toys). Examining children's use of base-rates in more naturalistic contexts would aid in uncovering how they come to use this information in their everyday environments.

Implications

Learning, and employing, a heuristic by the age of 6 is quite remarkable. Even though it can introduce some systematic bias, heuristic reasoning is efficient and often effective. The emergence and strengthening of biases between the ages of 4 and 6 suggests an interesting opportunity for an age at which intervening on the use of heuristics could be particularly effective. In the tradition of previous experiments that have employed statistical reasoning interventions to improve adult judgments and decision-making (Fong, Krantz, & Nisbett, 1986), some intuitive hands-on training with these very young children could go a long way. Demonstrating to children the errors that can arise when base-rate information is not adequately considered could help children have better intuitions about everyday problems that involve weighing social and statistical information.

Chapter 6: References

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35*(5), 303-314. 10.1037/0022-3514.35.5.303

Arthur, A. E., Bigler, R. S., Liben, L. S., Gelman, S. A., & Ruble, D. N. (2008). Gender stereotyping and prejudice in young children: A developmental intergroup perspective. In S. R. Levy, & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66-86). New York, NY, US: Oxford University Press.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321-324. 10.1111/1467-9280.00063

Baker, S. T., Gjersoe, N. L., Sibielska-Woch, K., Leslie, A. M., & Hood, B. M. (2011). Inhibitory control interacts with core knowledge in toddlers' manual search for an occluded object. *Developmental Science, 14*(2), 270-279.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44*(3), 211-233. 10.1016/0001-6918(80)90046-3

Bascandziev, I., Powell, L. J., Harris, P. L., & Carey, S. (2016). A role for executive functions in explanatory understanding of the physical world. *Cognitive Development, 39*, 71-85. 10.1016/j.cogdev.2016.04.001

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16*(3), 162-166. 10.1111/j.1467-8721.2007.00496.x

Boseovski, J. J., Chiu, K., & Marcovitch, S. (2013). Integration of behavioral frequency and intention information in young children's trait attributions. *Social Development, 22*(1), 38-57. 10.1111/sode.12008

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology, 42*(3), 500.

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T. L., & Gopnik, A. (2016). Children's causal inferences from conflicting testimony and observations. *Developmental Psychology, 52*(1), 9. 10.1037/a0039830

Bryant, P., & Nunes, T. (2012). *Children's understanding of probability: A literature review (full report).* London: Nuffield Foundation.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053. 10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? contributions of inhibitory control and working memory. *Infant and Child Development, 11*(2), 73-92. 10.1002/icd.298

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153.

Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly (1982-),* , 328-346.

De Neys, W., & Vanderputte, K. (2011). When less is not always more: Stereotype knowledge and reasoning development. *Developmental Psychology, 47*(2), 432-441. 10.1037/a0021313

Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children’s causal inferences: The sampling hypothesis. *Cognition, 126*(2), 285-300. 10.1016/j.cognition.2012.10.010

Denison, S., Konopczynski, K., Garcia, V., & Xu, F. (2006). Probabilistic reasoning in preschoolers: Random sampling and base rate. In R. Sun and N. Miyake. (Eds.) *Proceedings of the 28th Annual Conference of the Cognitive Science Society.* (pp. 1216-1221).

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49*(2), 243-249. 10.1037/a0028278

Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798-803. 10.1111/j.1467-7687.2009.00943.x

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130*(3), 335-347. 10.1016/j.cognition.2013.12.001

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135-168. 10.1146/annurev-psych-113011-143750

Edelbrock, C. S., & Sugawara, A. I. (1978). Acquisition of sex-typed preferences in preschool-aged children. *Developmental Psychology, 14*(6), 614-623. 10.1037/0012-1649.14.6.614

Eisenberg, N., Murray, E., & Hite, T. (1982). Children's reasoning regarding sex-typed toy choices. *Child Development, 53*(1), 81-86. 10.2307/1129639

Falk, R., Yudilevich-Assouline, P., & Elstein, A. (2012). Children's concept of probability as inferred from their binary choices--revisited. *Educational Studies in Mathematics, 81*(2), 207-233. 10.1007/s10649-012-9402-1

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18*(3), 253-292. 10.1016/0010-0285(86)90001-0

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31-60. 10.1037/0033-2909.134.1.31

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*, 115-140.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about SocialCategories. *Psychological Science, 10*(6), 489-493. 10.1111/1467-9280.00194

Girotto, V., Fontanari, L., Gonzalez, M., Vallortigara, G., & Blaye, A. (2016). Young children do not succeed in choice tasks that imply evaluating chances. *Cognition, 152*, 32-39. 10.1016/j.cognition.2016.03.010

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*(1), 325-344.

Gonzalez, C. M., Zosuls, K. M., & Ruble, D. N. (2010). Traits as dimensions or categories? developmental change in the understanding of trait terms. *Developmental Psychology, 46*(5), 1078-1088. 10.1037/a0020207

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children. *Psychological Review, 111*(1), 3-32. 10.1037/0033-295X.111.1.3

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science, 17*(9), 767-773.

Gualtieri, S., & Denison, S. (2018). The development of the representativeness heuristic in young children. *Journal of Experimental Child Psychology, 174*, 60-76. 10.1016/j.jecp.2018.05.006

Gualtieri, S., & Denison, S. (2019). A comprehensive examination of preschoolers' probabilistic reasoning abilities. *Proceedings of the 41st Annual Conference of the Cognitive Science Society.*

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development, 77*(3), 505-524. 10.1111/j.1467-8624.2006.00886.x

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology, 69*, 251-273. 10.1146/annurev-psych-122216-011710

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology, 77*(1), 1-19. 10.1006/jecp.1999.2555

Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*(1), 166-178. 10.2307/1130712

James, W. (1890). *The principles of psychology* Harvard University Press.

Jones, E. E. (1979). The rocky road from acts to dispositions. *The American Psychologist, 34*(2), 107-117.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3*(1), 1-24. 10.1016/0022-1031(67)90034-0

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237.

Kalish, C. W., Kim, S., & Young, A. G. (2012). How young children learn from examples: Descriptive and inferential problems. *Cognitive Science, 36*(8), 1427-1448. 10.1111/j.1551-6709.2012.01257.x

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), 35.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*(1), 1-17. 10.1017/S0140525X00041157

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*(6), 1261-1277. 10.1111/j.1467-8624.2005.00849.x

Koenig, M. A., & Sabbagh, M. A. (2013). Selective social learning: New perspectives on learning from others. *Developmental Psychology, 49*(3), 399-403. 10.1037/a0031619

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General, 136*(3), 430.

Kuhn, D., Nash, S. C., & Brucken, L. (1978). Sex role concepts of two- and three-year-olds. *Child Development, 49*(2), 445-451. 10.2307/1128709

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*(8), 1134-1140. 10.1177/0956797610376652

Lane, J. D., Conder, E. B., & Rottman, J. (2019). The influence of direct and overheard messages on children's attitudes toward novel social groups. *Child Development,*

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development, 78*(5), 1543-1558. 10.1111/j.1467-8624.2007.01082.x

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition, 131*(2), 284-299. 10.1016/j.cognition.2013.12.010

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition, 120*(3), 403-411. 10.1016/j.cognition.2011.02.003

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British Journal of Psychology (London, England: 1953), 92*(Pt 1), 239-255.

Martin, C. L., & Little, J. K. (1990). The relation of gender understanding to children's sex-typed preferences and gender stereotypes. *Child Development, 61*(5), 1427-1439. 10.2307/1130753

Martin, C. L., & Ruble, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science, 13*(2), 67-70. 10.1111/j.0963-7214.2004.00276.x

Matthews, D. (2014). *Pragmatic development in first language acquisition* John Benjamins Publishing Company.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology, 49*(3), 404-418. 10.1037/a0029500

O'Grady, S., & Xu, F. (2019). The development of nonsymbolic probability judgments in children. *Child Development, 0*(0)10.1111/cdev.13222

Over, H., & McCall, C. (2018). Becoming us and them: Social learning and intergroup bias. *Social and Personality Psychology Compass, 12*(4), e12384. 10.1111/spc3.12384

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*(5), 1216-1226. 10.1037/0012-1649.43.5.1216

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 544-554. 10.1037/a0034887

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), 751-783. 10.1037/0022-3514.90.5.751

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children* (L. Leake, P. Burrell & H. D. Fishbein Trans.). Oxford, England: W.W. Norton.

Poulin-Dubois, D., & Brosseau-Liard, P. (2016). The developmental origins of selective social learning. *Current Directions in Psychological Science, 25*(1), 60-64. 10.1177/0963721415613962

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., & Call, J. (2014). Apes are intuitive statisticians. *Cognition, 131*(1), 60-68. 10.1016/j.cognition.2013.12.011

Riggs, A. E. (2019). Social statistics: Children use statistical reasoning to guide their inferences about the scope of social behavior. *Developmental Psychology, 55*(1), 66-79. 10.1037/dev0000618

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology, 69*(1), 181-203. 10.1146/annurev-psych-122216-011805

Schwarz, N., Strack, F., Hilton, D., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "Irrelevant" information. *Social Cognition, 9*(1), 67-84. 10.1521/soco.1991.9.1.67

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? causal inference and the development of social attribution. *Child Development, 84*(2), 443-454. 10.1111/j.1467-8624.2012.01865.x

Serbin, L. A., & Sprafkin, C. (1986). The salience of gender and the process of sex typing in three- to seven-year-old children. *Child Development, 57*(5), 1188-1199. 10.2307/1130442

Shutts, K. (2013). Is gender special? In M. R. Banaji, & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 297-300) Oxford University Press.

Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives, 9*(4), 262-266. 10.1111/cdep.12154

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society, 14*(1), 35-62. 10.1080/15248372.2011.638686

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*(4), 779-797. 10.1037/a0034191

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review, 31*(2), 103-118. 10.1016/j.dr.2011.07.003

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences, 104*(48), 19156-19159.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science, 18*(5), 259-263. 10.1111/j.1467-8721.2009.01648.x

Trautner, H. M., Ruble, D. N., Cyphers, L., Kirsten, B., Behrendt, R., & Hartmann, P. (2005). Rigidity and flexibility of gender stereotypes in childhood: Developmental or differential? *Infant and Child Development, 14*(4), 365-381. 10.1002/icd.399

Trope, Y., & Higgins, E. T. (1993). The what, when, and how of dispositional inference: New answers and new questions. *Personality and Social Psychology Bulletin, 19*(5), 493-500. 10.1177/0146167293195002

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science (New York, N.Y.), 185*(4157), 1124-1131. 10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). Evidential impact of base rates. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153-160). New York: Cambridge University Press.

Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science, 17*(3), 443-451. 10.1111/desc.12134

Vygotski, L. S. (1929). II. the problem of the cultural development of the child. *The Pedagogical Seminary and Journal of Genetic Psychology, 36*(3), 415-434.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*(1), 97-104. 10.1016/j.cognition.2009.04.006

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*(13), 5012-5015. 10.1073/pnas.0704450105

Yost, P. A., Siegel, A. E., & Andrews, J. M. (1962). Nonverbal probability judgments by young children. *Child Development, 33*(4), 769-780.

Zukier, H., & Pepitone, A. (1984). Social roles and strategies in prediction: Some determinants of the use of base-rate information. *Journal of Personality and Social Psychology, 47*(2), 349-360. 10.1037/0022-3514.47.2.349