August 2017

What Are The Cognitive Mechanisms That Underlie Our Theory Of Mind? Potential Insights From Information Theory

Nellie Kamkar
The University of Western Ontario

Supervisor
Adam S. Cohen
The University of Western Ontario

Graduate Program in Psychology

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

© Nellie Kamkar 2017

Follow this and additional works at: https://ir.lib.uwo.ca/etd
Part of the Child Psychology Commons, Cognitive Psychology Commons, and the Developmental Psychology Commons

Recommended Citation
https://ir.lib.uwo.ca/etd/4723

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlswadmin@uwo.ca.
Abstract

Theory of Mind (ToM) is the ability to infer mental states. The purpose of Study 1 was to reduce performance demands on a ToM test for forty (22 females) children ($M_{age} = 4.604; SD_{age} = 1.128$). Here, a low-uncertainty condition included a behaviour repetition manipulation, intended to increase success rate—but results did not confirm our hypothesis. Potential reasons for the results of Study 1 are discussed and tested in Study 2. The purpose of Study 2 was to determine the mechanism by which ToM operates in fifty-seven (26 females) adult participants ($M_{age} = 20.632; SD_{age} = 3.368$) by altering informational richness more directly. Results of Study 2 confirm that the mechanism by which ToM operates is via uncertainty reduction. These data motivate Study 3 in which child-appropriate vignettes will be used to address the limitations of Study 1 and implement the design of Study 2.

Keywords

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Adam S. Cohen who spent countless hours providing me with his guidance and support. Words cannot express my gratitude for your generosity, kindness, and continuous mentorship on this project. Thank you for challenging me when I needed to think critically, encouraging me when I felt frustrated, motivating me when I felt demoralized, and inspiring me when I needed hope. I will miss discussing research, politics, food, and everything in between with you over hours that after passing, felt like minutes with your enthusiasm.

Next, I would like to thank my mother—Shohreh Shahi—who, to this day, leads by example to show me that impossible isn’t a word. To the woman who decorated her bedroom not with posh paintings and decorative ornaments, but with my grade school tests of achievement—thank you for showing me the value of my academic efforts. To my father, Javad Kamkar, who had the unfortunate experience of teaching me advanced mathematics and calculus over several excruciatingly painful years. Thank you for refusing to give up until I reached the highest level of excellence and mastery in these now exciting topics.

And most importantly, to my sister, Niki Kamkar, you’ll never know how much you mean to me. Niki, my first word, my one and only role-model, my idol and my reason for everything—I owe this and every one of my accomplishments to you. I’ll forever be your biggest fan. Thank you for jumping for excitement during my successes, and for comforting me when $p > .05$. To my life partner, Navid Akmal-Sisan, thank you for standing by me in my lowest of lows and loving me at my absolute worst. You have taught me how to love. To my favorite couple, Moin and Jacquelline Alberto-Sisan—thank you for bringing me joy, cake, and participants.

To my labmate Isu Cho, for enriching my Master’s experience with your positivity, dedication, and eager helpfulness. When I struggled with computer coding, you sat beside me and cheerfully encouraged me to keep trying. When I needed fuel, you took me out for sushi, and when I needed friendship, I looked beside me and there you were.

I would now like to thank Marlee Salisbury and Ritika Bishnoi, my friends and research assistants, who worked tirelessly in collecting the first wave of data collection for Study 1. I will never forget your generosity when you took control of this project after I had my wisdom teeth removed. To Andrea Sandstrom, Hunster Yang, and Patricia Michaan, whose research contributions to this project as a whole, I cannot emphasize enough. Thank you for coming in (rain, shine, or Canada) to run participants for me. Thank you for studying for your courses in between participant sessions and for setting ridiculous 8am alarms to visit local preschools with me. Thank you for smiling even when you were tired, for pushing through even when you were hungry, and for continuously and eagerly volunteering to help me in every one of my research endeavors. A special thank you to Delaney Cosma who spent the Spring of 2017 running participants for me, alone in a windowless basement.

And finally, I would like to thank the families and teachers who volunteered their time to help me pursue our research. Without your kindness and generosity, this paper would simply not exist.
# Table of Contents

Abstract...........................................................................................................................................i

Acknowledgements............................................................................................................................ii

Table of Contents............................................................................................................................iii

List of Tables.....................................................................................................................................vii

List of Figures......................................................................................................................................viii

Chapter 1: General Introduction and Motivation Behind Study 1..................................................1

1 « When Do We Develop a ToM? »...............................................................................................2

1.1 « A Look at the Preschooler’s Mind-Reading Capacity ».......................................................2

1.2 « Fundamental Conceptual Change View ».............................................................................4

2 « Mind-Reading Infants? »...........................................................................................................7

2.1 « Substantial Continuity View »..............................................................................................9

3 « How Does ToM Develop? Reconceptualising the Puzzling and Controversial Findings of Previous Research »..........................................................................................................13

3.1 « The Competence-Performance Distinction ».......................................................................14

3.2 « Inhibitory Control as a Performance Demand »...................................................................15

3.3 « Verbal Fluency as a Performance Demand»..........................................................................16

3.4 « Making the Explicit False-Belief Test Easier: A Look at More Recent Cutting-Edge Research Findings ».........................................................................................................................17

3.5 « Putting it all Together »........................................................................................................19

4 « A Novel Theoretical Perspective: Information Theory and the Current Study ».................19

4.1 « Study 1: Reducing Children’s Uncertainty through Repetition »...................................20
4.2 « How the Uncertainty Reduction Manipulation Operates Mechanistically »……..21
4.3 « Developmental Change in the Information Processing System of ToM »………23
4.4 « Study 1: Significance and Hypotheses »……………………………………………24

Chapter 2: Study 1 Methods……………………………………………………………………26
5 « Participants »………………………………………………………………………………26
6 « Materials »………………………………………………………………………………..29
  6.1 « Study Design »………………………………………………………………………..29
7 « Procedure »………………………………………………………………………………..33

Chapter 3: Study 1 Results……………………………………………………………………36
8 « Descriptive Statistics »……………………………………………………………………36
  8.1 « Hypothesized Pattern of Data »……………………………………………………36
  8.2 « Actual Pattern of Data »………………………………………………………………37
9 « Inferential Statistics »……………………………………………………………………39

Chapter 4: Discussion of Study 1 and Motivation behind Study 2…………………………40
10 « Study 1: Main Findings »……………………………………………………………….40
  10.1 « A Deeper Analysis of Study 1: Potential Reasons for Null Effects »…………40
11 « Motivation behind Study 2 »……………………………………………………………44
  11.1 « Study 2: Design and Hypotheses »………………………………………………45

Chapter 5: Study 2 Methods………………………………………………………………….46
12 « Participants »………………………………………………………………………………46
13 « Materials »………………………………………………………………………………..46
  13.1 « Randomization Technique for Vignettes »……………………………………47
  13.2 « Test Questions »………………………………………………………………………48
<table>
<thead>
<tr>
<th>Page</th>
<th>Section Title</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>13.3 « The Standard Task »</td>
<td>49</td>
</tr>
<tr>
<td>49</td>
<td>13.4 « The Low-Uncertainty Task »</td>
<td>49</td>
</tr>
<tr>
<td>50</td>
<td>13.5 « The High-Uncertainty Task »</td>
<td>50</td>
</tr>
<tr>
<td>51</td>
<td>14 « Design Features of Study 2 »</td>
<td>51</td>
</tr>
<tr>
<td>51</td>
<td>14.1 « Technique for Topic Matching »</td>
<td>51</td>
</tr>
<tr>
<td>51</td>
<td>14.2 « Accuracy Measure »</td>
<td>51</td>
</tr>
<tr>
<td>52</td>
<td>15 « Procedure »</td>
<td>52</td>
</tr>
<tr>
<td>54</td>
<td>Chapter 6: Study 2 Results</td>
<td>54</td>
</tr>
<tr>
<td>54</td>
<td>16 « Descriptive Statistics »</td>
<td>54</td>
</tr>
<tr>
<td>54</td>
<td>17 « Inferential Statistics »</td>
<td>54</td>
</tr>
<tr>
<td>57</td>
<td>Chapter 7: Study 2 Discussion</td>
<td>57</td>
</tr>
<tr>
<td>57</td>
<td>18 « Study 2: Limitations »</td>
<td>57</td>
</tr>
<tr>
<td>59</td>
<td>19 « Study 2: Main Findings and Interpretation »</td>
<td>59</td>
</tr>
<tr>
<td>61</td>
<td>Chapter 8: General Discussion</td>
<td>61</td>
</tr>
<tr>
<td>61</td>
<td>20 « Study 1 and Study 2: A Brief Comparison »</td>
<td>61</td>
</tr>
<tr>
<td>62</td>
<td>21 « Future Research Goals: The Necessity of Study 3 »</td>
<td>62</td>
</tr>
<tr>
<td>62</td>
<td>21.1 « Study 3: Design »</td>
<td>62</td>
</tr>
<tr>
<td>64</td>
<td>21.2 « Study 3: Predictions and Significance »</td>
<td>64</td>
</tr>
<tr>
<td>65</td>
<td>22 « Concluding Remarks »</td>
<td>65</td>
</tr>
<tr>
<td>65</td>
<td>22.1 « A More In-Depth Discussion of Each Study »</td>
<td>65</td>
</tr>
<tr>
<td>66</td>
<td>22.2 « Revisiting the Competence-Performance Distinction and Information Theory »</td>
<td>66</td>
</tr>
</tbody>
</table>
22.3 « Final Thoughts: Information Theory Across Development » ...................... 69
References .............................................................................................................. 71
Appendices ............................................................................................................. 77
Curriculum Vitae .................................................................................................... 93
List of Tables

Table 1. Study 1 Participant Characteristics ................................................................. 28
Table 2. Props Used in Interactive Puppet Shows .......................................................... 30
Table 3. Example Script from a Randomly Selected Puppet Show ................................ 31
Table 4. Series of Events in Each Condition .................................................................. 32
Table 5. Four Cells that Constitute Child Proportions in False-Belief Tasks .................. 36
Table 6. Actual Proportions of Children in Each Cell .................................................... 36
Table 7. Results of the Pairwise Comparisons between the Three Conditions .............. 55
List of Figures

Figure 1. Bar plot illustrating the lack of difference between the proportion of children passing the S-FB test (43%), relative to those passing the LU-FB test (45%) ..........................39

Figure 2. Bar plot illustrating the main findings of Study 2........................................55
Chapter 1 : General Introduction and Motivation Behind Study 1

In 1978, David Premack and Guy Woodruff conducted one of the most highly cited and controversial studies of animal cognition. Specifically, they were interested in learning whether or not a chimpanzee has the capacity to reason about the unobservable mental states of others. This ability has been formally termed as a Theory of Mind (ToM)—or, more informally, as mind-reading—and it involves the ability to make inferences about other individuals’ mental states including their perceptions, emotions, desires, beliefs, and intentions. Such inferences are deemed a theory because they are not directly observable and thus, are inevitably made with some level of uncertainty (Premack & Woodruff, 1978). Since its introduction in 1978, a plethora of similar studies on ToM have followed. But as is made evident in the title of this paper, this is not an essay in animal ToM but one examining human mental state reasoning.

An adequately functioning ToM is integral to facilitating our basic daily social and communicative interactions, thus many researchers have sought to delineate when in ontogenetic human development we acquire a ToM (Rosnay, Fink, Begeer, Slaughter & Peterson, 2013; Slaughter, 2015). Therefore, the introductory section of the present paper will adhere by the following sequential order: First, series of experiments examining young, preschool-aged children’s mental state reasoning capacities will be reviewed—followed by a more recent body of infant literature studying ToM in a considerably younger population; next, these two bodies of literature will be conceptualized together into one, coherent framework by addressing past
accounts on this issue by scholars in the field. And finally, a novel theoretical account will be highlighted with the aim of motivating the present experiment.

1 « When Do We Develop a ToM? »

1.1 « A Look at the Preschooler’s Mind-Reading Capacity »

In order to discern when young children develop a ToM—first, researchers must develop a means to test ToM in said populations. To this end, several tests of mental state reasoning have emerged. In a recent review, Virginia Slaughter (2015) provided an overview of various tests of children’s’ mind-reading capacity along with the conceptual understanding underlying each test type. Because ToM is a broad, umbrella-term under which various mentalizing capacities fall—this review also demonstrates that as ToM tests increase in their difficulty, the proportion of preschool-aged children who pass sharply declines. At its simplest level, tests known as *diverse desires tests* assess children’s conceptual understanding that *not everyone* likes and wants—desires—the same thing. Not surprisingly, the majority of 3-7 year-old children pass these types of tests.

Tests assessing children’s understanding of not the desires of others—but the *beliefs* of others—are considerably more difficult. For example, tests known as *false-belief reasoning tests* evaluate children’s ability to recognize that people can sometimes be wrong in their beliefs about something, ergo they can possess what is known as a *false-belief*. People’s behavior when faced with a false-belief, while seemingly strange, is actually in line with their beliefs. Let’s consider a
real-world example of a false-belief scenario. If Professor Smith were to place her keys into her leather purse but unbeknownst to her—her 3-year-old son removes her keys and places them in the trash can at home—Professor Smith will likely search frantically in her purse for her keys. Although her behavior is not consistent with the true state of affairs—it is consistent with her belief, her mistaken, false, belief. Intriguingly, while some preschool-aged children are able to pass tests of false-belief reasoning ability, others struggle with such tests (Slaughter, 2015).

One of the most commonly used types of false-belief reasoning tests is known as the Sally-Anne Test (Baron-Cohen, Leslie & Frith, 1985). In this test, children are first presented with a series of characters and objects; namely, a character named Sally, a character named Anne, a marble (or ball), a box, and a basket. After their initial introduction, children are shown a sequence of events in the form of a puppet show. First, Sally places the marble into the basket. Then, Sally leaves the room and during her absence, Anne moves the marble from the basket to the box—after which Anne leaves. And finally, Sally returns and children are asked the following critical question: “Where will Sally look for the marble?” in an effort to engage their ability to make an accurate prediction about Sally’s behavior.

Of course, we can clearly see that children are presented with a dichotomous choice—they can select the basket as the appropriate answer, or they can select the box. If children select the basket as the location in which Sally will look for the marble, then they pass and are said to have a ToM—or more precisely, the ability to reason about another individual’s false-beliefs. If, on the other hand, children select the box as the location in which Sally will look for the marble,
then they fail and are said to lack this ability. With this test, Baron-Cohen et al. (1985) found that four-year-old healthy, neurotypical children were able to pass this test, setting the foundation for a series of experiments corroborating the view that ToM ability undergoes a genuine and fundamental shift in early childhood.

1.2 « Fundamental Conceptual Change View »

Subsequent experiments on children’s false-belief reasoning used this test—the Sally-Anne test—or some similar variant of it, thus rendering it one of the most widely used tests of ToM (Slaughter, 2015). These types of ToM tests, of course, are not without flaws and indeed, have been subject to a significant and justified amount of scrutiny (Bloom & German, 2000). One such criticism is the mere fact that reasoning about mental states encompasses so much more than simply predicting people’s behavior on the basis of their mistaken beliefs. Despite the logical validity of arguments such as this, the Sally-Anne test is still a widely used test of ToM (Bishop-Fitzpatrick, Mazefsky, Eack & Minshew, 2017; Lábadi & Beke, 2017; O’Toole, Monks & Tsermentseli, 2017).

But two years prior to the development of the Sally-Anne test, a similar type of false-belief test was used by Heinz Wimmer and Josef Perner in 1983. Much like the Sally-Anne test, a character named Maxi (analogous to Sally) places chocolate (analogous to the marble) into cupboard x (analogous to the basket). In his absence, his mother (analogous to Anne) places the chocolate into cupboard y (analogous to the box). Children then must indicate where Maxi will
look for the chocolate. The results of this study demonstrated an interesting developmental shift whereby *none* of the 3-4 year-old children passed, 57% of 4-6 year-old children passed, and 86% of 6-9 year-old children passed. Based on these results, it was concluded that the ability to represent mental states *emerges* during the late preschool years (approximately 4-6 years-of-age; Wimmer & Perner, 1983).

Some years later, further evidence in support of the view that younger preschoolers lack a ToM (i.e. have a *conceptual deficit* in ToM) came with a slightly different type of false-belief reasoning test (Perner, Leekam & Wimmer, 1987). The false-belief tests discussed thus far are tests examining children’s ability to represent a character’s false belief about the *location* of an object. Therefore, they are typically referred to as *location false-belief tests*. But we can have an incorrect belief about other situations such as the contents of a container. This is precisely what Perner et al. (1987) tested. In addition to a location false-belief test, this study included a *contents false-belief test* in which children were brought into the laboratory with a friend. While the friend waited outside the room, an experimenter showed the child a box with the “Smarties” chocolate logo. When asked what they think is in the box, most children answered correctly, exclaiming “chocolate!” or “Smarties!” but to their surprise—the experimenter showed them that there was a pencil inside it. The test question was seemingly simple: “What will [name of friend] think is in here?” Again, the response options here are binary such that children can state one of two possible answer choices: Smarties chocolate, or a pencil. Consistent with the conceptual deficit view of ToM development, Perner et al. (1987) found that the majority of younger
preschoolers (3 year-olds) failed both location and contents false belief tasks. For both task types, younger children claimed that the agent will think that the location or contents of an object are consistent with reality. For what is now termed The Smarties Task, younger children consistently reported that their friend will think that there is a pencil in the box, despite the fact that she was outside of the room when the pencil was revealed to the child. The authors of this paper interpret these findings as indicating that “it is only when children are older than 3 years that they can assign conflicting truth values” (Perner et al., 1987, p. 136).

These initial studies sparked the inception of a myriad of similar studies in the 1990’s. In fact, the field of ToM development flourished so much that in 2001, Henry Wellman, David Cross, and Julanne Watson conducted a meta-analytic review of 178 separate studies. Their findings serve as one of the most robust and cogent arguments in support of the conceptual deficit view. Irrespective of the differences in task types used across various studies—Wellman et al. (2001) found the same typical pattern across development. Specifically, younger children performed at below-chance, whereas their older preschool-aged counterparts performed well above-chance. There is thus a clear discontinuous change across development such that in a seemingly sudden acquisition of mental state reasoning—4-5 year-old children undergo a fundamental change in their capacity to represent others’ mental states. A second meta-analysis several years later confirmed these findings with analyses on studies of Chinese children (Liu, Wellman, Tardif & Sabbagh, 2008). This review included data from over 3,000 children and found that Chinese children also progress in a similar manner to Anglo-European children
studied in previous reports. Like the findings of the 2001 meta-analysis with Anglo-European children, younger Chinese preschoolers performed below chance on tests of false-belief reasoning. These parallel findings highlight the universal developmental trajectory of false-belief reasoning. Furthermore, they substantiate the view that an understanding of the beliefs of others—and arguably, an understanding of the minds of others—undergoes a genuine and fundamental conceptual change in the early preschool years (Liu et al., 2008). Based on the research presented thus far, it is apparent that a fully-functioning mind-reading system emerges when we are approximately 4-5 years-of-age. But is this the current consensus on ToM development amongst researchers of early social development today? In recent years, a cascade of new findings have emerged that bring into question the once widely held view that early in life, our social cognitive capacity was marked by a deficit in attributing mental states to others.

2 « Mind-Reading Infants? »

In 2005, Kristine Onishi and Renée Baillargeon conducted one of the first and most groundbreaking studies of infant belief tracking. With a deceptively simple and unique paradigm—this study came to challenge and critique our previously held conceptions of ToM development. The researchers tested 15-month-old infants on their ability to track beliefs—and more specifically, false beliefs. If their results indicate that infants can, indeed, track the beliefs of others, then this would bring to question the genuine and discontinuous conceptual change view of ToM development. Onishi and Baillargeon (2005) presented infants with a series of
events much like those presented in the Sally-Anne test. Conceptually, the scenarios are very similar, as one character is presented, believing an object to be located in one location and during her/his absence—the object is moved to another location. Their paradigm involves an initial introduction during which infants are shown a woman facing a desk onto which a watermelon, a yellow box, and a green box are located. Then, the woman places the watermelon into the green box several times in what are termed familiarization trials. These trials make it clear to the infant that she believes the watermelon to be located in the green box. In false-belief trials, the woman leaves and during her absence, the watermelon magically moves from the green box to the yellow box. In a traditional false-belief task given to preschoolers, when the woman returns, children would be asked where she will look for the watermelon.

However, because many 15-month-old infants are preverbal—Onishi and Baillargeon (2005) were required to obtain a non-verbal response. To this end, they included two versions in which they completed the story. In the first version, they had the woman search in the green box for the watermelon. This, of course, is an expected course of action because the green box is the location in which she believes the watermelon to be. But in the second version, they had the woman search in the yellow box for the watermelon. While it is true that this is the true location of the watermelon—critically, it is not where the woman believes the watermelon to be. Therefore, this is an unexpected course of action because, as an adult seeing such an event, we might wonder: “Why is she looking in the yellow box? How does she know that the watermelon is in there?” This is precisely the logic of their study—a paradigm termed a Violation of
*Expectation (VOE) Paradigm* because in some conditions, the infant’s expectations are violated. It should then logically follow that infants who view a VOE condition should be surprised and thus, look longer during these conditions (relative to conditions during which their expectations are not violated).

This pattern of findings is precisely what Onishi and Baillargeon (2005) found. Remarkably, infants looked longer when the woman searched in the yellow box (the location in which she *does not believe* the watermelon to be) than when she searched in the green box (an action consistent with her beliefs) for the watermelon. These findings suggest that infants expect people to behave in ways that are consistent with their beliefs—including their false-beliefs. More broadly, these results indicate that rather than a fundamental and discontinuous shift from lacking a ToM to suddenly possessing a ToM by our fourth birthday—ToM is present early in life and exhibits a continual developmental trajectory from infancy to early childhood.

### 2.1 « Substantial Continuity View »

In a follow-up experiment using a similar VOE paradigm, a separate team of researchers tested 15-month-old infants on their ability to track the beliefs of others (Träble, Marinović & Pauen, 2010). Here, infants saw a similar series of events. First, a woman was situated in front of a ball and two containers. The woman believed the ball to be in one container but during her absence, it was moved to the other container. One critical difference between this paradigm and that of Onishi and Baillargeon (2005) was that rather than having the watermelon magically
move from one location to the other—both containers were balanced on a beam (much like a teeter totter seasaw in a children’s playground). Unbeknownst to the woman, the plank holding both containers shifted and thus, the ball rolled from the container that she believed it to be in to the other container. Like the findings of Onishi & Baillargeon (2005), Träble et al. (2010) found that infants looked reliably longer in VOE conditions than conditions in which the woman behaved in a way that was consistent with her beliefs. These results add to the flourishing literature in support of the view that false-belief understanding is present early in life and continually develops across ontogeny.

In a series of experiments adding further support for this substantial continuity view and against the fundamental conceptual change view, yet another group of researchers tested the infant’s ability to track mental states (Surian, Caldi & Sperber, 2007). This study also implemented a VOE paradigm but here, rather than facing the agent as was done in previous studies (i.e. the infant faced the woman), infants shared the field of view of the agent. Once again, results suggested that infants expect that the agent’s search behavior to be consistent with his/her beliefs (Surian et al., 2007).

At this juncture, one may reasonably argue that regardless of the internal validity of such experimental designs, in the real-world, when we encounter a situation in which we expect an object to be located in a given container and we find that it is missing, we should show some type of emotional reaction. In other words, we would naturally be surprised upon finding that we are mistaken about the location of an object—much like Professor Smith in the example presented at
the beginning of this chapter would be surprised to learn that her keys are missing from her purse. Thus in a recent series of experiments, Rose Scott (2017) included the emotional consequences of possessing a false-belief in a VOE paradigm with infants. Consistent with what you or I would presume, infants expected the agent to display the emotional expression of surprise upon encountering a false-belief. These results further substantiate the view that infants possess a rudimentary but robust capacity to reason about the mental states of others.

But in addition to VOE paradigms, a different type of looking paradigm is often used with infants. Namely, the *predictive or anticipatory looking paradigm* during which the infant’s spontaneous looking behavior is measured prior to the end of the typical locations false-belief scenario. In a study using an anticipatory looking paradigm, Southgate, Senju & Csibra (2007) tested infant’s false-belief tracking capacity. After presenting infants with an agent who holds a false belief about the location of an object, the researchers measured infants’ spontaneous looking using an eye tracker. They found that infants reliably looked at the location in which the agent *believes* the object to be located—rather than where it is actually located. This pattern of behavior suggests that infants anticipate the subsequent behavior of an agent who holds a false belief (Southgate et al., 2007). Similar findings have even been observed in an anticipatory looking paradigm in which rather than including an agent with beliefs—the researchers had simple geometric shapes who “had beliefs” (Surian & Geraci, 2012). Astonishingly, 17-month-old infants’ predictive looking reveals that they even spontaneously attribute beliefs to inanimate
objects, anticipating that the triangle will search for the disk in the location that s/he believes it to be.

The evidence so far—while compelling and robust—is based on the inferences and attributions that we adults make from observations of infants’ looking behavior. Although the researchers may have postulated a-priori hypotheses regarding infants’ looking behaviour— inferences of higher order cognition drawn from mere looking time data must be made with great caution. This, of course, is not a criticism specific to the infant literature on ToM development but may generalize to studies of infant cognition across various subjects and domains. In a recent study that, at least partly, addresses this criticism, infants were video-recorded as they watched a VOE false-belief paradigm (Moll, Khalulyan & Moffett, 2016). Interestingly, when infants viewed the actor behave in a way that is inconsistent with her beliefs (ergo a violation of the infant’s expectation), they were judged by independent raters to themselves display more facial tension. Now, the researchers are not simply coding looking times as indicative of a deeper cognitive function—but are assessing the infant’s emotional and facial display of suspense in response to the stimuli presented. Taken together, the data presented thus far make a cogent and compelling case for the existence of early false-belief reasoning in infants.
3 « How Does ToM Develop? Reconceptualising the Puzzling and Controversial Findings of Previous Research »

But the data thus far also present somewhat of a puzzle: If infants possess the ability to track, expect, anticipate, and respond to the false-beliefs of others—then why do early data with preschoolers demonstrate a conceptual deficit in 3-year-olds? To this very day, intense controversy surrounds the issue of when we develop the capacity to reason about the mistaken beliefs of others. Indeed, many researchers are still puzzled by the contrast between the growing body of infant literature on belief tracking and the initial research findings that younger preschoolers consistently fail tests of false-belief reasoning (Scott & Baillargeon, 2017). But perhaps part of the problem is that we are not asking the appropriate questions that may generate a cohesive answer to the conundrum of human ToM development. Rather than asking the question of when we develop a ToM—we should be asking the question of how we develop a ToM. That is, what are the cognitive mechanisms that underlie our capacity to mind-read? A more fruitful set of questions can be generated by taking a closer look at children’s performance on tests of false-belief reasoning such as the Sally-Anne test. For preschool-aged children who successfully pass false-belief reasoning tests—why do they pass? Is it because of a genuine mentalizing capacity? Or, are children deploying some other, non-mentalistic strategy to successfully pass? More importantly, when children fail tests of false-belief reasoning, why do they fail? Is it because they sincerely lack the ability to attribute mental states? In other words, is
it because they lack ToM competence? Or, is it possible that children who fail false-belief reasoning tests are faced with a series of performance demands that impede and hinder their ability to demonstrate their mind-reading capabilities?

3.1 « The Competence-Performance Distinction »

This distinction has been formalized as the Competence-Performance Distinction (Bloom & German, 2000; Chomsky, 1965; Surian & Leslie, 1999; Wellman et al., 2001; Yazdi, German, Defeyter & Siegal, 2006). Put simply, this distinction asserts that when completing any given task, there are at least two factors that predict success: the first is competence in that given area—in this case, a ToM; and the second is a series of other, non-focal task demands that interfere with successful performance. In false-belief reasoning given to preschoolers (like the Sally-Anne test), one of the most well-established critiques of it as a measure of ToM is that passing such tests often requires so much more than a ToM. That is, a series of performance demands often interfere and prevent children from demonstrating their true mindreading capabilities (Bloom & German, 2000). For example, demands on children’s inhibitory control as well as verbal and linguistic demands inherent in the task may impede and hinder their ability to demonstrate their true ToM.
3.2 « Inhibitory Control as a Performance Demand »

Regarding inhibitory performance demands, there is a strong and robust correlation between children’s performance on tests of inhibitory control and on tests of ToM (Carlson, Moses & Breton, 2002; Carlson, Moses & Claxton, 2004). In addition to these correlational studies, direct manipulations of inhibition have underscored the importance of inhibitory control when completing false-belief reasoning tests (Leslie, German & Polizzi, 2005). Here, rather than designing the scenario in such a way that Sally wants to approach the ball—Alan Leslie and his colleagues created a scenario during which the agent wanted to avoid the object (i.e. it made the agent sick). This deceptively simple manipulation increases the inhibition demands on performance, as it is more cognitively demanding because of the added computation inherent in the task. With this avoidance-desire task, the majority of four-year-old children who were able to pass the standard approach-desire false-belief reasoning test failed (Leslie et al., 2005). But inhibition is only one such performance demand that hinders children’s ability to succeed on the false-belief test. Indeed, in the reviewed research that follows, it should become apparent that manipulations that alter subtle—yet important—verbal aspects of the task can drastically alter children’s performance.
3.3 « Verbal Fluency as a Performance Demand »

In an interesting set of studies, a manipulation on the test question that children were asked yielded some fascinating results (Surian & Leslie, 1999). Recall that in the original Sally-Anne test, children are asked “Where will Sally look for the ball?” after which their response determines whether or not they are said to pass the test. In one of the most widely cited manipulations of the test, children were asked “Where will Sally look for the ball first?” (Surian & Leslie, 1999). Now, an overwhelming majority of 3-year-old children (who were once believed to lack a ToM) passed when given the look first test question. These findings confirm that young preschoolers’ difficulties on tests of ToM are largely due to linguistic performance factors inherent in the tasks themselves.

Despite the compelling evidence provided by Surian and Leslie in 1999—the 2001 meta-analysis by Henry Wellman and his colleagues—concluded otherwise. Recall that the results of this meta-analysis suggested that a genuine conceptual deficit of ToM exists in the young preschooler. Wellman et al. (2001) addressed manipulations aimed at reducing performance demands in tests of false-belief reasoning by stating that they disproportionately help older preschoolers and thus cannot explain deficits in younger cohorts. These findings render it possible, then, that young preschoolers fail not because of performance demands, but because of a genuine conceptual deficit (Wellman et al., 2001). But a few years later, a more in-depth investigation of manipulations such as the look-first manipulation revealed that it is not
necessarily true that this manipulation selectively benefits older children and that younger children can improve with it (Yazdi et al., 2006). Taken together, the results of the Surian and Leslie (1999) and Yazdi et al. (2006) studies strongly suggest that performance demands in the design of the false-belief reasoning tests are possible culprits for the young preschooler’s inability to pass. In recent years, this perspective has received more scientific support and is now one of the dominant and prevailing views of ToM development.

### 3.4 « Making the Explicit False-Belief Test Easier: A Look at More Recent Cutting-Edge Research Findings »

Given that 3-year-olds are just beginning to learn language, some aspects of the test question may be confusing (Rubio-Fernández & Geurts, 2013). For example, in the test question, “where will Sally look for the ball?”—the word will may be conflated with should, rendering the answer given to the question to be viewed as false. In reality, Sally should look in the location in which the ball really is, but she would look in the container in which she believes it to be. To help address the verbal ambiguity of the test question—the researchers of this study removed the test question altogether. Instead, after the series of events were presented, the experimenter simply handed the agent doll (i.e. Sally) to the child and asked her/him to complete the story. To reduce any confusion on the part of the child, they asked the following two questions, “what happens next?” and “what is she going to do now?” and found that 3-year-olds are now able to pass. Clearly, replacing the verbally demanding and cryptic canonical test question with more
suitable ones helps reduce performance demands for young preschoolers and enables them to demonstrate their true ToM (Rubio-Fernández & Geurts, 2013).

In a follow-up experiment published last year, Paula Rubio-Fernández and Bart Geurts (2016) found that the perceptual salience and representational strength of the marble in the Sally-Anne test may place performance demands on young children. Therefore, they recommend not mentioning the marble at all and instead, urge researchers to encourage young children to focus on the agent (i.e. Sally). To this end, they asked preschoolers questions like “where will Sally go now?” and found that most 3-year-olds are now able to pass (Rubio-Fernández & Geurts, 2016). Another recently published study also found support for the processing-demands view in which younger children’s failure on false-belief reasoning tests may be an artifact of performance-related task demands (Setoh, Scott & Baillargeon, 2016). Here, the authors argue that one reason why infants can pass false-belief reasoning tests while their older 3-year-old counterparts fail is because of some key differences between the task types given to infants and preschoolers.

Critically, infants are given non-traditional, implicit false-belief reasoning tests in which their implicit, non-verbal responses (i.e. looking behavior) are measured as a basis of their success rate. In contrast, preschoolers are given traditional, explicit false-belief reasoning tests in which they must produce a verbal response that itself may place cognitive demands on the child and hinder their performance (Setoh et al., 2016). To reduce such performance demands, the researchers of this paper provided young children with several opportunities to practice giving verbal responses to questions in a series of practice trials. Even with the cryptic behaviour-
prediction question of “where will Sally look for the ball?”—Setoh et al. (2016) found that following the added response-generation practice trials, even 2 and a half year old toddlers were able to pass!

### 3.5 « Putting it all Together »

At this juncture, it should be evident that the seemingly contradictory results of the preschool and infant literature can be resolved by taking into account the competence-performance distinction. In effect, these apparently puzzling bodies of evidence are simply reflective of the performance demands inherent in false-belief reasoning tests given to preschoolers. Taken together, these data overwhelmingly support the *substantial continuity view* and bring into question the *fundamental conceptual change view*.

### 4 « A Novel Theoretical Perspective: Information Theory and the Current Study »

But in addition to the substantial amount of evidence suggesting that other, non-focal skills (such as inhibition and verbal fluency) are necessary to pass, an account informed by information theory can further explain why younger children consistently struggle. From an information-theoretic perspective (Balsam & Gallistel, 2009), ToM processing generates a series of possible representations of another agent’s mental state, each with varying levels of certainty (Leslie, Friedman, & German, 2004). By returning to the Sally-Anne test, we can clearly see that
children are faced with a dichotomous choice—the basket, or the box. For each response option, children possess a specific level of certainty. If their level of uncertainty regarding the agent’s belief that it is in, say, the basket is high, then they will not choose that response option. From this standpoint, young children who fail false-belief tests may simply require more information to reduce their uncertainty about the character’s mental state (i.e., Sally’s belief about the ball’s location). In the standard FB task, Sally’s hiding of the ball into a basket only occurs once after which a series of other events occur, so it is possible that the amount of information pertaining to Sally’s belief about the ball’s location is inadequate. Children are required to make a mental state inference on the basis of one and only one piece of information; namely, that Sally put the ball into the basket. It is perhaps not surprising, then, that for the young child—making an inference about Sally’s beliefs is difficult. In essence, the impoverished nature of the information provided to children may be a critical factor that places performance demands on children. Consequently, we must enrich the amount of information that children are given, therefore reducing their level of uncertainty regarding the agent’s belief in standard tests of false-belief reasoning.

4.1 « Study 1: Reducing Children’s Uncertainty through Repetition »

Based on these considerations, we conducted an experiment that systematically manipulated the child’s uncertainty regarding another’s beliefs. Others have manipulated the story’s ambiguity by increasing the child’s uncertainty regarding the agent’s behavior and have
found that such manipulations undermine children’s performance (Scott & Roby, 2015). But in the current study, we wanted to know whether or not reducing children’s uncertainty and as such, the performance demands in a false-belief reasoning test will allow preschoolers who typically fail standard false-belief tasks to pass. In addition to the standard false-belief task (S-FB), we included a low-uncertainty false-belief task (LU-FB) during which Sally repeatedly hid the ball into one location, thereby increasing the child’s certainty regarding Sally’s belief of the location of the ball. With each added repetition during which Sally places the ball in a given location—the child should become more certain that this, indeed, is the location in which she believes the ball to be located.

4.2 « How the Uncertainty Reduction Manipulation Operates Mechanistically »

Representationally, the belief in the child’s mind is strengthened by the low-uncertainty manipulation via repetition. That is, the child first possesses no mind-reading hypotheses and with incoming information, the ToM computational system generates a set of possible mental states in an abstract hypothesis space. Much like Bayesian models of cognitive reasoning (Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Perfors, Tenenbaum, Griffiths, & Xu, 2011), incoming information alters the probability of each hypothesis being the correct one underlying the observation. This probability index is subjective and thus constitutes a certainty index that attaches to each belief representation and is updated based on the behavioural cues that
the child observes (i.e. the data input). The manipulation in the low-uncertainty condition encompasses a scenario during which the agent repeatedly engages in a behaviour in an effort to increase the certainty index attached to the correct belief hypothesis. Put more simply, the child first possesses no belief hypotheses pertaining to Sally’s belief about the location of the ball. After being presented with the standard false-belief scenario, the child’s ToM computational system consists of two possible hypotheses regarding Sally’s belief of the ball’s location: 1) Sally believes that the ball is in the basket; and 2) Sally believes that the ball is in the box. In the standard test, 1) is the correct answer and for a child to explicitly state this—their certainty index attached to hypothesis 1) should be fairly strong. In contrast, hypothesis 2) is the incorrect response but it is, indeed, the actual location of the ball. For the young child who fails such tests and responds with hypothesis 2) as representing Sally’s belief—this may be because of the perceptual salience of actual location of the ball, memory demands inhibiting the child’s ability to remember earlier events, or a combination of these possibilities. In the low-uncertainty condition of the present study—because Sally repeatedly places the ball into the basket—with each repetition, the child’s certainty index attached to hypothesis 1) should be strengthened. Thus in hypothesis space, there will be a shift towards hypothesis 1) as representing Sally’s belief because the child’s certainty regarding hypothesis 1) should be enhanced. This shift is expected to be a direct result of the repetition manipulation that facilitates successful processing of the explicit ToM system.
4.3 Developmental Change in the Information Processing System of ToM

Adequate processing of the explicit system requires an enriched information source and a significantly reduced level of uncertainty. The implicit system, on the other hand, does not operate in a similar fashion. Instead, it can make automatic mental state inferences using simple bottom-up processes that do not exert the executive resources that the explicit system exerts. In infancy, the implicit system is at play, thus rendering information enrichment a less relevant factor in its processing. But even if the implicit system requires a rich information source—it is often provided, as the overwhelming majority of infant studies include several familiarization trials during which the infant's certainty is increased via repetition.

In early childhood, the explicit system is typically tested. This system requires a myriad of cognitive abilities including working memory abilities, verbal comprehension and fluency skills, and inhibitory control (as outlined by performance accounts). Successful explicit belief-reasoning also requires information richness, as delineated by information theory. Thus in early childhood, a predeveloped explicit ToM system requires a highly enriched information source. By adulthood, this explicit system becomes better developed and can generate accurate inferences more readily. Here, a more responsive information updating system operates and again, because explicit responses are required, information richness plays a role in mental state inferences.
The five year old who passes explicit tests has now developed a more responsive and generative information updating system that can operate in spite of a relatively impoverished information source. In contrast, the three year old who fails requires a much richer information source to pass.

4.4 « Study 1: Significance and Hypotheses »

By applying an information-theoretic perspective, this study brings into question the impoverished nature of the information provided to children in standard false-belief tests. Relatedly, the study enriches the amount of information supporting a mental state inference, which will not only help delineate when children acquire a ToM—but critically, elucidate how, mechanistically, ToM operates. More broadly, this approach to studying ToM in children will explain a significant amount of prior research in one coherent framework, ultimately adding to the plethora of research findings that resolve the seemingly contradictory infant and preschool data.

In light of the features of Information theory outlined above, it is expected that ToM operates through gradual uncertainty reduction, rendering the repetition manipulation of the LU - FB test a theoretically appropriate and well substantiated manipulation on children’s performance. The main hypotheses of the current study are simple:
1) A large proportion of children who fail the S-FB test will pass the LU-FB test.

2) A considerably smaller proportion of children will show the reverse pattern (passing the S-FB test but failing the LU-FB test).

Please note that we also included true-belief conditions in Study 1 but our main hypothesis was pertaining to false-belief conditions. For a deeper explanation of true-belief conditions and their utility in past studies as well as the present study, please see Appendix A.
Chapter 2 : Study 1 Methods

5  “Participants”

Eighty-two (45 females) typically developing preschool-aged children between 3 to 6 years of age ($M = 4.199; SD = 1.077$) were recruited to participate in the current study. Children were recruited both by contacting their parents via telephone from Western University’s Developmental Participant Pool and by visiting local preschools in the London and surrounding area. Children were excluded from the analyses under the following conditions:

1) If they had a neurological or psychological condition that impeded their ability to complete the tasks. Examples include delays in verbal comprehension or verbal fluency that may hinder their ability to understand the task and/or to respond to the questions asked of them;

2) If they had a language barrier and no suitable translators were present to help them understand the tasks and/or the questions asked of them;

3) If they were unable to sit through all puppet shows (for false-belief analyses the two relevant puppet shows were the S-FB condition and the LU-FB condition; for true-belief analyses the two relevant puppet shows were the S-TB condition and the LU-TB condition);

4) If they answered either of our control questions incorrectly. Specifically, we included two control questions to ensure that children were attending to and ultimately able to
understand the series of events presented to them. The first question was a memory question in which children were asked “Where did Sally put the ball?” in order to assess children’s memory of Sally’s behaviour. The second was a reality question in which children were asked “Where is the ball now?” that assessed children’s understanding of the current state of affairs. Only children who passed both memory and reality questions were included in the final analyses.

Of the eighty-two children tested, seventy children were able to sit through both false-belief conditions and sixty-four children were able to sit through both true-belief conditions. With respect to the false-belief conditions, forty (22 females) children ($M_{age} = 4.604; SD_{age} = 1.128$) answered both memory and reality questions correctly and thus, were included in the final analysis. For true-belief conditions, thirty two (17 females) children ($M_{age} = 4.661; SD_{age} = 1.253$) answered both memory and reality questions correctly and thus, were included in the final analysis. For a simpler depiction of the participant characteristics, please see Table 1 (below).
Table 1. Study 1 Participant Characteristics

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Children Recruited</th>
<th>Children Able to Complete Both Standard and Low-Uncertainty Tasks</th>
<th>Children Who Answered Both Memory and Reality Questions Correctly (Final Sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>False-Belief</strong></td>
<td>N = 82</td>
<td>N = 70</td>
<td>N = 40</td>
</tr>
<tr>
<td></td>
<td>45 Females</td>
<td>40 Females</td>
<td>22 Females</td>
</tr>
<tr>
<td></td>
<td>$M_{\text{age}} = 4.199$</td>
<td>$M_{\text{age}} = 4.271$</td>
<td>$M_{\text{age}} = 4.604$</td>
</tr>
<tr>
<td></td>
<td>$SD_{\text{age}} = 1.077$</td>
<td>$SD_{\text{age}} = 1.087$</td>
<td>$SD_{\text{age}} = 1.128$</td>
</tr>
<tr>
<td><strong>True-Belief</strong></td>
<td>N = 82</td>
<td>N = 64</td>
<td>N = 32</td>
</tr>
<tr>
<td></td>
<td>45 Females</td>
<td>38 Females</td>
<td>17 Females</td>
</tr>
<tr>
<td></td>
<td>$M_{\text{age}} = 4.199$</td>
<td>$M_{\text{age}} = 4.268$</td>
<td>$M_{\text{age}} = 4.661$</td>
</tr>
<tr>
<td></td>
<td>$SD_{\text{age}} = 1.077$</td>
<td>$SD_{\text{age}} = 1.125$</td>
<td>$SD_{\text{age}} = 1.253$</td>
</tr>
</tbody>
</table>
6 « Materials »

6.1 « Study Design »

The study design was that of a 2 (condition: standard, low-uncertainty) x 2 (belief status: true, false) with-subjects design. All children were shown a total of four puppet shows—the S-FB task, the LU-FB task, the S-TB task, and the LU-TB task in no particular (random) order.

To account for the possibility that children may become bored or that story scenarios will transfer across conditions—we introduced new characters and props for each puppet show. Our goal was to maintain the child’s interest and so, each puppet show included a slightly different—albeit theoretically similar—storyline. Character names were generated by using an online baby name generator and objects were every-day, household objects that children typically interact with. We included both male and female puppet characters to ensure that children of both genders can identify with all puppets. For a complete list of all characters and props used in Study 1, please see Table 2 (below).
Table 2. Props Used in Interactive Puppet Shows

<table>
<thead>
<tr>
<th>Character Names</th>
<th>Objects</th>
<th>Containers (Locations)</th>
<th>Task that First Character Engages in While Second Character Moves the Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sally</td>
<td>Ball</td>
<td>Basket</td>
<td>“...takes a nap”</td>
</tr>
<tr>
<td>Blake</td>
<td>Car</td>
<td>Treasure Chest</td>
<td>“...goes for a walk”</td>
</tr>
<tr>
<td>Nina</td>
<td>Hair Brush</td>
<td>Barrel</td>
<td>“...goes to the park”</td>
</tr>
<tr>
<td>Peter</td>
<td>Crayon</td>
<td>Wooden Box</td>
<td>“...brushes his teeth”</td>
</tr>
<tr>
<td>Lilly</td>
<td>Guitar</td>
<td>Cup</td>
<td>“...takes a bath”</td>
</tr>
<tr>
<td>Tom</td>
<td>Block</td>
<td>Little Box</td>
<td>“...eats some food”</td>
</tr>
<tr>
<td>Anne</td>
<td></td>
<td>Blue Box</td>
<td></td>
</tr>
<tr>
<td>Ron</td>
<td></td>
<td>Green Box</td>
<td></td>
</tr>
</tbody>
</table>

For each puppet show, the following were selected: two characters, one object, two locations, and a given task that the first agent completes while the second agent moves the object. To ensure that research assistants interacting with children remembered the series of events and to reduce the possibility of human error—they were provided with a generic script that they committed to memory. For an example of the specific script used, please see Table 3 (below).
Table 3. Example Script from a Randomly Selected Puppet Show

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Script</th>
<th>Test Questions (in order)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-Uncertainty False-Belief Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>a)</strong> This is Sally [wave Sally] and this is Anne [wave Anne]. This is their ball [have both Sally and Anne touch ball simultaneously]. This is their box [have both Sally and Anne touch the box simultaneously] and this is their basket [have both Sally and Anne touch the basket simultaneously].&lt;br&gt;&lt;br&gt;<strong>b)</strong> Look at Sally! Sally takes the ball and puts it into the basket [have Sally do this; play out other events that follow as well].&lt;br&gt;&lt;br&gt;<strong>c)</strong> Then, Sally takes the ball out of the basket.&lt;br&gt;&lt;br&gt;<strong>d)</strong> Repeat <strong>b)</strong> and <strong>c)</strong> until Sally has placed the ball into the basket 3 times.&lt;br&gt;&lt;br&gt;<strong>e)</strong> Then, Sally gets very sleepy and leaves to go and take a nap [remove Sally from stage]. Now that Sally is taking a nap, can she see what we’re doing in here? Can she hear what is happening here? [Confirm with child that she can neither see nor hear what is happening in the room; do not continue until they verbally confirm “no” to each question in <strong>e])</strong>.&lt;br&gt;&lt;br&gt;<strong>f)</strong> While Sally is sleeping, look what happens! Anne takes the ball out of the basket and puts it into the box [do this] and then Anne leaves [remove Anne from stage].&lt;br&gt;&lt;br&gt;<strong>g)</strong> Remember Sally? Sally woke up from her nap and she’s come back to play! Sally wants to play with the ball [bring Sally back and have her “scratch her head as if thinking”].</td>
<td>1) Where did Sally put the ball?&lt;br&gt;&lt;br&gt;2) Where is the ball now?&lt;br&gt;&lt;br&gt;3) Does Sally know it’s there?&lt;br&gt;&lt;br&gt;4) Where will Sally look for the ball?&lt;br&gt;&lt;br&gt;5) Why will she look in there?&lt;br&gt;&lt;br&gt;6) Where does Sally think the ball is?</td>
</tr>
</tbody>
</table>

Importantly, prior to presenting any puppet show(s), we ensured that the following were randomized: 1) the props being used; 2) the side of the “stage” that each character is presented on; 3) the location in which the first character places the object; and 4) the order of the puppet
shows presented. We also asked a series of other test questions but for the purposes of this study, we were only interested in the canonical, critical test question that is bolded in Table 3.

For a clearer depiction of the main test conditions (S-FB and LU-FB) presented to each child—please see Table 4 (below).

**Table 4. Series of Events in Each Condition**

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard False-Belief Task</td>
<td>1) Sally throws the ball up and down 3 times and then she puts the ball into the basket.</td>
</tr>
<tr>
<td></td>
<td>2) In Sally’s absence, Anne moves the ball from the basket to the box and leaves.</td>
</tr>
<tr>
<td></td>
<td>3) Sally returns.</td>
</tr>
<tr>
<td>Low-Uncertainty False-Belief Task</td>
<td>1) Sally puts the ball into the basket (and takes it out) 3 times.</td>
</tr>
<tr>
<td></td>
<td>2) In Sally’s absence, Anne moves the ball from the basket to the box and leaves.</td>
</tr>
<tr>
<td></td>
<td>3) Sally returns.</td>
</tr>
</tbody>
</table>

As evident in Table 4, Standard and low-uncertainty conditions differ in the type of repeated behaviour presented. In the LU-FB condition, Sally repeatedly hides the ball into the basket, thereby strengthening the child’s representation of Sally’s belief that it is in the basket. However, in order to standardize the amount of time between standard and low-uncertainty conditions, we had Sally repeatedly throw the ball up and down at the beginning of the standard task. Importantly, the repetition in the S-FB task does nothing to influence Sally’s belief about the ball’s location but that in the LU-FB task should strengthen the child’s certainty of Sally’s belief.
about the ball’s location (she really believes that it is in the basket because she put it in there so many times).

7  « Procedure »

We obtained informed consent from parents and verbal consent from children to participate. Children were situated in a testing room adjacent to an observation room in which their parents quietly sat and watched. Between both rooms was a two-way observational mirror such that from the testing room—it appeared to be a mirror. All puppets and props were placed in a closed suitcase underneath the table in order to ensure that children are not distracted by other toys and props during the stories presented. Research assistants were given the opportunity to bring a laptop onto which a PowerPoint slide of each randomized puppet show is presented. This was to ensure that participants are not eliminated due to researcher error. But to ensure that the added laptop in the room does not distract children, we asked research assistants to show the laptop to children at the outset of the study and to state the following: “this here just tells me what story we’re going to see next, okay?” after which children nodded. We found that simply explaining to children why there is a laptop behind them quenched their curiosity about it and thus, they were able to focus on the puppet shows presented.

Next, the research assistant would go through each of the four puppet shows in random order. After being asked the test question (i.e. “where will Sally look for the ball?”), children
were given the opportunity to point to, touch, or verbally state which location they believe is correct.

Meanwhile, in the observation room, parents watched their child interact with the research assistant. Once the child had completed all of the puppet shows, they were reunited with their parent(s) and given an opaque bag to choose a “mystery surprise toy” from. They were also given a $15.00 gift card to Indigo/Chapters/Coles and thanked for their time and participation. Each session took approximately 45 minutes to complete.

During visits to local preschools, all procedures were identical to that of in-laboratory sessions. To ensure the children’s safety, we asked that a teaching assistant employed by the preschool be present during each session (analogous to parents watching children from the observation room during in-lab sessions). Children from preschools were thanked with a Curious George book along with a $15.00 gift card to Indigo/Chapters/Coles for their participation.

We registered the procedure and analysis plan of Study 1 on the Open Science Framework (OSF; an online database that is part of a movement aimed at making scientific research more accessible and open). The OSF page for this study can be accessed via the following link: https://osf.io/4c7gm/. We have uploaded a word file on our OSF study page that addresses a series of typically asked questions regarding our study plan. This template was derived from the as predicted website (https://aspredicted.org/). All analyses discussed below are part of a-priori, hypothesis-driven analysis plans. Any other exploratory analyses that have not
been recorded on our OSF document (if conducted) are presented in the Appendices of this paper (rather than the main results sections).
Chapter 3: Study 1 Results

8 « Descriptive Statistics »

8.1 « Hypothesized Pattern of Data »

Our data consisted of proportions of children in each condition. For false-belief reasoning conditions, there are four cells into which all 40 children fell. The four possible cells are depicted in Table 5 (below).

Table 5. Four Cells that Constitute Child Proportions in False-Belief Tasks

<table>
<thead>
<tr>
<th>Passed S-FB Test</th>
<th>Failed S-FB Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed LU-FB Test</td>
<td>Proportion P</td>
</tr>
<tr>
<td>Passed Both</td>
<td>Passed Both</td>
</tr>
<tr>
<td>Passed Both</td>
<td>Failed Both</td>
</tr>
<tr>
<td>Failed Both</td>
<td>Hypothesis</td>
</tr>
<tr>
<td>Odd Pattern</td>
<td>Failed Both</td>
</tr>
</tbody>
</table>

As illustrated in Table 5, there are four possible proportions in the two false-belief tasks. For example, a proportion of children may pass both the standard and low-uncertainty false-belief tasks (those in Proportion P; here, the “P” represents the word passed). We expected some children to comprise this proportion, as those who passed the standard test should surely pass the considerably easier low-uncertainty test. The opposite pattern can be seen in Proportion F in which children fail both the standard and low-uncertainty tests. We did not expect many children to comprise this proportion because if they failed the standard test, we expected them to now
pass our low-uncertainty test. The third possible cell is depicted in Proportion O in which children display an odd pattern of behavior. Specifically, children in this proportion pass the standard (and arguably more difficult) test of false-belief reasoning but strangely fail the low-uncertainty (and theoretically easier) test of false-belief reasoning. This is a rather peculiar pattern of behavior and so, we did not predict that a substantial proportion of children would compromise Proportion O. Instead, we predicted that a significantly larger proportion of children would comprise Proportion H in which they fail the standard false-belief test but are able to pass the low-uncertainty test of false-belief reasoning. Because this experiment was predominantly aimed at understanding children’s false-belief understanding—we have focused our results on data from false-belief conditions. However, for an explanation of the hypothesized and actual pattern of data from true-belief conditions, please see Appendix B.

8.2 « Actual Pattern of Data »

Contrary to our hypotheses—the actual pattern of data do not appear to indicate that the manipulation helped children pass the low-uncertainty false-belief task. For descriptive statistics of the proportions of children in each condition, please see Table 6 (below).
Table 6. Actual Proportions of Children in Each Cell

<table>
<thead>
<tr>
<th>Passed S-FB Test</th>
<th>Failed S-FB Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed LU-FB Test</td>
<td>15</td>
</tr>
<tr>
<td>Failed LU-FB Test</td>
<td>3</td>
</tr>
</tbody>
</table>

Proportions of Children in Each Cell by Age

<table>
<thead>
<tr>
<th>Passed S-FB Test</th>
<th>Failed S-FB Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed LU-FB Test</td>
<td>15</td>
</tr>
<tr>
<td>Failed LU-FB Test</td>
<td>3</td>
</tr>
</tbody>
</table>

Proportions of Children in Each Cell by Sex

<table>
<thead>
<tr>
<th>Passed S-FB Test</th>
<th>Failed S-FB Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed LU-FB Test</td>
<td>15</td>
</tr>
<tr>
<td>Failed LU-FB Test</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. Younger and older cohorts of children were delineated by conducting a median split of age (in months). Older and younger children constituted the upper and lower halves of the data, respectively.
We conducted McNemar’s $\chi^2$ test using the `mcnemar.test` statistical package in R (version 3.3.2). Relative to children demonstrating the odd pattern of behavior (in which they pass the S-FB test but fail the LU-FB test), we did not find that a larger proportion of children who fail the S-FB test will pass the LU-FB test, McNemar’s $\chi^2(1) = 0.200$, $p = .655$, effect size (odds ratio) = 0.667. For a graphic illustration comparing the proportion of children passing the standard to the low-uncertainty false-belief test, please see Figure 1 (below).

![Bar plot illustrating the lack of difference between the proportion of children passing the S-FB test (43%), relative to those passing the LU-FB test (45%).](image)

Figure 1. Bar plot illustrating the lack of difference between the proportion of children passing the S-FB test (43%), relative to those passing the LU-FB test (45%).

For inferential statistics in true-belief conditions, please see Appendix B.
Chapter 4 : Discussion of Study 1 and Motivation Behind Study 2

10 « Study 1: Main Findings »

The results of Study 1 clearly do not support our prediction. As made evident by both the descriptive and inferential statistics of Study 1—the proportion of children who fall into Proportion H (in which we hypothesized many children to fall) was not significantly higher than those who fall into Proportion O (the odd pattern of behavior that we predicted very few children to show). Simply put, we can clearly see that our manipulation did not exert an effect on children, as those who pass the S-FB test typically also pass the LU-FB test and those who fail the S-FB test also tend to fail the LU-FB test. The goal of the sections that follow is to better understand the results of Study 1 by critically assessing the design, implementation, and theoretical perspective applied to this study. This discussion of our findings in Study 1 will set the foundation and motivation behind Study 2.

10.1 « A Deeper Analysis of Study 1: Potential Reasons for Null Effects »

One may reasonably argue that it is possible that—while it is true that our original sample size of children tested was large (N = 82)—our final sample size was small (N = 40). Due to time constraints and other practical considerations (i.e. successful completion of data collection for the purposes of this thesis), we did not meet our sample size goal. Within our OSF plan of study,
we conducted a power analysis using GPower 3.1 that resulted in a sample size goal of N = 89 ([https://osf.io/4c7gm/](https://osf.io/4c7gm/)). Thus it is possible that one reason why we did not observe the effects that we predicted is because we were underpowered to detect such an effect. This criticism is a valid one and one that we could not avoid, as within a limited time constraint (of approximately one year’s time), we ambitiously recruited and tested over 80 preschool-aged participants. However, we lost over 50% of our data after applying our exclusion criteria. We could not avoid using these criteria (such as the requirement that children respond correctly to both memory and reality questions)—as they are the critical and canonical exclusion criteria used in past studies. We needed to ensure that the data that we obtain is as accurate a reflection of children’s mindreading capacity as possible and in this effort, we needed to be sure that each child understood the series of events presented in the puppet show(s). The problem of small final sample sizes in preschool aged populations is not specific to our study, as many other studies with young children also suffer of this limitation. For example, the Baron-Cohen et al. (1985) study had a mere 27 typically developing children in their study. More recently, the study of false-belief reasoning with reduced processing demands (via response-generation practice trials) had 32 participants in experiment 1 and 16 participants in subsequent studies (Setoh et al., 2016). Thus the problem of small sample size is not unique to our study but is still, of course, a problem nonetheless. This was an inevitable limitation of our study and is problematic across studies with young populations.
However, in light of the fact that only three children (of the forty included in the final analysis) showed the pattern of behavior that we predicted—it is unlikely that increasing the sample size will result in a noticeable and significant shift in the findings. Because so few children demonstrated the pattern of behavior that we predicted—the likelihood that our results are solely due to our data being underpowered is potentially quite low.

A second possible explanation for the null effects observed in Study 1 is that children who fail tests of false-belief reasoning genuinely lack the ability to reason about others’ mental states and we are unable to help them pass. But in light of the overwhelmingly amount of empirical evidence on infant belief tracking (Onishi, Baillargeon, 2005; Moll et al., 2016; Southgate et al., 2007; Surian et al., 2007)—this explanation is unlikely. The notion that preschoolers in our study who failed the S-FB test simply lack a ToM altogether (and thus cannot pass the LU-FB test) brings us back to the seemingly puzzling and contradictory infant and preschool data presented in chapter one. This explanation does not account for any of the infant data presented in the past nor does it address studies that have reduced performance demands for younger children and thus, increased their success rate. It is therefore unlikely that the genuine conceptual deficit explanation is the true culprit underlying our findings.

A third possibility is that there are some inherent flaws in our repetition manipulation that limit its ability to exert any effect. Justifiably so, one may argue that perhaps our manipulation was not potent enough or even that its effects were simply different than those that we had originally intended. This possibility is actually quite likely, as all of my research assistants and I
noticed that following the repetition manipulation of LU-FB conditions—most children burst into uncontrollable laughter. This was particularly true for younger children who—instead of attending to and carefully thinking about Sally’s repeated hiding behavior—clearly thought the manipulation was silly. Our initial reaction to the first few children who displayed this behavior was that this might be specific to this child but as our sample size increased (and the tendency of children to laugh with the manipulation remained), we began carefully considering the reason(s) why they found it funny. As adults, we may not understand this as readily but in actuality, the manipulation is quite silly and funny. To repeatedly place an object into a location (by putting it in the basket, then taking it out of the basket; then putting it in the basket etc.) is strange, silly, and odd. Children recognized this and giggled uncontrollably when they saw this pattern of behavior. It is thus quite possible that (especially for the young preschooler) our manipulation did not exert the effect that we anticipated. When designing the study, we did not consider this outcome and (as is an issue across studies with young children), we were surprised to see children’s reactions when we implemented it. But in reality, they are right for laughing, as it is a silly, strange, and even funny course of action to take (to repeatedly place something into a container)! Thus in future investigations, we will aim to manipulate children’s certainty in ways that are hopefully more clear and that directly alter children’s certainty (rather than exerting some other, albeit emotionally positive, effect).

And lastly, the fourth and final possible explanation for our null results pertains to the theoretical motivation behind the study. Namely, that ToM (and thus false-belief reasoning)
genuinely does not operate through information updating; as such, an information-theoretic approach to studying ToM through gradual uncertainty reduction may not be justified. This explanation—while seemingly well-substantiated and valid—is unlikely because uncertainty reduction via information updating is quite possibly the mechanism by which ToM operates in adults. It is therefore highly conceivable that the system operates similarly in early development. At this juncture, one may ask the following question: “on what basis are we basing the inference that adult ToM operates through uncertainty reduction via information updating (i.e. with an information-theoretic perspective)?”—and reasonably so. This statement is not on the basis of intuition or personal experience but on the foundation of the findings of Study 2.

11  « Motivation behind Study 2 »

In an effort to better understand and interpret the null findings of Study 1—we reflected back on the design, implementation, and theoretical framework of Study 1. This critical and reflective process inspired us to reflect on why we feel that the mechanism by which we make mental state inferences of others’ minds is through uncertainty-reduction algorithms. In theory, this might sound intuitive, obvious, and indisputable but indeed, this gradual belief updating on the basis of the informativeness of the stimulus may not be the mechanism by which we make belief attributions. We therefore applied an information-theoretic perspective to understanding how adults make inferences about other people’s mental states. We also addressed the potential problem(s) of boredom and humor associated with the repetition manipulation by more directly
manipulating the amount and type of information that participants are given when reading short vignettes about various characters.

11.1 « Study 2: Design and Hypotheses »

We included standard and low-uncertainty conditions during which participants read vignettes containing a moderate, or rich source of information to base a mental state inference on. But because adults have a considerably longer attention span than young preschoolers—this wealth of time afforded us the ability to also increase participants’ uncertainty with ambiguous, informationally impoverished, high-uncertainty conditions.

The design of Study 2 was that of a one-way (condition: standard, low-uncertainty, high-uncertainty) within-subjects design. All participants were exposed to 5 versions of the three possible story types making a total of 15 stories that each participant was exposed to.

We made the following three hypotheses regarding our findings from Study 2: *Using a within-subjects design, we will see an incremental pattern such that participants’ performance will be as follows: 1) Moderately high in standard conditions; 2) Significantly higher in low-uncertainty conditions; and 3) Lowest in high-uncertainty conditions.*
Chapter 5 : Study 2 Methods

12 « Participants »

Sixty (29 females) healthy adult participants between 18 to 31 years of age ($M = 20.600; SD = 3.288$) were tested in Study 2. All participants were undergraduate students at Western University and had normal or corrected-to-normal vision. Three participants were excluded from the final analyses because they did not provide a complete dataset. The final sample consisted of fifty-seven (26 females) healthy adult participants between 18 to 31 years of age ($M = 20.632; SD = 3.368$).

13 « Materials »

The materials used in Study 2 consisted of a series of short vignettes (stories about fictional characters) presented on a computer screen. Specifically, a total of 15 short vignettes were given to each participant to read. After carefully reading each vignette, participants were shown a series of questions that assessed their ability to reason about the mental state(s) of the main character in the vignette. All vignettes were presented (and subsequent responses to questions pertaining to each vignette were collected) using OpenSesame Experiment Builder 2.9.7.
13.1 « Randomization Technique for Vignettes »

With respect to content within the vignettes, we wanted to ensure that personal biases do not influence the topics of each story and so, all aspects of each story were generated randomly using the following four steps:

1) Each story consisted of a character whose gender was decided using a random number generator and whose name was decided using the following random baby name generator online: http://www.randomnames.com/.

2) The element onto which the character would have a set of mental states about could be a person, place, thing, or phenomenon and this was decided using a random number generator.

3) If it was a person, then the name and gender were decided in the same way that those of the first character were generated. If it was a place, thing, or phenomenon, then the following random generators online were used: https://www.randomlists.com/random-world-cities, https://www.randomlists.com/things, and http://www.ratespeeches.com/t=Speech-Topics, respectively.

4) Once the element onto which the agent has a mental state was decided, using a random number generator, whether or not the character agrees, likes, wants (i.e. approaches) or disagrees, dislikes, hates (i.e. avoids) the element was generated randomly.
For example, a male (step 1) named Lewis (step 2) has a set of mental states (beliefs, thoughts, knowledge, emotions, and behavioral intentions) about a second character named Clara’s piano playing (step 3); in particular, Lewis likes (approach; step 4) Clara’s piano playing.

13.2 « Test Questions »

After participants carefully read each vignette, they were shown a series of test questions assessing their ability to reason about the character in the vignette’s mental state(s). In order to more comprehensively capture the constituent aspects of ToM, participants were asked questions that assessed their ability to reason about the character’s beliefs, thoughts, knowledge, emotions, and behavioral intentions. For belief questions, the word “believes” was included in a true or false formatted question such as the following: “Lewis believes that Clara is musically talented”. For questions about others’ thoughts and knowledge, the words “thinks” and “knows” were included in a similar format. For questions evaluating participants’ ability to reason about the emotions of others, the six basic emotions of happiness, sadness, anger, fear, surprise, and disgust (Ekman, 1992) were used in multiple choice style questions such as the following: “when Lewis listens to Clara play the piano, how does he feel?” along with the six basic emotions as response options ranging from a) to f).
13.3 « The Standard Task »

The standard task was intended to be moderately difficult in that it provides participants with a reasonable (but not substantial) amount of information onto which participants can base a mental-state inference. Importantly, the information that participants were given included *behaviours* and not direct statements about the given character’s mental states. This way, we could ensure that participants are not simply memorizing scripts from the text provided but are actually making a mental state attribution based on the behaviour and disposition of the character(s) presented. Standard conditions began with a brief introduction followed by a total of six sentences—not two of which presented information that was *consistent* with the agent’s mental state(s), while the remaining *four sentences* served as *length-fillers* that ensure consistency across conditions in the amount of text displayed in each vignette. It is noteworthy that the order of each sentence type (be that consistent or filler) was randomized. For an example of a story within the standard condition, please see Appendix C.

13.4 « The Low-Uncertainty Task »

The low-uncertainty task was intended to be exceptionally easy in that it provides participants with a substantial amount of information onto which participants can base a mental-state inference. These conditions were identical to standard conditions with the exception that rather than *two* consistent-information sentences—participants were given *six* sentences that are
consistent with the agent’s mental state(s). Like standard conditions, low-uncertainty conditions began with a brief introduction, but a total of six sentences followed—all of which presented information that was consistent with the agent’s mental state(s). Here, as new information comes in, there should be reduced ambiguity, confusion, and uncertainty regarding the agent’s mental state(s). For an example of a story within the low-uncertainty condition, please see Appendix D.

13.5 « The High-Uncertainty Task »

The high-uncertainty task was intended to be exceptionally difficult in that it provides participants with a cryptic set of information designed to increase participants’ confusion and essentially, their level of uncertainty. In addition to the typical introductory sentence, participants were given conflicting information such that three sentences provided participants with consistent information—consistent with the agent’s set of mental state(s)—and three sentences provided participants with inconsistent information (directly contradicting consistent sentences). In these vignette types, as new information comes in, there should be increased ambiguity, confusion, and uncertainty regarding the agent’s mental state(s). For an example of a story within the high-uncertainty condition, please see Appendix E.
14  « Design Features of Study 2 »

14.1  « Technique for Topic Matching »

Despite the fact that vignette topics were generated using the randomization technique outlined in 12.1—it is still possible that story topics for low-uncertainty vignettes (by chance) happen to be simpler than those in, say, high-uncertainty vignettes. We therefore decided to take an additional precaution to ensure that story topics are matched across conditions. Thus for each story topic (i.e. the story about Lewis), we generated a standard, low-uncertainty, and high-uncertainty version. Participants each saw a total of fifteen vignettes, five of which were standard story versions, five of which were low-uncertainty story versions, and five of which were high-uncertainty story versions.

14.2  « Accuracy Measure »

To obtain an accuracy measure for each condition from participants, we obtained their responses to each of the five (believe, think, know, emotion, and intention) test questions presented after each vignette. For the standard condition, for example, each participant read five standard vignettes after which five questions were asked. Therefore, for this condition, participants each provided responses to a total of twenty-five questions—resulting in a calculated accuracy score out of twenty-five for the standard condition. The same logic was applied to the low and high-uncertainty conditions. It is noteworthy that we also generated reaction time
measures but our main hypotheses surrounded the extent to which the amount and type of information that participants are given influences their final mental-state attribution(s)—their performance accuracy.

15 « Procedure »

Prior to collecting any data for Study 2, we preregistered our study on OSF. The preregistration page as well as the as predicted document for this study can be found via the following link: https://osf.io/89wr4/. All analyses discussed in the results section are part of a-priori, hypothesis-driven analysis plans. Any other exploratory analyses that have not been recorded on our OSF document (if conducted) are presented in the Appendices of this paper.

After providing written consent to participate in the study, participants were seated in front of a computer screen and told that in a few moments, they will see a series of short stories (“paragraphs of text”) appear on the screen. They were instructed to read each story as carefully as possible and were reassured that they will have ample time to read all of the stories presented. They were then told that once they are comfortable with their knowledge of each story, they may press any key on the keyboard and this will generate a series of questions pertaining to the story they just read. They were specifically told to focus their attention on “getting the answers right” rather than quickly reading and answering questions. Once they confirmed that they understand the instructions, the research assistant started the computer task and participants began reading. To answer True or False type test questions, participants pressed the “t” or “f” keys to indicate
that the statement is true or false, respectively. To answer *Multiple Choice* style questions, participants simply pressed the key corresponding to their choice (i.e. “a” to choose option a). Once finished, participants were given one full research course credit for their participation and were thanked for their time. The entire session took approximately 45 minutes to complete.
Chapter 6 : Study 2 Results

16 « Descriptive Statistics »

Descriptive statistics of Study 2 indicate trends in support of our predictions. That is, performance (accuracy) was moderately high in the standard condition ($M = 75.000; SD = 10.590$), highest in the low-uncertainty condition ($M = 81.965; SD = 8.606$), and lowest in the high-uncertainty condition ($M = 52.754; SD = 11.891$).

17 « Inferential Statistics »

A repeated-measures Analysis of Variance (ANOVA) conducted using the Statistical Package for the Social Sciences Software (SPSS) revealed that there is a statistically significant effect of condition on accuracy, $F(1,56) = 115.064, p < .0001, \eta^2 = .673$. Follow-up pairwise comparisons revealed that at all three levels, there is a significant difference. Results of the pairwise comparisons are depicted in Table 7 (below).
Table 7. Results of the Pairwise Comparisons between the Three Conditions

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Mean Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard vs. Low-Uncertainty</td>
<td>-6.965*</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Standard vs. High-Uncertainty</td>
<td>22.246*</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Low-Uncertainty vs. High-Uncertainty</td>
<td>29.211*</td>
<td>&lt;.0001***</td>
</tr>
</tbody>
</table>

For a graphical illustration of the results of Study 2, please see Figure 2 (below).
Because our main findings concerned participants’ performance accuracy across condition types—we presented accuracy data for Study 2 in this chapter. For reaction time data, please see Appendix F.
Chapter 7 : Study 2 Discussion

18  « Study 2: Limitations »

Unlike study 1, we not only met but exceeded our sample size goal with study 2. Our OSF preregistration document for Study 2 includes a power analysis using GPower 3.1 that resulted in a sample size goal of N = 45 (https://osf.io/89wr4/). Our actual sample size for Study 2 was N = 57. Thus although sample size is clearly not a limitation of study 2—the generalizability of the sample under study is. All participants included in the final analyses for study 2 were recruited from Western University and consisted of students taking undergraduate psychology courses or of graduate students in the department of psychology. This sample clearly does not reflect the general population, as all participants are from a relatively young, educated sample of North American participants. Although generalizability is often a limitation with studies in psychology and related fields—the mere fact that it is a common practice to recruit university students as participants does not justify it. We recruited an unrepresentative sample of the general population for several reasons: convenience, efficiency, compensation cost, and necessity. This sample was a convenient sample that could be efficiently tested with no cost (they were given course credit) in a considerably short amount of time.

Another critique of study 2 worth addressing is the external validity of the tasks used. In real-world interactions, we are not faced with vignettes in the form of a paragraph of text on a computer screen from which to make a mental-state inference. Instead, we are given a multitude
of cues from which to base a ToM attribution. Missing from our stimuli are facial cues, bodily gestures, and changes in an agent’s vocal pitch—all of which may contribute to participants’ certainty regarding the agent’s belief status. The reasons why we implemented tasks with such low external validity pertain to the ease of administration and convenience of using these tasks. We also feel that our tasks are internally valid in that there is a considerable amount of control in the details pertaining to each story that are used. The randomization technique and matched story content across conditions were some ways in which we increased the internal validity and reduced the potential for noise influencing the results of study 2. Despite the virtues of the internal validity and control of the vignettes used in study 2, follow-up research with greater external validity is warranted.

A third critique of study 2 pertains to the test questions asked after the presentation of each vignette. Of the five question types asked (emotion, believe, think, know, and behaviour prediction), the first question type (emotion reasoning) suffers from a great limitation. In an effort to reduce the possibility of linguistic ambiguity in the emotions presented—we presented participants with only the six basic emotions (Ekman, 1992). However, for some scenarios, it may have very well been possible that the likely way in which the character in the vignette was feeling was not happy, sad, angry, disgusted, surprised, or afraid. This renders it possible that participants simply chose the best possible option and not the way in which they truly believed the character to feel. To address this, open-ended questions should be used in future studies in which participants enter an emotional feeling (i.e. excited) that they attribute to the character.
Afterwards, participant responses can be qualitatively analyzed by independent coders. But to increase our efficiency, we included close-ended questions in multiple-choice format.

To address all of the concerns discussed in this section, more representative samples of the population should be tested using more externally valid tasks. Examples of such tasks include video vignettes that portray facial expressions, body language, and vocal intonation as a basis for participants to generate mental state inferences on. And finally, dependent variables should be open-ended, allowing participants to think beyond the limited set of options given in close-ended questionnaires.

19  « Study 2: Main Findings and Interpretation »

Despite these limitations, the main findings of Study 2 are interesting and insightful, nonetheless. Study 2 manipulated participants’ levels of uncertainty when making mental state attributions by altering the informativeness of a series of vignettes. If the underlying mechanism by which ToM operates in adults is that of gradual uncertainty reduction with the onset of new information—then we should observe a difference in the final ToM inferences that adults make as a function of the condition of each vignette. Results of Study 2 confirmed our hypotheses by revealing that performance is moderately high when participants are provided with a small (albeit consistent) amount of information to base mental state attributions on. Performance then increases considerably as the amount of consistent information is increased—thereby enriching the information source by which participants form a ToM. We also show in Study 2 that we can
hinder performance by altering the final mental state attribution that participants make with an informationally ambiguous, uncertainty-provoking condition. Taken together, this set of findings substantiate the view that ToM operates via a gradual uncertainty-reduction, belief-updating mechanism in which final inferences are updated as new, consistent information is obtained. In light of these data with adults, it is possible that a similar (albeit potentially less sophisticated) mechanism operates early in development.
Chapter 8: General Discussion

20 « Study 1 and 2: A Brief Comparison »

Due to the design and implementation limitations outlined in chapter 4—Study 1 could not adequately assess the utility of information-theory in understanding preschoolers’ mental state reasoning capacities. The way in which Study 1 enriched preschoolers’ belief certainty of another agent was via repetition in a standard Sally-Anne test of false-belief reasoning. The manipulation of repetition was based on the logic that with each iterative display during which Sally hid the ball in the basket, we can systematically reduce the child’s uncertainty regarding Sally’s belief about the ball’s location. This design—while in theory and seemingly in practice well-justified—produced an interesting, but unintended, effect. In practice, children found the manipulation silly and laughed (rather than carefully attending to the scenario). It is also possible that the repetitive nature of the manipulation prevented children from obtaining a rich body of information from multiple sources (as was done in the adult study). Recall that in the adult study, participants in the low-uncertainty condition did not simply repeatedly read the same sentence—but were given multiple pieces of convergent information onto which a cohesive set of mental state attributions can be generated. Thus relative to the adult study (Study 2), the preschool study (Study 1) did not enrich participants’ information source in an effective, potent, or comprehensive manner.
Instead, a similar (but child-appropriate) set of vignettes that systematically enriches children’s information source is warranted. Returning back to the literature on ToM development, the field is increasingly adopting the *Substantial Continuity View* instead of the previously held *Fundamental Conceptual Change View*. This shift was largely a consequence of the plethora of research findings indicating that children far below 4-5 years of age (in many cases, young infants) have the capacity to track mental states (Scott & Baillargeon, 2017). The *competence-performance distinction* (in which young children who fail explicit tests of false-belief reasoning simply cannot overcome performance demands in the task—but nonetheless have ToM competence) helps resolve the seemingly conflicting body of infant and preschool data (Setoh et al., 2016; Surian & Leslie, 1999; Yazdi et al., 2006). Missing in the literature, however, is the perspective that the impoverished nature of information provided to young preschoolers serves as a performance demand, preventing their ability to make a verbally explicit belief-attribution. This perspective is motivated by information-theory (Balsam & Gallistel, 2009) and the field is wanting of its application in the domain of ToM.

21.1 « Study 3: Design »

In Study 3, we hope to address the necessity for a study that systematically reduces children’s uncertainty regarding the mental states of others. In Study 3, we aim to more
effectively manipulate the richness of children’s information source with child-appropriate vignettes that differ in the amount and consistency of information that they provide children with. In addition to a set of standard cartoon vignettes (taken from the *Theory of Mind Battery*; Hutchins, Prelock & Chace, 2008), we will include a set of low-uncertainty and high-uncertainty vignettes. The purpose of the low-uncertainty vignettes is to assess the extent to which young children (who fail standard tests of mental state reasoning) can pass with an informationally enriched set of similar vignettes. And finally, using the same logic applied when creating the high-uncertainty vignettes of Study 2, we will determine whether or not older children (who pass standard tests of mental state reasoning) fail with an ambiguous, uncertainty-provoking set of information.

Specifically, vignettes in the standard condition will provide children with a moderate amount of consistent information—after which the remaining sentences will serve as length fillers. For example, the character Patty from one of the existing stories in the ToM-Battery (Hutchins et al., 2008) thinks that her glasses are on the table. In the standard condition, children will be given one piece of information consistent with this mental state; namely, that “this morning, Patty saw her glasses on the table”. This is the standard amount of information—identical to that provided in the existing vignette within the ToM-Battery. To increase length such that vignettes in the standard condition are equivalent in length to those in other conditions, we will include filler sentences (as was done in the standard condition of Study 2). Vignettes in the low-uncertainty condition, on the other hand, will provide children with a substantial amount
of consistent information—enriching the child’s certainty of the character in the story’s set of mental states. For example, in addition to the information provided in the standard condition, children will be given additional information such as the following: “Yesterday, Patty’s babysitter asked her why she left her glasses on the table”. These additional sentences provide the child with more consistent information—consistent with the agent’s mental states (i.e. that Patty thinks her glasses are on the table). The third and final types of vignettes (those in the high-uncertainty condition) will provide children with conflicting information such that some sentences are consistent with the character’s mental states (i.e. that Patty thinks her glasses are on the table) and others that are inconsistent with the character’s mental states (i.e. that Patty does not think that her glasses are on the table). Examples of such sentences are as follows: “this morning Patty saw her glasses on the table” and “but by supper time, Patty saw that her glasses were not on the table”. For example stimuli from the standard, low-uncertainty, and high-uncertainty conditions, please see Appendices G, H, and I, respectively.

### 21.2 « Study 3: Predictions and Significance »

We plan on implementing the same design characteristics as those applied in Study 2. That is, we will use a within-subjects design and create story topics using the same randomization and matching procedure as outlined in chapter 5. Like the hypotheses outlined for study 2 (in chapter 4), we predict the following:
Like the findings of Study 2, Study 3 will discover an incremental pattern such that children’s performance will be as follows: 1) Moderately high in standard conditions; 2) Significantly higher in low-uncertainty conditions; and 3) Lowest in the high-uncertainty conditions.

This will be the first well-controlled study of ToM development to apply an information-theoretic perspective in delineating the performance demands of information richness when children complete tests of mental state reasoning.

22 « Concluding Remarks »

22.1 « A More In-Depth Discussion of Each Study »

The purpose of Study 1 was to apply an information-theoretic approach to understanding ToM. Findings from Study 1 suggest that in children, the repetition manipulation as a means of reducing uncertainty has no effect. Instead, Study 1 elucidated the necessity for robust and effective manipulations of uncertainty reduction. The purpose of Study 2 was to discern how ToM operates in adults. The results of Study 2 confirmed that manipulating informational richness more directly using vignettes does influence adults’ final mental state attributions. More specifically, Study 2 demonstrated that a) the mechanism by which adults form mental state attributions is via a gradual and integrative updating process of ToM-relevant and consistent information; b) final mental state attributions can be manipulated via uncertainty-reduction and uncertainty-amplification; and c) vignettes in which various types of information converge to
generate a set of mental states serve as an effective and robust manipulation of uncertainty in tests of ToM. And finally, the purpose of Study 3 is to discern how we can better manipulate informational richness more directly with children. The goal of Study 3 is to create child-appropriate vignettes in which we apply the logic of the design and implementation of Study 2 to address the outstanding research questions that motivated the theoretical foundation of Study 1. Once complete, all three studies will provide important scientific insights to both the research and general community of the mechanisms by which—or *how*—ToM develops.

### 22.2 « Revisiting the Competence-Performance Distinction and Information Theory »

As demonstrated by the literature reviewed in Chapter 1, a considerable amount of debate and controversy surrounds the issue of ToM development. Early research studies on ToM development found strong support for the notion that there is a discontinuous shift during early preschool in which children suddenly acquire the ability to infer mental states. This fundamental conceptual change view—while substantiated by evidence that young preschoolers fail explicit tests of ToM—has come under criticism by a growing number of contradictory research findings. That is, recent findings from implicit ToM tests are increasingly suggesting that some belief-tracking abilities exist in young infants. These findings bring into question the fundamental conceptual change view and warrant the following developmental research question: If infants can track mental states, why do young preschoolers fail tests of mental state reasoning?
One explanation for these seemingly opposing views is that perhaps there is a substantial amount of continuity in ToM abilities throughout the lifespan. These abilities are captured well with implicit ToM tests administered to infants but explicit tests given to preschoolers pose additional task demands that prevent children from passing. These additional, non-focal task demands may include demands on working memory, verbal comprehension, verbal fluency, and inhibitory control. Importantly, these demands may hinder children from demonstrating their true mind-reading capacity and thus, their performance is compromised. One task demand that has yet to be investigated pertains to the impoverished nature of the information provided in explicit ToM tests given to young children. In accordance with information theory, when information provided by the stimulus is low, uncertainty will concomitantly increase. From this standpoint, young children who fail explicit ToM tests may simply require a richer source of information in order to reduce their level of belief-uncertainty. By using this information-theoretic perspective, we may better understand the seemingly puzzling findings of research on ToM development. This was the motivation behind Study 1 and although results did not confirm that manipulating uncertainty via repetition has an effect on children—there are several reasons why we should take caution before concluding that young children cannot pass ToM tests because they simply lack a ToM.

First, to conclude that young children cannot be helped and simply cannot pass because of a conceptual deficit in ToM ignores the overwhelmingly large and cogent body of infant data on belief-tracking. Second, this conclusion would also disregard the experiments conducted with
preschoolers outlined in Chapter 1 during which various task demands were reduced and resultantly, many preschoolers were then able to successfully pass. And third, concluding that the findings of Study 1 support the fundamental conceptual change view would neglect the many practical considerations and critiques of the design and implementation of Study 1. Critically, many young children found the repetition manipulation of Study 1 to be humorous and laughed, rather than carefully attending to the scenario presented to them. Upon further reflection, it became evident to us that young children’s reasoning for finding the manipulation to be silly is well-warranted (as it, indeed, is silly to repeatedly place a ball into a basket). This brings into question the validity of the repetition manipulation in Study 1 as a means for reducing uncertainty. It is for these reasons that the results of Study 1 do not negate the utility of information theory in understanding ToM but simply warrant more direct manipulations of uncertainty.

Study 2 manipulated informational richness more directly by administering vignettes that alter the amount and type of information provided in the stimulus. The goal of Study 2 was to assess whether or not the mechanism by which ToM operates is, in actuality, via uncertainty-reduction. Applying an information-theoretic perspective to understanding ToM may not have been appropriate, as perhaps ToM genuinely does not operate through such a mechanism. Thus, testing information-theory more directly in adults was well-warranted and tested in Study 2. In adults, the type and amount of information provided in the stimulus had a profound role in altering the final mental state attributions made. It is logical to assume that at the outset of Study
2, participants had no mental state attributions about the characters in each vignette. Thus they had no prior ToM hypotheses about the agents in each vignette. In a similar manner as the operation of Bayesian models of cognitive reasoning (Gopnik et al., 2004; Perfors et al., 2011), we found that incoming information altered the final mental state attributions made. It is critical to note that there are three stages underlying this process of changing mental state attributions: 1) informational-richness provided in the stimulus; 2) uncertainty-reduction/amplification; 3) final mental state attributions made. In Study 2, we altered 1) and found concomitant changes in 3) but importantly, we did not test whether or not this was because of 2). However, it is a reasonable and logically valid assumption that altering informational richness will alter the certainty index which will, in turn, result in a given set of mental state attributions. In future investigations, it would perhaps be fruitful to test this by simply asking participants to rate their certainty level when making mental state inferences.

22.3 « Final Thoughts: Information Theory across Development »

The findings of Study 2 confirm that adequate processing of the explicit system requires, at least in part, an enriched information source which should, in theory, significantly reduce the level of uncertainty. Whether or not the ToM system operates in a similar fashion in early development (i.e. during early preschool) warrants testing in Study 3. However, the data provided by Study 2 do provide useful insights as to the mechanism by which the explicit system
operates. These data can potentially help in understanding the contradictory preschool and infant data on the development of ToM. Recall that in infancy, the implicit system is at play, whereas in preschoolers—the explicit system is tested. One view may be that evidence of belief-tracking in infants is simply reflective of the system engaged—the implicit ToM system. This system may seamlessly operate even in the face of an impoverished set of information. While it is possible that the implicit system can function in a rudimentary, automatic sense in the absence of an enriched information source—this possibility has yet to be tested. However, even if the implicit system did require a rich information source—this is often provided in infant studies in which familiarization trials provide additional consistent information via repetition.

The explicit system, on the other hand, requires a plethora of cognitive abilities (as addressed by performance accounts) in addition to rich information source (as outlined by Study 2). Thus children who fail explicit tests of ToM may not have a fully-developed explicit mental state reasoning system and therefore require a highly enriched information source. Children who pass, on the other hand, may have a more developed explicit system in which they are able to extract relevant pieces of information from a social scenario and generate a cohesive set of mental state inferences—albeit from an informationally impoverished source. These children (who typically tend to be older), mimic the explicit mind-reading system of adults. By adulthood, the explicit system is enhanced and is able to extract ToM-relevant pieces of information from various sources of information more readily. In doing this, the system is able to generate a set of accurate and cohesive mental-state inferences on the basis of the informational richness of the stimulus.
References


Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of


https://doi.org/10.1017/S0305000913000433.


https://doi.org/10.1007/s13164-015-0290-z.


https://doi.org/10.1371/journal.pone.0142405.


https://doi.org/10.1111/1467-8624.00304.

https://doi.org/10.1016/j.cognition.2005.05.004.
Appendices

Appendix A: A Deeper Explanation of True-Belief Conditions

Consistent with previous studies of false-belief reasoning, we also included true-belief conditions as control trials during which the agent does not have a false-belief. Here, Sally sees Anne move the ball and thus, she has a true-belief about its location. True-belief conditions are considerably easier than false-belief conditions because there are now no processing demands requiring the child to overcome the temptation to locate the ball’s actual location (as in false-belief conditions). That is, the true location of the ball is also the location in which Sally believes it to be located. Of course, when children pass true-belief conditions it is unclear whether or not it is because they are truly reasoning about Sally’s beliefs or because they are simply locating the ball based on their own beliefs.

For matched true-belief conditions, we included the standard true-belief (S-TB) task in addition to a low-uncertainty true-belief (LU-TB) task. Consistent with previous studies, during the S-TB task, Sally placed the ball into the basket after which she left. She then returned and saw Anne move the ball from the basket to the box. During the LU-TB task, after Sally’s return, she saw Anne move the ball from the basket to the box repeatedly. Here, the LU-TB condition is matched in its series of events to the LU-FB condition with the exception that in true-belief conditions, Sally possesses a true belief. Based on the aforementioned literature reviewed in
Chapter 1—most preschoolers should pass true-belief tasks. The series of events in the true-belief conditions of the present study were as follows:

<table>
<thead>
<tr>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard True-Belief Task</td>
</tr>
</tbody>
</table>
| 1) Sally throws the ball up and down 3 times and then she puts the ball into the basket.  
2) In Sally’s presence, Anne moves the ball from the basket to the box and leaves.  
3) Sally remains. |
| Low-Uncertainty True-Belief Task |
| 1) Sally puts the ball into the basket.  
2) In Sally’s presence, Anne moves the ball from the basket to the box and takes it in and out of the box 3 times and leaves.  
3) Sally remains. |

It is noteworthy that the repetition in the standard condition is not meant to alter Sally’s belief status but simply to act as an equivalent control to the low-uncertainty task. The low-uncertainty task, however, does include repetition that should strengthen Sally’s belief (she saw it go into the box 3 times and therefore should believe it to be there with more certainty). Children were expected to do well in both standard and low-uncertainty true-belief tasks.
Appendix B: Hypothesized and Actual Data from True-Belief Conditions

**Hypothesized Pattern of Data**

For true-belief reasoning conditions, our prediction was simple—the majority of children will pass *all* tests of true-belief reasoning. But because the manipulation in the LU-TB condition strengthened children’s certainty regarding Sally’s true-belief of the ball’s location—more children should pass the LU-TB condition (relative to the S-TB condition). Importantly, however, we did not predict that these results will be as robust as those in false-belief conditions. Furthermore, if we found that more children pass the LU-TB condition (as compared to those passing the S-TB condition), we would not be sure whether or not they passed because they were more certain about Sally’s belief—or because they were more certain about where the ball *really* is because the repetition here can influence either or both factors.

**Actual Pattern of Data: Descriptive Statistics**

Of the 32 children tested, 19 passed both the S-TB and the LU-TB test, while only 5 children failed both the S-TB and LU-TB test. We expected a large proportion of children to pass true-belief conditions and so, this finding was consistent with our predictions. Two children passed the S-TB test and failed the LU-TB test and the remaining six children failed the S-TB test but passed the LU-TB test. Critically, it is unclear whether or not these six children passed the LU-TB test because they were more certain of Sally’s true belief of the ball’s location, or
because they, themselves were more certain of the ball’s location and simply used their own knowledge as a basis for their answer.

**Actual Pattern of Data: Inferential Statistics**

When we ran McNemar’s test, we did *not* find that these six children comprise a larger proportion than those showing the reverse pattern of behavior, McNemar’s $\chi^2(1) = 2.00$, $p = .157$, effect size (odds ratio) = 0.333.
Appendix C: A Sample of Stimuli used in the Standard Condition of Study 2

**Introduction:** Meet Lewis. Lewis has a roommate named Clara. Clara plays the piano.

**Filler 1.** Lewis speaks four languages because growing up, his father was a linguist who encouraged Lewis to learn multiple languages.

**Filler 2.** Lewis and Clara’s apartment is in downtown Chicago, which can get very expensive. Luckily, their landlord covers the cost of utilities.

**Consistent 1.** Lewis spends most of his evenings listening to Clara play the piano. After several hours of listening to her play, Lewis often goes to the kitchen.

**Filler 3.** He then makes himself a cup of camomile tea to help him fall asleep.

**Consistent 2.** One day, Lewis came across an advertisement to join the local orchestra. Lewis took a copy of the advertisement home and gave it to Clara.

**Filler 4.** That night, Lewis’s mother called and when they spoke, Lewis asked his mother if they would be visiting Chicago over the holidays.

*Please note that bolded and underlined sections of text (above) were not included in the text presented to participants but are depicted here for illustrative purposes in demonstrating the utility of each sentence.*
Appendix D: A Sample of Stimuli used in the Low-Uncertainty Condition of Study 2

**Introduction:** Meet Lewis. Lewis has a roommate named Clara. Clara plays the piano.

**Consistent 1.** Lewis spends most of his evenings listening to Clara play the piano.

**Consistent 2.** After several hours of listening to her play, Lewis often smiles and applauds.

**Consistent 3.** When Clara’s piano needed tuning, Lewis paid to have it done professionally.

**Consistent 4.** Lewis regularly organizes social get-togethers in their apartment and asks Clara to play the piano for their friends.

**Consistent 5.** He even once videotaped Clara playing the piano and uploaded the video on Youtube for the world to see.

**Consistent 6.** One day, Lewis came across an advertisement to join the local orchestra. Lewis took a copy of the advertisement home and gave it to Clara.

*Please note that bolded and underlined sections of text (above) were not included in the text presented to participants but are depicted here for illustrative purposes in demonstrating the utility of each sentence.*
Appendix E: A Sample of Stimuli used in the High-Uncertainty Condition of Study 2

**Introduction:** Meet Lewis. Lewis has a roommate named Clara. Clara plays the piano.

**Consistent 1.** Lewis spends most of his evenings listening to Clara play the piano.

**Inconsistent 1.** After several hours of listening to her play, Lewis often gets a headache.

**Inconsistent 2.** He once went to the mall across the street and bought a pair of noise-cancelling headphones.

**Consistent 2.** When Clara’s piano needed tuning, Lewis paid to have it done professionally.

**Inconsistent 3.** During a social get-together at their apartment, Clara played the piano and Lewis wore earplugs during her performance.

**Consistent 3.** One day, Lewis came across an advertisement to join the local orchestra. Lewis took a copy of the advertisement home and gave it to Clara.

*Please note that bolded and underlined sections of text (above) were not included in the text presented to participants but are depicted here for illustrative purposes in demonstrating the utility of each sentence.*
Appendix F: Reaction Time Data for Study 2

**Descriptive Statistics**

Descriptive statistics of reaction time data for Study 2 are as follows: participants’ reaction time (in seconds) indicate that they are moderately fast when answering questions for the standard condition ($M = 5.208; SD = 1.733$), fastest in the low-uncertainty condition ($M = 5.063; SD = 1.823$), and slowest in the high-uncertainty condition ($M = 6.229; SD = 2.435$).

**Inferential Statistics**

A repeated-measures (ANOVA) in SPSS revealed that there is a statistically significant effect of condition on reaction time, $F(2,56) = 25.064, p < .0001, \eta^2 = .309$. Follow-up pairwise comparisons revealed that this is mainly driven by the longer response times in the high-uncertainty condition, as it differed significantly from both the standard ($M_{\text{difference}} = 1.022; p < .0001$) and low-uncertainty conditions ($M_{\text{difference}} = 1.166; p < .0001$). However, the standard and low-uncertainty conditions did not differ significantly ($M_{\text{difference}} = 0.145; \text{ns}$).
Appendix G: Sample Stimuli for the Standard Condition of Study 3

This is Patty.

Patty has a dog named Teddy.

Teddy gets very excited when someone knocks at the door.

Patty's neighbour Madison also has a dog.

Madison's dog's name is Shaggy because of his long, shaggy fur.

This morning Patty saw her glasses on the table.

Now she wants her glasses.

Where does Patty think her glasses are? Does Patty know where her glasses are?

Note. Stimuli were modified stories from the ToM-Battery (Hutchins et al., 2008).
Appendix H: Sample Stimuli for the Low-Uncertainty Condition of Study 3

This is Patty.

This morning Patty saw her glasses on the table.

Yesterday, Patty’s babysitter asked her why she left her glasses on the table.

Patty responded that because she reads storybooks on the table, she always leaves her glasses on the table.

Where does Patty think her glasses are? Does Patty know where her glasses are?

Note. Stimuli were modified stories from the ToM-Battery (Hutchins et al., 2008).
Appendix I: Sample Stimuli for the High-Uncertainty Condition of Study 3

Note. Stimuli were modified stories from the ToM-Battery (Hutchins et al., 2008).

This is Patty.

This morning Patty saw her glasses on the table.

But by supper time, Patty saw that her glasses were not on the table.

So Patty asked her babysitter where her glasses are.

Patty’s babysitter said that she doesn’t know where Patty’s glasses are.

Where does Patty think her glasses are? Does Patty know where her glasses are?
Appendix J: Ethics Documents

Western University Non-Medical Research Ethics Board
NMREB Annual Continuing Ethics Approval Notice

Date: February 01, 2016
Principal Investigator: Dr. Adam Cohen
Department & Institution: Social Science/Psychology

NMREB File Number: 103137
Study Title: Cognitive development of theory of mind: Behavioral studies
Sponsor:

NMREB Renewal Due Date & NMREB Expiry Date:
Renewal Due - 2017/01/31
Expiry Date - 2017/02/06

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed the Continuing Ethics Review (CER) form and is re-issuing approval for the above noted study.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), Part 4 of the Natural Health Product Regulations, the Ontario Freedom of Information and Protection of Privacy Act (FIPPA, 1990), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.
**Project Title:** The Development of Social Abilities in Children  
**Principal Investigator:** Adam Cohen, PhD, Department of Psychology and The Brain and Mind Institute, University of Western Ontario

---

**Letter of Information**

Dear Parent,

I am an Assistant Professor and researcher in the Department of Psychology and The Brain and Mind Institute at The University of Western Ontario. I am conducting research on the development of social abilities in children along with some of the students from my research laboratory. In the following letter of information, the pronouns 'you' and 'your' should be read as referring to the participant (in this case, the child) rather than the parent/guardian/next of kin who is signing the consent form for the participant.

**Purpose of this Study**
You are being invited to participate in this research study because we are exploring how children your age begin to make sense of the social world and interact within it and how this understanding changes across development. The purpose of this letter is to provide you with information required for you to make an informed decision regarding participation in this research.

**Research Studies and Procedures**
If you participate in this study, you will be asked to listen to a story, watch a short video, or watch a scene acted out with toys. The stories and events will often include people doing everyday tasks (putting things away, talking to other people, and so on). Then, you will be asked to answer questions about what you just heard or saw, including recalling events and predicting/explaining what someone might do based on what you saw or heard.

It is anticipated that the entire task will take between 30 and 45 minutes, over one session. The task(s) will be conducted in Westminster Hall on the Western University campus. All materials will be safe for children to handle, and children will be supervised by a trained researcher at all times. We will only run sessions at pre-arranged times convenient for the parent and child.

**Possible Risks and Harms**
There are no known risks associated with participation in this study. However, if you start feeling tired while you are participating, please tell the experimenter and he or she will give you a break or stop the study session.
Possible Benefits
You may not directly benefit from participating in this study but information gathered may provide benefits to society as a whole by advancing our knowledge of cognitive development. This knowledge could inform programs for educational institutions, and perhaps assist in designing interventions for children overcoming certain developmental difficulties and/or delays.

Compensation
You will be compensated with a $15 gift certificate and a small prize for your participation in this study (for example, stickers, pencils, or a book).

Voluntary Participation
Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions, or withdraw from the study at any time.

Confidentiality
All data collected will remain confidential and accessible only to the investigators of this study. If the data are stored on a computer, they will be stored in such a way that your name will not be connected to the results. If the results of the study are published, your name will not be used and no information that discloses your identity will be released or published without your specific consent. Representatives of the University of Western Ontario's Research Ethics Board may require access to your records for the purpose of monitoring the research. The Research Ethics Board at the University of Western Ontario may contact you directly to ask about your participation in the study. If we find information we are required by law to disclose, we cannot guarantee confidentiality.
Because the results are only meaningful with respect to groups and are uninformative about individuals, parents and children will not be notified about individual performance. After the studies are completed, we will send out a newsletter to parents reporting the overall results of the research.

On behalf of my research team and myself, I would like to extend our gratitude to you for reading this letter and considering participating in our study. If you would like to participate, simply sign and return the consent form. You do not waive any legal rights by signing the consent form.

We look forward to hearing from you.

Best wishes,
Consent Form

Please keep this identical copy of the consent form for your records

Project Title: The Development of Social Abilities in Children

Study Investigator’s Name: Dr. Adam Cohen

I do not wish to participate in this study __________

If you wish to participate in this study, please fill out the information below

I have read the Letter of Information, have had the nature of the study explained to me, and I agree that ________________ may participate. All questions have been answered to my satisfaction.

__________________________
(Child’s name)

__________________________ ____________________________
(Name of parent) (Signature of parent) Date

The University of Western Ontario
Faculty of Social Science, Department of Psychology
Nellie Kamkar

Education

Master of Science in Developmental Psychology, Western University 2015-present
Honours Specialization in Psychology, Western University 2008-2015

Honours and Awards

Recipient of the Canada Graduate Scholarship, Master’s Program 2016-2017
Natural Sciences and Engineering Research Council of Canada
Value of $17,500

Recipient of the Western Graduate Research Scholarship Fall 2016
Western University
Value of $4,067

Recipient of the Western Graduate Research Scholarship Summer 2016
Western University
Value of $4,700

Recipient of the Western Graduate Research Scholarship Winter 2016
Western University
Value of $3,950

Recipient of the Western Graduate Research Scholarship Fall 2015
Western University
Value of $3,950

Recipient of the Certificate of Teaching Excellence 2015-2016
Western University
Value of $50
*The Canadian Psychological Association*
Recipient of the Clark and Mary J. Wright Scholarship 2014-2015
*Department of Developmental Psychology, Western University*

Value of $1,000
Placed on the Dean’s Honor List 2012-2015
*Western University*
Laureate of the Scholarship of Distinction 2008-2009
*Western University*

### Publications


### Poster Presentations


examination of mental state reasoning in preschool aged children. Poster presented at the 36th annual meeting of the Southern Ontario Neuroscience Association (SONA), Waterloo, ON.

**Kamkar, N., & Morton, J. B. (2015, April). The timing of stress influences whether or not children learn more from rewards or losses.** Poster presented to the Department of Psychology for the Honours Thesis Poster Presentation Day, Western University, London ON.

**Invited Talks**

**Kamkar, N.** (April 2017). Why Come to Graduate School? Graduate School and the Cohen Laboratory. *Bring an Undergrad to Grad School.* Western University, London ON.


**Teaching Experience**

**Teaching Assistant**

*Winter 2017*

*Introduction to Developmental Psychology (Psychology 2410A)*

Course Instructor: Dr. Lynne Zarbatany, *Western University*

- Taught material for two tutorial classes on a weekly basis.
- Graded written and oral presentations of a large research project.

**Teaching Assistant**

*Winter 2015*

*Special Topics in Psychology (Psychology 3991G)*

Course Instructor: Dr. Adam S. Cohen, *Western University*

- Participated in active weekly discussions for an upper-year seminar course.
- Graded written work on the topic of evolutionary developmental psychology.

**Teaching Assistant**

*Fall 2015*

*Introduction to Developmental Psychology (Psychology 2410A)*

Course Instructor: Dr. Adam S. Cohen, *Western University*

- Taught material for two tutorial classes on a weekly basis.
- Graded written and oral presentations of a large research project.
Certified Student

The Teaching Assistant Training Program (TATP)

Western University

- Successfully received certification of completion for TATP.
- TATP is an interdisciplinary course offered through the Teaching Support Centre.
- The program is intended for graduate teaching assistants and teaches various strategies and practices of University teaching.

Supervising Experience

Independent Study Student

Psychology 3996F

Department of Developmental Psychology, Western University.

Primary Supervisor: Dr. Adam S. Cohen

- Created course syllabus.
- Met regularly with student in order to develop an empirically motivated and feasible research project for course credit.
- Graded all written work pertaining to the final project.

Independent Study Student

Psychology 3997G

Department of Developmental Psychology, Western University.

Primary Supervisor: Dr. Adam S. Cohen

- Created course syllabus.
- Met regularly with student in order to develop an empirically motivated and feasible research project for course credit.
- Graded all written work pertaining to the final project.
**Volunteer Experience**

**Developmental Brown Bag Organizer** 2015-2016

*Department of Developmental Psychology, Western University.*

- Organized a series of weekly departmental talks by both graduate students and faculty in Developmental Psychology and Neuroscience.

**Undergraduate Journal Reviewer** 2015-2016

*Western Undergraduate Psychology Journal (WUPJ), Western University.*

- Reviewed empirical papers, research proposals, and theoretical review papers submitted by undergraduate students.

**Secondary School Outreach Director** 2015-present

*The Western Women in Neuroscience (Western WINS) Organization*

Co-Founders: Dr. Jessica Grahn., Niki Hosseini-Kamkar., Emily S. Nichols., & Ramina Adam. Western University.

- Organized a series of outreach talks to various secondary schools in the London and surrounding area with the purpose of engaging the interest of young scholars in the fields of Psychology and Neuroscience.

- Corresponded with both secondary school staff members as well as a diverse sample of graduate students in Psychology and Neuroscience with the intent of conveying the implicit idea that researchers can be both male and female and can come from various demographic backgrounds.
Inspiring Young Women in STEM Conference Administrator 2015-2016

The Western Women in Neuroscience (Western WINS) Organization
Co-Founders: Dr. Jessica Grahn., Niki Hosseini-Kamkar., Emily S. Nichols., & Ramina Adam. Western University.

- Organized speaker presentations from various faculty members including Dr. Christine Tenk and Dr. Anabel Quan-Haase at Western University for a conference aimed at inspiring young undergraduate women to pursue careers in Science, Technology, Engineering, and Mathematics (STEM).

- Rated poster presentations of undergraduate students in various disciplines of Science and Engineering for receipt of a set of awards.

Psychology and Neuroscience Outreach Event Organizer 2015-2016

The Canadian Association for Girls in Science (CAGIS)
Co-Founders: Dr. Evelyn Vingilis., Niki Hosseini-Kamkar., & Erin Shumlich. Western University.

- Organized an outreach program aimed at engaging the interest of young pre-adolescent girls in Psychology and Neuroscience.

Volunteer Speaker 2014-2015

The Western Women in Neuroscience (Western WINS) Organization
Co-Founders: Dr. Jessica Grahn., Niki Hosseini-Kamkar., Emily S. Nichols., & Ramina Adam. Western University.

- Voluntarily gave several presentations on selected topics in Psychology and Neuroscience at various secondary schools.
Fundraising Committee Volunteer

Students Rebuilding Health in Rwanda, Western University.

- Assisted in the organization of various fundraising events aimed at improving health care facilities available in Rwanda.

Research Experience

Honours Student

Dr. Bruce Morton’s Developmental Cognitive Neuroimaging Laboratory

Western University

- Conducted an experiment aimed at understanding the relationship between stressful life events and reward learning profiles in young children with a critical focus on the timing of stressful experiences.

Research Assistant

Dr. Bruce Morton’s Developmental Cognitive Neuroimaging Laboratory

Western University

- Preprocessed a set of neuroimaging data obtained using functional Magnetic Resonance Imaging (fMRI) from healthy participants and participants with Multiple Sclerosis (MS) using Statistical Parametric Mapping (SPM) software.

- Recruited participants from a developmental population pool by providing prospective participants with detailed information pertaining to a developmental neuroimaging study of genetic and environmental influences on cognitive and behavioural self-regulation.

- Ran child participants (aged 9-12) and their parents through a mock scanning procedure to ensure the child's comfort with fMRI and assisted in running participants through an fMRI scanning procedure at the Robarts Research Institute.

- Transferred written questionnaire data onto Microsoft Excel and converted raw scores into statistically appropriate values for formal behavioural analyses in a longitudinal study for PhD candidate Heather Wilk.
Research Assistant

2010-2013

Dr. William Fisher's Social Psychology Laboratory, under the supervision of Dr. Kohut

Western University

- Compiled an online questionnaire using the Qualtrics database for a study in collaboration with Dr. Lorne Campbell

- Provided appropriate statistical formulas for a post-doctoral study of Dr. Salisbury

- Had the opportunity to format and review a now published research article in accordance with proper APA 6 standards.

- Had exposure running participants in the laboratory in accordance with the Ethical guidelines (obtained informed consent, respected confidentiality, and debriefed each participant appropriately upon completion of the study)

- Met regularly to enhance inter-rater reliability by extensively discussing and correcting any potential biases in coding

- Independently completed qualitative analysis on responses from a large scale participant survey using Microsoft Excel

Other Professional Experience

Administrative Assistant,  

2009-2010

Department of Statistical and Actuarial Sciences

Western University

- Performed daily administrative tasks for graduate students at Western University

- Assisted in the admissions process for globally aspiring MSc and PhD students

- Prepared and administered examinations in the field of statistical and actuarial sciences

- Scheduled appointments and assigned tasks to teaching assistants to ensure their coursework, research, and teaching priorities do not conflict

- Reviewed computations of financial budgets allocated to the department through government grants
Strengths

• Attended a senior level neuroimaging course with Dr. Peter Bandettini for neuroscientists at the 44th annual meeting of the International Neuropsychological Society (INS) in Boston, Massachusetts.

• Capable of preprocessing and analyzing neuroimaging data obtained from fMRI studies using Statistical Parametric Mapping (SPM) software.

• Exceptional oral and written presentation skills (scored in the 99th percentile on the analytic writing portion of the Graduate Record Examination)

• Superb arithmetic skills (earned a final grade of 87% in Methods of Matrix Algebra)

• Computer coding experience using MATLAB (earned a final grade of 94% in a graduate course on scientific computing)

• Experienced with various computer applications including MATLAB, SPM8, R, SPSS, OpenSesame, Psychopy, and Qualtrics.

• Extremely familiar with 6th edition APA formatting

• Excellent administrative skills including an 96 Word Per Minute (WPM) type rate

• Thorough understanding of human social psychology (earned a final grade of 92% in Social Psychology)

• Considerable familiarity with both ethical and procedural guidelines concerning laboratory work

• Intrinsically motivated to perfect the task or project assigned