Children's Understanding of the Opaque and Transparent Uses of Language

by

Deepthi Kamawar

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Human Development and Applied Psychology
Ontario Institute for Studies in Education of the University of Toronto

© Copyright by Deepthi Kamawar 2000
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Children's Understanding of the Opaque and Transparent Uses of Language

By

Deepthi Kamawar, Doctor of Philosophy, 2000
Department of Human Development and Applied Psychology
University of Toronto

Abstract

A developmental study of children's sensitivity to opaque contexts is reported. Sentences are considered opaque if the substitution of a term with a co-referential term can change the truth of the sentence (e.g., "John thinks that Venus is the Morning Star"). Opaque contexts often involve propositional attitude verbs such as 'know' and 'think'. Sensitivity to opacity requires understanding that a referent can be represented in different ways, and that the specific description that is used is important.

Two studies were conducted investigating children's ability to deal with opacity. In the first, 4- to 8-year-olds were given a variety of opacity tasks (referential, linguistic and intentional), false belief tasks, digit span, and vocabulary measures. The results of the first study indicated that children were more successful on opaque contexts with action referents than those with object referents. The results also demonstrated that children find referential opacity easier than linguistic opacity, and both easier than intentional opacity. Nonetheless, performance across the three types of opacity was related even when age, vocabulary and digit span were statistically controlled, supporting the view that there is an underlying commonality for the three contexts.

The second study examined whether performance on referentially opaque contexts could be predicted by metarepresentational ability and metalinguistic awareness. Metarepresentational ability was measured using false belief tasks and
metalinguistic ability was measured using tasks requiring children to compare and evaluate statements containing referring expressions. In this second study, 5- to 7-year-olds were given opacity (referential and quotational), false belief, metalinguistic awareness, digit span, and vocabulary measures. Hierarchical regressions indicated that even after the variability from children's age, vocabulary scores and digit span performance is taken into account, metarepresentational ability and metalinguistic awareness still significantly, and independently, explained some of the remaining variability in performance on referential opacity tasks. These results are taken as support for the view that both metarepresentational ability and metalinguistic awareness are necessary in order to deal with referentially opaque contexts.

These studies are of interest to anyone concerned with: (1) children's understanding of the representational nature of language; and (2) the necessary abilities for sensitivity to opaque contexts.
Acknowledgements

I would like to thank Professor David Olson for his advice and encouragement. His comments and questions were always helpful. One of the reasons I decided to pursue a Ph.D. was because of how much fun David made the learning process; I am grateful to him for making my time at OISE/UT enjoyable. I also want to thank Professor Janet Astington for sharing her time and ideas with me. Over the years, she has been a source of support, guidance and expertise, for which I am very grateful. I also appreciate the comments that Professor Phil Zelazo offered, both during the thesis preparation, and during the defence. I would like to thank Professor Ron Smyth for his thought-provoking (and challenging) questions during the defence, and Professor Daniela O'Neill for raising interesting questions for discussion.

I would like to acknowledge the generous help of the principals, teachers, and students of The Scarborough Board of Education, especially my parents-in-law, Don and Donna MacLeod, who welcomed me into their schools. I am also grateful to Sue Elgie for sharing her statistical knowledge with me; her help was invaluable.

Thanks also to my parents for their ongoing support and for their continuous words of encouragement. Their pride was always evident, and much appreciated.

Finally, to my husband Mark: thee, thee, thee.
# Table of Contents

Abstract ....................................................................................................................... ii
Acknowledgements .................................................................................................... iv
Table of Contents ........................................................................................................ v
List of Tables ................................................................................................................ ix
List of Figures .............................................................................................................. xi
List of Appendices ....................................................................................................... xii

Chapter 1. Introduction .............................................................................................. 1
  1.1 Goal of Study 1 ................................................................................................. 3
  1.2 Goal of Study 2 ................................................................................................. 3
  1.3 Overview of Chapters ....................................................................................... 4

Chapter 2. Philosophical and Theoretical Background .............................................. 7
  2.1 Sense-Reference Distinction ............................................................................ 7
  2.2 Opaque Contexts ............................................................................................. 8
  2.3 Referential Opacity .......................................................................................... 10
  2.4 Intentional Opacity ......................................................................................... 11
    2.4.1 Searle ..................................................................................................... 11
    2.4.2 Davidson ................................................................................................. 16
    2.4.3 Comparison of Searle and Davidson ....................................................... 18
    2.4.4 Intentionally Opaque Contexts .............................................................. 20
  2.5 Linguistic Opacity ........................................................................................... 20
  2.6 Quotational Opacity ....................................................................................... 20
  2.7 Difference Between Linguistic and Quotational Opacity ............................... 21

Chapter 3. Cognitive Requirements .......................................................................... 23
  3.1 Referential Opacity ......................................................................................... 23
3.1.1 Metarepresentational Ability ........................................... 23
3.1.2 Partial Knowledge ...................................................... 27
3.1.3 Metalinguistic Awareness ............................................... 28
3.1.4 Comparison of false belief and referential opacity .......... 31

3.2 Intentional Opacity ................................................................. 32
3.3 Linguistic Opacity................................................................. 33
3.4 Quotational Opacity ............................................................... 33

Chapter 4. Related Research .......................................................... 34
4.1 Multiple Naming ................................................................. 34
4.2 Referential Opacity .............................................................. 35
4.3 Intentional Opacity ............................................................... 42
4.4 Linguistic Opacity ............................................................... 47
4.5 Quotational Opacity .............................................................. 51
4.6 Metalinguistic Awareness ...................................................... 52

Chapter 5. Study 1 ....................................................................... 60
5.1 Purpose ................................................................................... 60
5.2 Hypotheses ............................................................................. 60
    Hypothesis 1 ........................................................................ 60
    Hypothesis 2 ........................................................................ 61
    Hypothesis 3 ........................................................................ 61
    Hypothesis 4 ........................................................................ 61
    Hypothesis 5 ........................................................................ 62
    Hypothesis 6 ........................................................................ 62
    Hypothesis 7 ........................................................................ 62
    Hypothesis 8 ........................................................................ 62
5.3 Participants .............................................................................. 62
5.4 Materials ................................................................................ 63
    5.4.1 Opacity Tasks .............................................................. 63
    5.4.2 Memory Measure ....................................................... 68
    5.4.3 Language Measure ..................................................... 68
    5.4.4 False Belief Tasks ....................................................... 69
5.5 Procedures and Design ............................................................ 70
Chapter 6. Study 2

6.1 Purpose

6.2 Hypotheses

Hypothesis 1
Hypothesis 2
Hypothesis 3
Hypothesis 4
Hypothesis 5
Hypothesis 6

6.3 Participants

6.4 Materials

6.4.1 Opacity Tasks
6.4.2 Memory Measure
6.4.3 Language Measure
6.4.4 False Belief Tasks
6.4.5 Say/Refer Task
6.4.6 Say/Mean Task

6.5 Procedures and Design

6.6 Results

6.6.1 Methods of Analysis
6.6.2 ANOVA Results
6.6.3 Hierarchical Regressions
6.6.4 Correlational Analyses
6.6.5 False Belief Tasks
6.7 Discussion ................................................................. 136
   6.7.1 Age ................................................................. 136
   6.7.2 Opacity ............................................................. 136
   6.7.3 Predicting performance on referentially opaque contexts .................................. 137

6.8 Summary and Conclusion ........................................... 139

Chapter 7. General Discussion and Conclusions ..................... 141
   7.1 Study 1 ............................................................... 141
   7.2 Study 2 ............................................................... 144
   7.3 Studies 1 and 2 ....................................................... 151
   7.4 Future Research ..................................................... 152
   7.5 Conclusion ........................................................... 155

References ........................................................................... 156

Appendices ........................................................................ 161
List of Tables

Table 1: Descriptive statistics for referent, opacity and referent by opacity scores for Study 1. .................................................................73

Table 2: Performance on intentional opacity by performance on referential opacity across referent types ..................................................178

Table 3: Performance on intentional opacity by performance on referential opacity for object referents ......................................................178

Table 4: Performance on intentional opacity by performance on referential opacity for action referents ......................................................179

Table 5: Performance on linguistic opacity by performance on referential opacity across referent types ...................................................179

Table 6: Performance on linguistic opacity by performance on referential opacity for object referents .....................................................180

Table 7: Performance on linguistic opacity by performance on referential opacity for action referents .....................................................180

Table 8: Performance on the false belief tasks by age .............................................180

Table 9: Correlation matrix for all measures ........................................................182

Table 10: Partial correlation matrix: Age, vocabulary and digit span performance partialled out .................................................................91

Table 11: Descriptive statistics for opacity and metalinguistic tasks for Study 2 ......................................................................................122

Table 12: Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and the say/refer task .........................................................................................126

Table 13: Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and say/mean – judgment .........................................................................................128

Table 14: Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and say/mean – evaluation .................................................................................130

Table 15: Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and metalinguistic awareness measures ..............................................131

Table 16: Correlation matrix for all measures ........................................................133
Table 17: Performance on the false belief tasks by age ........................................ 196
List of Figures

Figure 1: Mean scores (+SE) on referential, linguistic and intentional opacity by age ................................................................. 75
Figure 2: Mean scores (+SE) on object and action referents by age .......... 77
Figure 3: Mean scores (+SE) on referential, linguistic and intentional opacity by age ................................................................. 78
Figure 4: Mean scores (+SE) on linguistic and intentional opacity by ESL status .............................................................................. 80
Figure 5a: Mean scores and 95% confidence intervals for intentional opacity by ESL status .......................................................... 82
Figure 5b: Mean scores and 95% confidence intervals for linguistic opacity by ESL status ............................................................. 83
Figure 6a: Interaction among opacity, ESL status and referent type (age = 6 years) ................................................................. 85
Figure 6b: Interaction among opacity, ESL status and referent type (age = 8 years) ................................................................. 86
Figure 7: Mean scores (+SE) on false belief tasks by age ......................... 90
Figure 8: Mean scores (+SE) on referential and quotational opacity by age .. 123
Figure 9: Mean scores (+SE) on false belief tasks by age ......................... 135
## List of Appendices

Appendix A: Information Letter and Consent ......................................................... 161
Appendix B: Opacity Stories and Questions for Study 1 ........................................... 163
Appendix C: False Belief Tasks ............................................................................. 176
Appendix D: Tables for Study 1 (Chapter 5) ............................................................. 178
Appendix E: Correlation Matrix for All Measures in Study 1 ................................. 182
Appendix F: Opacity Stories and Questions for Study 2 ......................................... 183
Appendix G: Say/Refer Task .................................................................................. 189
Appendix H: Say/Mean Task .................................................................................... 193
Appendix I: Table for Study 2 (Chapter 6) ............................................................... 196
Chapter 1: Introduction

In recent years, a great deal of research has examined children's developing ability to metarepresent, that is, their ability to represent a representation (Leslie, 1988) or "... to represent that something (another organism) is representing something..." (Perner, 1991, p. 7). According to many, children are said to have reached this milestone when they realize that another's thoughts are only representations of the world, and as such, can differ from it or be false of it. This is no small achievement. The development of the ability to consider another's beliefs, desires, and knowledge of the world (in other words, another's representations) when predicting her actions or her expectations has often been referred to as the acquisition of a Theory of Mind (ToM) (e.g., Astington, 1993; Astington, Harris & Olson, 1988; Wimmer & Perner, 1983). While there has been much debate as to whether what is going on actually constitutes the acquisition of a theory (e.g., Harris, 1991), the fact remains that the child has stumbled upon something that we, at least, have a theory about. In explaining children's new found proficiency with mental representations, we appeal to the notion of metarepresentation. In the case of ToM, we say that the child has become able to represent another's beliefs (which are representations themselves) and can predict the actions (or expectations) of another based on those beliefs (desires, etc.). Thus, for many researchers (e.g., Astington, 1993; Perner, 1991), the ability to recognize a false belief in another is taken as an indication of the ability to metarepresent.

However, until recently, there has not been a great deal of research examining
other tasks in which the ability to metarepresent is equally important to successful performance, nor has there been much research determining whether performance on one metarepresentational task is related to performance on others, or whether the relation holds only for tasks within the area of theory of mind. Thus, the general focus of this thesis is on the relation between performance on false belief tasks and performance on another group of metarepresentational tasks, namely those involving opaque contexts.

In brief, opaque contexts are contexts in which you cannot 'see through' the description of the referent to the referent itself, hence the metaphorical name 'opaque'. In such situations, one should not secure reference to the referent simply by determining what a referring expression is denoting; in such contexts, the referent is being picked out in one way, and not necessarily others. Opaque contexts, more formally, are those which contain both a proposition and a mental attitude toward the proposition; for example, 'Mary knew [mental attitude] that the cat was in her favourite tree [proposition]'. In such a context, the substitution of a coreferential term (e.g., 'the oldest oak tree on Major Street' for 'her favourite tree') can affect the truth value of the resulting sentence as Mary may not know anything about the age or type of her favourite tree, and therefore will not have represented it as 'the oldest oak tree on Major Street' (this topic will be discussed in detail in Chapter 2). Tasks involving opaque contexts will require children to judge whether a particular propositional attitude can be ascribed to another. For example, if 'Mark knows that he took the big ball' but does not know that the big ball is in fact yellow, we can ask 'Does Mark know that he took the yellow ball?' (correct answer: 'no'). Such tasks will be used: (a) to determine children's development on such tasks; and (b) to determine the relation between performance on one type of metarepresentational task (e.g., false belief, opacity) and performance on the others.
1.1 Goal of Study 1

We know from previous work (Apperly & Robinson, 1998; de Villiers & Fitneva, 1996; Kamawar, 1996) that successful performance on false belief tasks is not equivalent to successful performance on referentially opacity tasks (contexts with mental verbs such as 'think' and 'know'). Referentially opaque contexts are generally found to be more difficult. However, there exists an assortment of opaque contexts which share the defining feature of not allowing substitution of coreferential terms in the proposition without possibly affecting the truth value of the resulting sentence. It is also the case that the referent of the key term in the proposition can be an object or an action. This brings us to two interesting questions: (1) Is it the case that once a child is able to perform successfully on one type of opacity task (such as a referentially opaque context), she is then able to perform successfully on all other types of opacity?; and (2) Is performance for a given type of opacity identical for both object and action referents? If the answer to first question is affirmative, then the ability to deal with opacity can be viewed as an all-or-nothing competence. Alternatively, it may prove to be the case that children become proficient with some types of opaque context earlier than others. If the answer to the second question is affirmative, then performance on opaque contexts is not affected by type of referent (object vs. action). Alternatively, it may be the case that one type of referent is easier than the other for reasons that remain to be explored. Thus, the goal of Study 1 is to use a variety of opacity tasks (referential, linguistic, and intentional), which at the very least require metarepresentational ability for successful performance, and both action and object referents in an effort to answer the above questions.

1.2 Goal of Study 2

Given that there exists a difference between performance on false belief tasks and referentially opaque contexts, we are left with the task of accounting for the
discrepancy in performance between these two metarepresentational tasks. If successful performance on tasks assessing referentially opacity required only the ability to metarepresent, we should find an equivalence relation between performance on the two tasks but we do not (Apperly & Robinson, 1998; de Villiers & Fitneva, 1996; Kamawar, 1996). This brings us then to the goal of Study 2, which set out to answer what, in addition to metarepresentational ability (as evidenced by successful performance on false belief tasks), is required for successful performance on tasks of referential opacity. Given that opacity tasks, in addition to being metarepresentational tasks, are also fundamentally metalinguistic1 tasks, it is expected that metalinguistic awareness will play an important role in being sensitive to opaque contexts. The secondary goal of this study is to examine the developmental trajectory of children's performance on contexts that are opaque by virtue of containing direct quotes (quotational opacity).

The questions addressed by this research fit both within the framework of the philosophies of language and mind, the areas concerned with issues such as how mental verbs affect the substitutability of terms in sentences and the role of intentions and Intentionality2 when describing behaviour (e.g., Davidson, 1971/1980; Frege, 1892/1991; Quine, 1955/1991; Searle, 1983). Further, they fit within the theoretical framework of the psychology of mind, an area that addresses issues such as representation, metarepresentation, beliefs, intentions, etc. (e.g., Astington, Harris & Olson, 1988; Perner, 1991).

1.3 Overview of Chapters

Chapter Two will begin by discussing the philosophical background pertaining to a variety of opaque contexts including referential, linguistic, intentional, and

---

1 In an opaque context, the mental attitude is toward a proposition, which is a linguistic entity.

2 In keeping with convention, I will used a capital T Intentionality when using the term in its philosophical sense.
quotational opacity; more specifically, it will lay out the concepts and the relevant contributions made by philosophers such as Frege, Quine, Searle and Davidson. It will then go on to address some theoretical issues relating to theory of mind.

Chapter Three outlines the cognitive requirements that I expect to be necessary to perform successfully on the key tasks (i.e., the opacity tasks). Chapter Four consists of a review of the relevant research literature on theory of mind, referential opacity, intentional opacity, linguistic opacity, quotational opacity and metalinguistic awareness.

Chapter Five presents Study 1 which set out to address the first goal of the thesis. In this first study, 4- to 9-year-olds were given referential, linguistic, and intentional opacity tasks along with false belief, vocabulary and digit span measures. It was found that children's performance on the three opacity tasks and the false belief tasks differs significantly, with false belief easiest and intentional opacity hardest. Another result was that children found opaque contexts with action referents significantly easier than those with object referents.

Chapter Six presents Study 2 which addresses the question of what else is needed, in addition to metarepresentational ability (as evidenced by successful performance on the false belief tasks), in order to succeed on the easiest opaque context (referential). In this second study, 4- to 7-year-olds were given referential opacity, quotational opacity, false belief, digit span, and vocabulary measures. Children were also given two tasks designed to measure their metalinguistic awareness: (1) the say/mean task (Robinson, Goelman & Olson, 1983); and (2) a task designed specifically for this study that requires children to be able to identify when an utterance matches one, both or neither of the two utterances presented before it (say/refer task). By performing a regression analysis, the results of this study indicate that even after the variability from children's age (measured in months), their vocabulary scores and their digit span tasks is taken into account, metarepresentational ability (false belief
performance) still accounts for a significant amount of the remaining variability on referential opacity tasks. Further, even after all of the above variables are entered into the regression equation, metalinguistic awareness (i.e., say/refer) still significantly explains some of the remaining variability.

Chapter Seven presents a general discussion of the two studies and completes the thesis with some concluding thoughts and ideas for future research.
Chapter 2: Philosophical and Theoretical Background

2.1 Sense-Reference Distinction

In order to examine when children realize that the meaning of a word or name does not necessarily coincide with what it refers to in the world, it is necessary to find some way to describe the distinction that is developed. This distinction has been examined in the philosophy of language.

Frege (1892/1991), in "On Sense and Nominatum", demonstrates that the Reference Theory cannot be an accurate representation of language. He does so by using the following name and descriptors as an example: Venus (name); 'the Morning Star' (description 1); and 'the Evening Star' (description 2). Frege points out that since all of these names/descriptions refer to the same object, and are therefore coreferential, they are said to have the same meaning according to the Reference Theory. Therefore, according to that theory, it should be possible to substitute one of these terms for the other without changing the meaning of a sentence. However, saying that 'Venus is the Evening Star' means something different than 'Venus is Venus' even though a coreferential term has been replaced for 'the Evening Star'. Frege concluded that meaning is not located in the object, and therefore, the Reference Theory of meaning is inaccurate and insufficient. A theory to account for meaning must include something more.

Since it seems that "every object can be referred to by many words, and many words can represent several different objects" (Bryen, 1982), it is necessary to examine
how this can be so. There must exist something that mediates between the word and the object: this something is meaning. Frege (1892/1991) posited that the mediating factor was the *sense* of a word. The *sense* represents a particular way of connecting a referent to its sign/name, and is objective (unlike the *idea*). The object need not exist in order for the name to refer (for example, 'Pegasus'). The sense of a name or description is a function of how it picks out an object; Venus is picked out differently when it is referred to as: (a) the 'Evening Star', the first star visible in the evening; (b) the 'Morning Star', the last star visible in the morning; or (c) the second planet from the sun. The ways in which the descriptions refer are different; therefore, they have different senses. However, it is possible for names to have the same sense; for example, 'snow' and 'neige' (French).

In order to account for the subjectivity of experience and the fact that individuals may have their own specific meanings for a word that are not shared, Frege introduces the further concept of an *idea*. An idea is a subjective image that is different for each individual because it arises from each person's memories of sense impressions and all the actions she has performed. By designating subjectivity as being in the realm of the idea, Frege can maintain that the sense (or senses) is objective; it is available to all.

As mentioned, it may be argued that children initially operate using the Reference Theory of meaning. They begin by thinking that all that a name does is pick out an object. By the time they are adults they are able to make the sense-reference distinction as evidenced by sensitivity to referentially opaque contexts (see below).

2.2 *Opaque Contexts*

Opaque contexts can be best understood by contrasting them with transparent contexts; ones in which the truth value of a sentence does not change when coreferential terms are substituted for one another. For example, if 'John purchased a picture of the Morning Star' is true, then so is, 'John purchased a picture of Venus'.
these statements are about what John did, not about what he thought he did. A context is opaque if replacing a term within it with a coreferential term can affect the truth value of the whole sentence (Hookway, 1988). However, in one opaque context, namely, that involving direct quotations, the substitution of a term necessarily alters the truth value of the resulting sentence. For example, if John said, 'I have the cat', it would be false to say 'John said, 'I have a cat'.'

Three categories of opaque contexts have been widely noted by philosophers. The first type of context is those with propositional attitudes (mental verbs). For example, while it might be true that 'John thinks that Venus is the Morning Star', it is not necessarily true that 'John thinks that Venus is the Evening Star' as John may not know the term 'Morning Star' is coreferential with 'Evening Star'. The second type consists of statements which contain another's quoted speech. For example, if 'John said, "I would like some supper"', even though the terms 'I' and 'John' are coreferential, the resulting statement 'John said, "John would like some supper"' would not be true. The third type of context often addressed by philosophers is that which is opaque by virtue of containing a modal verb. Modals are "auxiliary verbs that express notions like permission, possibility, obligation and futurity (e.g., 'may', 'must', 'can', 'will')" (O'Grady & Dobrovolsky, 1987, p. 452). For example, consider the following two sentences from Quine (1960):

1) Necessarily 9 > 4. [true]
2) Necessarily the number of major planets > 4. [false]

Since '9' and 'the number of major planets' are coreferential terms, we should be able to substitute one for the other without affecting the truth value of the resulting sentence (if this were a transparent context). Since this is not the case, as the first statement above is true while the second is false, we can conclude that the "... necessity operator is opaque" (Quine, 1960, p. 197).

While many solutions have been offered for the problem of referential opacity,
for the moment I will only mention Frege's (1892/1991). His explanation is that when terms occur in an opaque context, they do not have their usual referent but instead refer to what is usually their sense (meaning). Therefore, in order to replace one term with a coreferential term in an opaque context, the terms must each have the same sense.

For the purposes of this thesis, I am concerned with four types of opaque contexts: the first two types of opacity addressed by philosophers of language (contexts containing propositional attitudes and direct quotes), as well as two other types. The first, a variant of propositional attitude opacity, deals with contexts that are opaque because they contain a verb expressing an intention (i.e., purposeful behaviour) toward a referent under a particular representation. The second, a variant of quoted opacity, deals with contexts that are opaque by virtue of containing indirect quotations of another's speech. I will now address each of these in turn.

2.3 Referential Opacity

The term 'referential opacity' will be used when describing propositional attitude opacity and applies specifically to contexts containing the mental verbs 'know' or 'think'. The use of this term is common to a number of researchers who study this topic (e.g., Apperly & Robinson, 1998; de Villiers, 1994).

Understanding that objects can be referred to by multiple descriptions/names is a necessary, but not sufficient, condition for understanding when certain contexts are referentially opaque. Remember the example we saw earlier: While it might be true that 'John thinks that Venus is the Morning Star', it is not necessarily true that 'John thinks that Venus is the Evening Star' as John may not know the term 'Morning Star' is coreferential with 'Evening Star'.

In this context, the substitution of one coreferential term has the potential for changing a true sentence into a false one. Thus, the verb 'think', and other like mental
verbs, produce opaque contexts.

2.4 Intentional Opacity

Blackwell's *A Companion to the Philosophy of Mind* (1995), has three separate entries for intentionality. Two of the entries are about Intentionality defined as: "that property of the mind by which it is directed at, about, or of objects and states of affairs in the world. ... [It] includes such mental phenomena as belief, desire, intention, hope, fear, ... and intentional action" (Searle, 1995, p. 379). The other entry is about intentionality as it is commonly used; that is, as a term used to refer to mental states or actions that have a plan or a purpose, and are explainable by what the agent desired and believed. "To classify an action as intentional, or as done with a certain intention, is, then, to say it has the appropriate *relation* to the agent's desires and beliefs" (Bratman, 1995, p. 375). I will discuss each of these in turn.

As can be seen in the quote from Searle above, intentions (in their common sense) are one sort of Intentional state. However, before discussing intentions specifically, it would be useful to discuss Intentionality in its technical, philosophical sense.

2.4.1 Searle

In trying to the answer the question of what it is to be an Intentional state, Searle posits a model based on speech acts. He states that "Intentional states represent objects ... in exactly the same sense that speech acts ..." do (Searle, 1978/1982, p. 260). He notes that the distinction made between illocutionary force and propositional content of speech acts applies just as well to Intentional states because all Intentional states consist of a representative content in addition to a psychological mode" (Searle, 1983). For example, when I say 'Mariko goes to school', the speech act has a propositional content (that Mariko goes to school) and an illocutionary force (Assertive). Analogously, when I
believe that Mariko goes to school, my Intentional state has a representative content (that Mariko goes to school) and the psychological mode of belief. I can perform different speech acts that have the same content but different illocutionary forces. For example, by commanding Mariko to go to school, I have changed the illocutionary force (Directive) but not the propositional content.

The way that one’s mental states are related to the world depends upon their direction of fit (Searle, 1983). Beliefs have a ‘mind to world’ direction of fit; that is, the mind has to fit what actually exists in the world in order for the belief to be true. If this is not the case, it is not the world that needs to change, but the mind (the belief about the world). In contrast, intentions and desires have a ‘world to mind’ direction of fit, and as such, cannot be said to be true or false; rather, they are fulfilled or not fulfilled. In order to fulfill a desire or intention, it is not the Intentional state that has to change, but the world. For example, desire can be seen as having a ‘simple link’ to fulfillment whereas intention has a more complex one (Astington, 1993):

\[
\begin{align*}
\text{Desire:} & \quad \text{desire} \rightarrow \text{outcome} \\
\text{Intention:} & \quad \text{desire} \rightarrow \text{intention} \rightarrow \text{action} \rightarrow \text{outcome}
\end{align*}
\]

The desired outcome for an intention is achieved through the action of an agent whose intention causes that action. The representation of intention is not \textit{intend}(X happens) nor \textit{intend}(I do it) but, \textit{intend}(I do it because I intend to do it). In other words, while my desire to have the door closed is fulfilled if the door is closed (regardless of who does it), my intention to close the door is only fulfilled if I close the door. In this way, intentions are causally self-referential, since their propositional content refers to the intention of the agent (Searle, 1983). If they were not causally self-referential, they would be the same as desires and my intention to close the door would be fulfilled regardless of who closes it.

This use of intention, the common sense one, is most relevant to this thesis.
Searle (1980, 1983) makes a distinction between two aspects of intention, 'prior intention' and 'intention in action', and offers an explanation of the relation between intention and Intention. Intention in action is intentional and is the experience of acting that one has while performing an intentional action; it only causes the bodily movement. It has the characteristic linguistic form of "I am doing A" and is the Intentional component of the action. Prior intentions are the intentions one may have before performing the action; they cause the entire action. Prior intentions have the characteristic linguistic form of "I will do A" or "I am going to do A"; they are intentions formed prior to the act itself. In order to clarify the relation between intentions in action and prior intentions, Searle (1983, p. 94) offers the following diagram:

```
 prior intention ----> intention-in-action ----> bodily movement
                      action
```

The above model depicts a successfully performed intentional action; that is, an action that came about 'in the right way' so as to satisfy the intention. An intentional action consists of an intention in action and bodily movement where the movement is considered the Intentional object of the action. Alternatively, people may perform an intentional action without having formed prior intentions, and they can form prior intentions to perform an action, and then not act on them. As can also be seen in the diagram, by definition, there are no actions without intentions. The prior intention represents the whole action as the rest of its conditions of satisfaction, but the intention in action presents, but does not represent, the physical movement and not the whole action as the rest of its conditions of satisfaction. In the former case the whole action is the 'Intentional object'; in the latter case the movement is the 'Intentional object'. The intention in action, like the prior intention, is self-referential in the sense that its Intentional content determines that it is satisfied only if the event that is its condition of satisfaction is caused by it (p. 93).

Searle (1983) discusses a situation that is directly relevant to the notion of intentional opacity; namely, the accordion effect. The accordion effect refers to a
situation in which there are complex intentions and multiple descriptions of a physical action (i.e., bodily movement). One example Searle uses is that of Oedipus marrying his mother. Of this he states that even where there is an unintentional action such as Oedipus’s marrying his mother, that is only because there is an identical event which is an action he performed intentionally, namely, marrying Jocasta. There are many states of affairs without corresponding beliefs and many states of affairs without corresponding desires, but there are in general no actions without corresponding intentions. (p. 82)

He also provides another (and perhaps better known) example, that of Gavril Princip and his murder of the Archduke Ferdinand. Of this murder it can be said of Princip that he

- pulled the trigger
- fired the gun
- shot the Archduke
- killed the Archduke
- struck a blow against Austria
- avenged Serbia (p. 98)

Searle notes that each item on the list is related to those preceding and succeeding it, for example, he fired the gun by means of pulling the trigger etc. Some, but not all, of these relations are causal. It is the case that all of the above could be considered true answers to the question "What are you now doing?" when asked of Princip (where that question asks, "What intentional action are you now performing or trying to perform?"). The test that shows whether or not they are to be counted as part of the content of the intention in action is "what counts as succeeding or failing?" (p. 99). For example, if Princip’s only intention had been to fire the gun (so as to startle someone) as long as that happened he could consider himself successful even if that action could not be redescribed as 'avenging Serbia' or 'striking a blow for Austria'. The other descriptions are just unintended occurrences that happened as a result of Princip’s action.

One problem that this leads to is that there are other ways to describe what
happened that we aren't concerned with, such as 'Princip moves some molecules around' and the question becomes "How do we distinguish between those aspects of the complex event under which it is an unintentional action and those aspects which are so far from the intention that under them it is not an action at all?" (p. 102). Dascal and Gruengard (1981; as cited in Searle, 1983) present a rough possible criterion for such a distinction: we count an action as unintentional under those aspects which, though they may not be intended, are however, within the field of possibility as being intentional actions of the agent (from our perspective).

Searle wants to make it clear that he does not think that saying that 'an event can be intentional under some descriptions but not under others' means anything. Rather, he wants to say that a complex event which constitutes the action will have other features that are not presented by the Intentional content of the intention in action; for example, Oedipus intended to marry Jocasta but when he married Jocasta he was marrying his mother. 'Marrying his mother' was not a part of the Intentional content of the intention in action but it was something that happened anyhow. According to Searle, while it can be said that the action was intentional under the description "marrying Jocasta" but not under the description "marrying his mother", all this means is that

the total action had elements which were parts of the conditions of satisfaction of the intention in action and other elements that were not. It is misleading to state these facts about actions in terms of descriptions of actions because it suggests that what matters is not the action but the way we describe the action, whereas, according to my account, what matters are the facts that the descriptions describe. ... The sense in which one and the same event or sequence of events can be both an intentional action and an unintentional action has no intrinsic connection with linguistic representation but rather with Intentional presentation. (p. 101)

The main point to take from the above excerpt is that for Searle, the linguistic description used to refer to an action has no fundamental connection with the way in which a given event can considered both intentional and unintentional. It appears that,
for Searle, the intention is bound up with the nature of the action (and the Intentional content of the intention in action), and not with its linguistic form. Other theorists, specifically, Donald Davidson, offer a quite different approach to the same issue.

2.4.2 Davidson

Davidson too has addressed the accordion effect (Feinberg, 1965; cited in Davidson, 1971/1980). He gives the example of a man who, when entering a room, flicks the switch, turns on the light, and startles a burglar (Davidson, 1971/1980). Davidson notes that the man has performed one physical action, but that the action can be redescribed in many ways, and only under some descriptions is the action considered intentional. While it can be said that 'he meant to flick the switch' and 'he meant to turn on the light', it cannot be said that 'he meant to startle the burglar' as he did not even know that there was one. Davidson's approach to the accordion effect is somewhat different than Searle's. This difference is a result of Davidson's particular theory of events, actions and intentions, to which I will now turn.

An important distinction that Davidson makes is that between actions and events (Evnine, 1991). All actions, whether they be typical mental actions (like calculating a sum) or typical physical actions (like moving a table), are events; in other words, they are particular, dated occurrences. However, not all events are actions. An event is an action only if it is intentional under some description. Thus, actions can be seen as a subset of events:

```
  events
    _______________________
    |                      |
    |    actions           |
    |______________________|
```

In order for anything to be an action, something someone does, rather than something that merely happens, the event must have some description under which it is
intentional (done with purpose). While as a particular event, a given action can be described in many ways, not all of the descriptions are equally revealing. Some of these descriptions will reveal the event as an action; that is, as an event performed for a reason or with an intention. Or, more formally:

(C) For an event $e$, under a description $d$, to be an intentional action, it must have been caused by something which was a reason for $e$ under description $d$. (p. 47)

Davidson also discusses the example of Gavrilo Princip. Of it he says that if Princip intentionally fired the gun, but did not intentionally kill the archduke, and if his firing the gun is the same event as his killing the archduke, then one and the same action is intentional under some descriptions and not under others. It is here that we see the greatest difference between Searle and Davidson. Davidson maintains that "being intentional is not something which qualifies events themselves, but events as described in one or another way" (Evnine, 1991, p. 41, my italics).

The descriptions that are used to pick out some event may refer to many different things, people, or other events which themselves are not parts or participants in the event that is being referred to (Evnine, 1991). The alternate descriptions may be in terms of consequences or effects as an agent is considered to cause what his actions cause. Evnine notes that

This stretching out and squeezing of descriptions of events is often effected by the use of the preposition 'by'. I kill the archduke by firing the gun, and I fire the gun by moving my finger. Moving my finger, firing the gun, and killing the archduke are all the same action, though in each case more and more of the consequences of what I do are subsumed by the description. This one action, with its three different descriptions, should not be confused with two other events, the gun's firing, and the archduke's dying. These two events are the effects of my action and are distinct from and subsequent to it. The event of my killing the archduke by firing the gun is not the same as the events of the gun's firing and the archduke's dying. But if an action of mine causes those events, then I fire the gun, and I kill the archduke, and I do both of these by moving my finger (p. 42)

In summary, Davidson holds the view that when interpreting someone's actions, we
describe events in ways which connect them with the mental states we attribute to the agent. The events are the intentional actions of that agent. Since causation is a relation which holds between events, it is natural to see these actions as causally related to other, mental, events and this provides the key to explaining why some propositional attitudes are the reason why an agent performs an action, rather than merely a reason for performing it (Evine, 1991, p. 45).

In Davidson's account, actions are things that people do with purpose, not things that merely happen, so therefore not all descriptions of the event will reveal it as an action (i.e., as purposeful).

Next, I will discuss some ways in which Searle and Davidson differ in their approaches to the accordion effect.

2.4.3 Comparison of Searle and Davidson

The important difference between the above two accounts is that Searle contends that the way in which an action is described is not important and that what is important are the facts the descriptions describe (Searle, 1983, p. 101); this position is concerned with the Intentional presentation and the way in which conditions of satisfaction are met. Davidson holds an opposing view: being intentional is not something that qualifies an event itself, but events as described one way or another (Evine, 1991, p. 41).

While these two positions seem to be at odds with each other, I want to present a possible way to bring them together. To Searle's concern that it is there is no intrinsic connection between the linguistic representation and whether an action can be considered intentional or unintentional I would have to agree that there is no intrinsic connection. For Searle, the mental state is intrinsically intentional whether or not one has a language. So while Searle thinks dogs have beliefs (1983, p. 5), Davidson does not (1985). However, I do agree with Davidson's point that how an action is described

---

3 I realize that this issue is debated among philosophers. I do not present the above as a solution to their problem, but rather, as the way in which I understand Searle and Davidson's theories to be related. This is done so as to not make this debate too large a part of the thesis.
reveals something about it. The key to showing that the two views are not completely contradictory lies is Searle's use of the word 'intrinsic'. I do not think that Davidson's view leads us to the position that the connection is intrinsic, but rather, it leads us to think about the way language is used to talk about intentions. I think Davidson's point is that the description that is used highlights certain features of the action, namely, those that could count as the Intentional content of the action, and those that could not. The role of linguistic description then is to make public (via language) the aspects of the action the agent had as her Intentional content. Thus, the descriptions and intentions are not intrinsically connected, but rather, they are meaningfully connected by the way the description draws attention to certain aspects of the action.

Another way in which the two views differ is that Davidson does not make Searle's distinction between prior intention and intention in action. I believe it is possible to map Davidson's examples on to this structure.

There is an intention in action for the event under the first description as the agent could say to the question "what are you now doing?" (Searle's test question), "I am turning on the light" while performing the action (a part of the definition of an intention in action). However, the agent would not say "I am scaring away the burglar"; she could not say this as she has no knowledge of the burglar. Therefore, while the bodily movement is constant for both descriptions, they differ in terms of their intentions.

\[ \text{prior intent} \xrightarrow{\text{causes}} \text{intention in action} \xrightarrow{\text{causes}} \text{bodily movement} \]

\[ \text{turning on the light} \]

\[ \text{scaring away the burglar} \]

---

4 It is important to note that for Davidson, intentionality is not just a way of speaking; it is not that events become actions if they are described in a particular way. Rather, it is that some ways of describing events reveal them as being actions, and therefore as being rational and having been done with intention.
2.4.4 **Intentionally Opaque Contexts** Thus, when the description of an event (or object) comes after an expressed intention by means of such words as 'meant to' or 'intended to', it cannot be substituted with a coreferential description without possibly affecting the truth value of the sentence. Therefore, while it may be true that 'Mark meant to turn on the light', it may not be the case that 'Mark meant to scare away the burglar' (because he knows nothing about the existence of the burglar) even though the two descriptions (underlined) make reference to the same set of events.

2.5 **Linguistic Opacity**

Indirect discourse, which is also referred to as *oratio obliqua* in the philosophy of language, is the use of words to report what another has said by way of an indirect quotation (McKay, 1995). It is this type of context I have been referring to with the term 'linguistic opacity'. Davidson⁵ (1968/84) provides the example, 'Scott said that Venus is an inferior planet' (p. 94). He notes that the singular term 'Venus' should be replaceable with other coreferential terms without affecting the truth value of the resulting sentence. For example, 'Scott said that the Evening Star is an inferior planet'. However, this is not always necessarily the case (as Scott may not know that Venus is the Evening Star). Therefore, in such contexts, the substitution of coreferential terms is not allowed if the speaker of the original utterance has no knowledge of the given referent under the substituted description (or name).

2.6 **Quotational Opacity**

Quotationally opaque contexts are those in which a person is directly quoted. In his paper, "Quotation", Davidson (1979/84) defines quotation as

> a device used to refer to typographical or phonetic shapes by exhibiting samples, that is, inscriptions or utterances that have those shapes. This characterization is broad and vague: broad enough to include not only written quotation marks, and spoken phrases like 'and I quote', but also

---

⁵ Davidson’s interest in such contexts was focused on formulating an adequate account of the logical forms of such sentences, but as that issue is not directly relevant, I will not address it in this thesis.
the finger-dance quotes [that are] often used... In quotation not only does language turn on itself, but it does word by word and expression by expression ... (p. 79)

He then goes on to quote Quine (1940) who had said that quotation

... must be regarded as a single word or sign, whose parts count for no more than serifs or syllables. A quotation is not a description, but a hieroglyph; it designates its object not only by describing it in terms of other objects; but by picturing it. The meaning of the whole does not depend upon the meanings of the constituent words (Davidson, 1979/84, p. 80)

In other words then, nothing can be substituted for what falls between the quotation marks, in contexts with direct quotations. Quite simply, it is false to report someone as having said something they did not in fact say. In such contexts, there can be no substitution without changing the truth of the whole sentence. For example, take 'Mark said, "I turned on the light"'. Though it is true that the event referred to by 'turning on the light' is the same as that referred to by 'flicking the switch' and by 'scaring away the raccoon', we cannot substitute these coreferential descriptions and accurately say that 'Mark said, "I flicked the switch"' or 'Mark said, "I scared away the raccoon"'. Further, we could not even substitute any other portion of the utterance; it would also be false to quote Mark as saying, "I scared away a raccoon" or "I turned on a light".

2.7 Difference Between Linguistic and Quotational Opacity

The quotes from Davidson and Quine above highlight both what quotationally opaque contexts share with the other opaque contexts I have discussed, but also the ways in which they differ. The way in which all of the contexts I have mentioned are similar, is that they are all opaque contexts, and by definition, the substitution of one coreferential term for another can alter the truth value of the resulting sentence (after the 'think', 'know', 'mean to', 'say that', 'say'). However, the way in which they differ is rather significant.

In the quotationally opaque contexts, no information about the speaker's mental states (knowledge) is necessary. Any kind of substitution within the quotation marks
results in an untrue report. In other opaque contexts, not all substitutions necessarily change the truth value of the new sentence; it depends on the speaker's knowledge of the situation and therefore what he can be said to have thought, known, etc. It may turn out to be the case that one does not need to explicitly understand mental verbs, or be aware of the knowledge another has of an object/action, in order to perform successfully on linguistic and quotational opacity tasks. However, such knowledge may make such tasks much easier. If one is aware that a person knows of an object under only one description of it, it becomes easy to reject a (mis)quote (even if it is indirect) of that person's speech in which a different description/name has been used for that object. It could not have been the case that someone had used a description that she/he was not even aware of. So, to review, while being able to metarepresent may not be theoretically necessary for successful performance on linguistic and quotationally opaque contexts, it may nonetheless assist one's performance on such tasks.
Chapter 3: Cognitive Requirements

3.1 Referential Opacity

I suggest that the ability to deal appropriately with referentially opaque contexts rests upon a number of other inter-related abilities, namely: (1) the ability to represent another's knowledge and beliefs, in other words, the ability to deal with representations, as measured by passing false belief tasks (i.e., metarepresentational ability); (2) the ability to recognize that a person may represent an object or event under one description without necessarily representing it under some alternative description (i.e., the ability to deal with another's partial knowledge); and (3) the ability to treat a referring statement as an object of thought, and therefore be capable of making judgments about the similarity, or difference, among various referring statements (i.e., metalinguistic awareness). I will now turn to the task of addressing each of these crucial skills and how they contribute to the child's ability to succeed on referentially opaque contexts.

3.1.1 Metarepresentational Ability. Some researchers have disputed the use of the false belief task as the best indication of the possession of a theory of mind (e.g., Leslie 1987; 1988; Wellman 1988; Wellman & Estes; 1986). For example, Leslie (1987, 1988) points out the similarity of pretense and false belief and Wellman (1988; Wellman & Estes; 1986) presents some evidence that children understand some aspects of thought before they acquire a theory of mind. While these are important milestones in developing an understanding of representation, as well as understanding others as
mental agents, this thesis deals specifically with the achievement children have made when they are able to pass false belief tasks and can be said to be able to deal with metarepresentation (Perner, 1988). This understanding of false belief has come to serve as the standard diagnostic criterion for attributing a *representational* ToM to children. In keeping with other such researchers, I too will treat performance on false belief tasks as being indicative of children's proficiency with metarepresentation. In saying this, I must now make clear what is meant by the term 'theory of mind' and why having such a theory can be taken as an indication of being able to metarepresent.

A theory of mind is "... the set of concepts, including belief, desire and intention that [people] can use systematically in explaining the behavior of themselves and others" (Olson, 1988, p. 419). Flavell (1988) advanced a theory about the stages leading to a theory of mind. He distinguishes between two types of knowledge a child has about the mind. The first type of knowledge that children have is of their *cognitive connections* to the world. This emerges at the age of 2 or 3 years. During this period, children are said to have only a 'primitive' theory of mind. They are bound by stimulus events and believe not so much that everyone represents the world in the way that they themselves represent it, but something more basic – that there is only one way in which the world *can* be represented. Children act as if their representations have a direct correspondence (connection) to the world. They also behave as if everyone had the same representation of an object/action (which happens to be the same as the child's).

When children realize that there exists an independent representation of the world in other people's minds, one that may or may not agree with the concrete world, they are said to have reached the level of *mental representation* (Flavell's second type); an 'advanced' theory of mind. Children at this point realize that while there is only one world 'out there', people can have different representations of it. This is the realization
that the world can be represented in different ways.

Perner's level of *metarepresentation* (1988; Perner, 1991) is very similar to Flavell's (1988) level of mental representation. A child is said to have reached this level when she has the "... ability to model the semantic relationship between a model and what it models." (1988, p. 150). This is the child's realization that what she has in her mind is a possible representation of the world which has a particular relation to the world, and which therefore may be true or false of it. Once she understands this, she will soon (or simultaneously) realize that other people also have representations, which are not necessarily in a one-to-one correspondence with the world. Comparing one's own representations to those of another becomes possible at this metarepresentational level.

By the end of their fifth year, most children have reached the stage of mental representation, or metarepresentation, and are said to be aware that people act not necessarily according to the actual state of the world, but according to their beliefs about it, or representations of it. It is at this time that children can pass standard false belief tasks such as a task that requires them to predict another's behaviour based on the other's (mis)representation of the situation. For example, in one story, a character hides her chocolate in a cupboard, and then goes outside to play. While she is outside playing, her mother finds her chocolate, uses some of it, and puts the rest into the refrigerator. The character returns and wants her chocolate. The participating child is asked where the little girl will look for her chocolate. The correct answer for the question is that the character will look for the chocolate in the cupboard (which is where she put it, and therefore where she represents it to be). To pass, children must realize that the story character will act on her belief, which in this case is inaccurate, rather than on the true state of the world. This understanding allows the drastic change from a child's initial understanding of herself and others; the child, we say, now has a theory of mind. The ability to deal appropriately with referentially opaque contexts requires this
Recall that in a referentially opaque context, someone is said to have a mental attitude (e.g., one of knowing or belief) toward an object/action in the world. In such contexts, one cannot simply substitute one coreferential term with another and guarantee that the resulting sentence will be true (given that the first sentence was true, of course). Substitution is in these contexts is constrained by the knowledge of the 'believer', or 'knower'. Remember the earlier example: 'John thinks that Venus is the Morning Star' is true, but it is not necessarily true that 'John thinks that Venus is the Evening Star' as John may not know the term 'Morning Star' is coreferential with 'Evening Star'. However, when not using mental or communicative verbs (such as verbs referring to physical actions or spatial relations), substitution is permissible regardless of what is known. For example, if 'John stood next to a pile of stones' is true, it is also true that 'John stood next to a memorial of a fallen warrior' even if he does not know that the pile of stones is a memorial. In this latter type of context, we are only concerned with what happened, and not at all interested in how someone represented what happened.

A child who is unable to think metarepresentationally will treat the above contexts as equivalent and allow substitution into both of them; she is paying no attention to another's representation. In order to succeed on a referentially opaque task, the child needs to realize that one's knowledge of a referent is under a particular representation of it, and therefore, the representations of a particular referent must be kept distinct from one another (e.g., 'the Morning Star' and 'the Evening Star', not the 'the Morning/Evening Star'). One cannot substitute coreferential terms in the context of a mental attitude such as 'believe' or 'know'. Each description recruits a specific mental representation. Hence, understanding referential opacity requires an understanding of the very form of representation such that alternative representations of even the same
referent are excluded. The child must be able to distinguish a referent from another's representation of it, and compare various representations of a single referent. Hence, successful performance on opacity tasks requires the ability to metarepresent.

To summarize, in order to link false belief performance to what children must understand in coping with referentially opaque contexts, we need only look to the fact that in such contexts, people have mental attitudes towards a referent under a particular representation (description/name) and not others. The child must be able to maintain multiple, distinct representations of an object (event etc.) and recognize (at least implicitly) that the representations differ in the particular ways in which they pick out the referent. The child must also realize that a referentially opaque context expresses someone's beliefs, not about the referent directly, but about the referent as represented in a particular way. Hence, understanding opacity requires an understanding of representation. Therefore, children's understanding of referentially opaque contexts will be closely tied to a representational theory of mind (Kamawar & Olson, in press).

3.1.2 Partial Knowledge. Another skill that is critical for being able to deal with referentially opaque contexts, one that is not required for dealing with false belief, is that of recognizing that people may have only partial knowledge about a referent. In other words, the child must understand that having access to an object under one description does not ensure access to that object under all descriptions (Kamawar, 1996). It is important to recognize that this is not the same skill required by false belief tasks, which require the child to distinguish truth from falsity.

In false belief tasks, there is a contrast between reality and the child's representation of reality based on an incorrect belief (of the location of the chocolate, for example). In referentially opaque contexts, the contrast is not between an accurate and an inaccurate representation of the situation, but between two accurate, but incomplete,
representations of the referent* (e.g., 'Morning Star' and 'Evening Star', for example); therefore, such contexts require the ability to deal with partial information. False belief tasks set up the options as a 'forced choice' scenario, with the hidden object at location A or B (or in the case of an unexpected contents task, the contents being X or Y), where only one representation is correct at a time; that is, the child cannot represent the hidden object to be at location 'A & B'. However, opacity tasks do not allow the child to mark one representation as false as both representations are true, but only one is true for a particular individual. In the case of opaque contexts, multiple representations of a referent can be combined to make a composite representation, such as 'C and D', where both components succeed in referring to reality (again, the referent).

Therefore, in order to be sensitive to opaque contexts, one must understand that having access to the referent under one description does not ensure access to it under all descriptions. Additionally, the child must understand that an object is represented differently under each of its descriptions. Thus, while someone may have a (mental) attitude towards an object under one description, she may not have a (mental) attitude towards it under a different description. Therefore it is crucial that children be able to keep multiple distinctions separate so that they can distinguish the representation (description) used in a referentially opaque context from other equally true representations. Part of becoming proficient with referentially opaque contexts is becoming able to distinguish between contexts for which the specific representation is important (e.g., propositional attitudes), and those for which it is not (e.g., spatial relations such at 'under', 'next to').

3.1.3 *Metalinguistic Awareness*. Recall that the defining feature of an opaque context is that substitution of coreferential terms can change a true sentence into a false

---

*Note: These representations are embedded within a sentence context which can change from true to false when a substitution is made.*
one. In contexts that are opaque by virtue of containing a mental verb (i.e., referentially opaque contexts), the mental attitude (e.g., 'John believes..') is directed not toward the referent, but rather, toward the proposition (e.g., 'Venus is the Morning Star'), which is a linguistic entity. Also recall that referentially opaque contexts require the understanding that different descriptions may represent a referent in distinct ways, and that while someone may have a mental attitude toward an object under one description, she may not have that attitude toward it under a different one. Stated this way, it becomes clear that when dealing with referentially opaque contexts, some metalinguistic awareness is required. Metalinguistic awareness has been defined as "[the] ability to make language forms opaque, and attend to them in and for themselves ..." (Cazden, 1974; cited in Bialystok, 1993). While the wording of this definition at first glance seems to be axiomatic, it is not. The definition, as I interpret it, is not claiming that metalinguistic awareness is the ability to deal with opaque contexts; rather, it uses the word 'opaque' in a different, though related, way. The idea here is that in developing metalinguistic awareness, one must become able to 'see' the words, and not just 'see through them' to their referents; this is the sense in which the child must be able to make language forms opaque. Another related definition is offered by Tunmer and Herriman (1984), who define metalinguistic awareness as

the ability to reflect upon and manipulate the structural features of spoken language, treating language itself as an object of thought, as opposed to simply using the language system to comprehend and produce sentences. To be metalinguistically aware is to begin to appreciate that the stream of speech, beginning with the acoustic signal and ending with the speaker's intended meaning, can be looked at with the mind's eye and taken apart. (p. 12)

Put together, these definitions highlight the fact that to be metalinguistically aware, one must be able to deal with the language forms themselves, and make them objects of thought.

Obviously, for dealing successfully with opaque contexts, it is not sufficient to
have metalinguistic awareness of phonemes, syllables or words. Rather, the type of metalinguistic awareness that is necessary is that which allows the child to make propositions, as well as names/descriptions, objects of thought. In other words, one must be able to pay attention to these linguistic entities somewhat independently of their referents. One must be able to distinguish between statements that share reference, but do so by using different descriptions. For example, in the say/refer task I developed for Study 2 (which I will discuss in detail in Chapter 6), the child must be able to identify whether a third utterance (which contains a referring expression and refers to one of eight pictures) matches both, one of, or neither of the two utterances heard before it (which each also refer to one of the same eight pictures). In order to succeed on this task, the child must make the utterances the focus of her attention, and not try to match utterances according to their reference.

Metalinguistic awareness should also become evident through successful performance on the say/mean task (Apperly & Robinson, 1998; Robinson, Goelman, & Olson, 1983; discussed in detail in Chapter Six). One type of say/mean task requires the child to make judgments about an utterance containing a referring expression by stating whether the utterance provided enough information to unambiguously select a referent. A second type of say/mean task requires the child to be able to judge whether an utterance (which contains a referring expression) matches an earlier utterance (which also contained a referring expression). In order to succeed on this task, the child must both be able to recognize when the second utterance does not match the first, and be able to recall the first utterance.

To link referential opacity performance with metalinguistic awareness, we need to recognize that in opaque contexts, people have mental attitudes towards a proposition (which contains a referring expression). The child must be able to turn that proposition into an object of thought and pay attention to the particular way in which
the referent is referred to, and not just 'see through' the description to the object/action in the world.

3.1.4 Comparison of false belief and referential opacity. It is important to note that while false belief tasks are similar to referential opacity tasks, they differ in significant ways, and the former should not be thought of simply as a subset of the latter. While false belief tasks may use verbs like 'know' or 'think' (e.g., "Where does the little girl think her chocolate is?") referentially opaque contexts by definition must contain such verbs. So, while it is possible to measure false belief understanding without using mental verbs (e.g., "Where will the little girl look for her chocolate?"), the same cannot be said for measuring the ability to deal with referentially opaque contexts.

Stated another way, it can be said that while false belief tasks can be completed successfully without necessarily understanding the cognitive implications of terms such as 'know', referential opacity tasks cannot.

False belief and opacity tasks are also similar in that both involve questions concerning what a character knows, or will do, when she has a partial or incorrect representation of a situation. Yet, they differ on this point as well, as one (false belief) is easier than the other (Kamawar & Olson, in press). False belief tasks require the ability to distinguish between two quite different objects (e.g., Smarties vs. pencils) or locations (e.g., cupboard vs. refrigerator), whereas opacity tasks involve the more subtle requirement of dealing with the same object or action, but under different, but equally

---

7 I am contrasting false belief with referential (as opposed to linguistic, intentional, or quotational) opacity because questions assessing false belief knowledge occasionally use mental verbs (e.g., "Where will he look for the chocolate?" vs. "Where does he think the chocolate is?") and it is only referentially opaque contexts that are opaque by virtue of having a mental verb such as 'know' or 'think' preceding a proposition.

8 Recall that by definition, referentially opaque contexts are those which are linguistic structures that are opaque by virtue of containing propositional attitudes. Propositional attitudes are "typically expressed in language containing a psychological verb ('think', 'deny', 'doubt', etc.) followed by a that-clause" (Wagner, 1995, p. 659).
valid, descriptions. In false belief tasks, children must allow for the fact that beliefs (concerning a particular state of affairs) might be false. In referential opacity tasks, on the other hand, children have to recognize that beliefs are directed towards propositions, and not states of affairs. Therefore, even if an embedded proposition is itself true, the sentence as a whole may be false. What is at issue is not the truth or falsity of the embedded proposition, but rather, whether a particular mental attitude is held toward that proposition, and in turn, toward the referent under that particular description.

The difference between false belief and referential opacity becomes even clearer when stated in the following way: opacity is a property of certain kinds of linguistic constructs (propositional attitudes) whereas false belief is a property of certain kinds of mental states. So, while it seems apparent that false belief and opacity are related conceptually, it is also evident that they differ enough that it would be incorrect to classify false belief tasks as a type of opacity task.

3.2 Intentional Opacity

I propose that success on the intentional opacity tasks requires everything that is required to succeed on referentially opaque contexts, but it also requires an understanding of intentions. Specifically, to succeed on these contexts, the child must understand that intentions are directed toward actions/objects as they are represented (described) and that while under one description an action may have been intentional, it may not be so under another.

The expectation is that if sensitivity to referentially and intentionally opaque

---

9 What makes opacity tasks different from misleading appearance tasks (e.g., the sponge-rock) is that for the former, both of the descriptions are true descriptions of the referent; it is not the case that the referent looks like one thing but is really something else (e.g., it is both 'yellow' and 'big'). Therefore, these tasks differ from appearance/reality tasks.

10 Again, this is by definition, as the terms "'referentially opaque' and 'referentially transparent' are used primarily to classify linguistic contexts ..." (MacKay, 1995, p. 681, my italics).
contexts rests solely on understanding representation, both should develop simultaneously. However, if there is an additional understanding necessary before one can succeed on intentionally opaque contexts, we should find that referential opacity is easier than intentional opacity (assuming, of course, that the tasks used to access each are matched in all other ways).

3.3 Linguistic Opacity

The requirements for success on linguistically opaque contexts are very similar to those for success on referentially opaque contexts. The only difference may be that for linguistic opacity, the child must be better able to remember a specific utterance in addition to representing a character's knowledge state (i.e., representation). This may make such tasks more difficult than referentially opaque contexts.

3.4 Quotational Opacity

I propose that becoming sensitive to quotationally opaque contexts requires the ability to 'see' speech and not 'see through it', as is the case for the other types of opacity discussed. However, I believe this type of opaque context to be different from the three above because it does not necessarily require an understanding of mental states (such as beliefs, intentions etc.); it looks only at the linguistic form of the expression, and therefore, it should be possible, in theory at least, to succeed on this task without having a theory of mind. As mentioned earlier (in Section 2.6), having a theory of mind may not be theoretically necessary for success on quotationally opaque contexts, but it may make the task considerably easier.
Chapter 4. Related Research

While many studies have examined children's developing ability in each of the areas addressed by this thesis, only those considered to be most significant will be discussed below.

4.1 Multiple Naming

Some of the ideas advanced by this thesis have been examined in terms of Piagetian class-inclusion problems (e.g., Inhelder & Piaget, 1958). Jordan (1980) examined children's knowledge of the fact that an object can be described in more than one way at the same level of a hierarchy. The purpose of her experiment was to assess children's understanding of the permanence and conservation of kinship-kinship concepts (e.g., mother-daughter) and kinship-social concepts (e.g., mother-teacher). The method for testing kinship-kinship roles consisted of showing participants paper dolls of the same size which were ambiguous as to their sex. The dolls were divided into two groups and the participants told that one group consisted of mothers and the other of daughters. Next, a doll was moved from the mother group and placed in the daughter group. The participant was told that this mother also wanted to be a daughter, and then asked if she can be both a mother and a daughter at the same time (cf. Piaget's "are there more petunias or more flowers?"). The procedure for testing kinship-social roles was the same as outlined above. Jordan found that younger children were unable to recognize that someone could hold more than one role at a time, or in other words, be described by the two descriptions 'mother' and 'daughter' simultaneously. Jordan's results showed
that children developed this understanding between the ages of five and seven years.

This study is of interest as it offers some support that around age 5 children start to realize that: (a) there is more than one way to refer to, or name, an object in the world; and (b) these names can point out the object in different ways, highlighting various features without changing the object, only one's representation of it. One can be a daughter in one context and a mother in another.

Olson (1970; Ford & Olson, 1975) looked at children's ability to highlight different features of an object as a way of referring to it in multiple ways. In Olson & Ford's (1975) first experiment, participants were presented with an array of blocks ranging in number from two upwards. A gold star would be placed underneath one of the shapes (for example, a round, white disk). The participant's task was to convey to the listener a description of the block under which the gold star was located. For example, if the target item was presented with a black disk, the participants should refer to it as 'the white one'. If it is then presented with a white, square block, the target item should be referred to as 'the round one'. The number of dimensions of difference (for e.g., shape, colour, size) between the target item and the other blocks varied. The second experiment was very similar to the first except that it used pictures of shapes instead of blocks. This allowed for greater manipulation of dimensions (for e.g., stripes or dots). Ford and Olson (1975) found that "even 4-year-olds do not give an invariant 'label' ... of an object but rather represent that object in terms of the context of alternative objects" (pg. 371) thus showing that young children are capable of referring to an object in different ways. However, such a finding does not reveal whether such children had any awareness of the fact that different representations could result in opaque contexts.

4.2 Referential Opacity

Being able to deal with referential opacity requires more than the simple ability to use multiple descriptions of a single referent; it requires understanding the
implications of substituting an alternative name or description of a referent for a subject's mental representation. In an earlier study (Kamawar & Olson, in press), one hundred twenty 3- to 7-year-olds were given three referential opacity and three false belief tasks. The opacity tasks were short stories about two characters (Mark and Anna) followed by questions about what the characters did and thought. Each opacity story included a referent that could be represented in two distinct ways. One character had knowledge of the referent under only one representation (description), and it was made clear that she/he lacked knowledge about it under other representation(s) (descriptions). For example, in one story, the two characters went a doctor's office so that Anna could get treatment for her (slightly) hurt knee. Mark stood next to the doctor (description A) while the doctor put a bandage on Anna. Unbeknownst to Mark, the doctor was in fact Anna's mom (description B). At the end of the story, the following key questions were asked: "Did Mark stand next to Anna's mom?" (transparent context) and "Does Mark know that he stood next to Anna's mom?" (referentially opaque context). The correct answer for the first was 'yes', and for the second, 'no'. The first question (transparent) ensured that participants knew that it was permissible to substitute one description for another in some situations. The second question (referentially opaque) was designed to determine whether participants were sensitive to the fact that substitution was not allowed because Mark's mental attitude of 'knowing' was toward the proposition that represented the referent as 'the doctor' and not as 'Anna's mom'.

Children earned a score of up to three points (one per story) for answering the referential opacity and transparent questions correctly. The also earned a score out of three based on their performance on the three false belief tasks. They were also given a score of one for each story for answering both the transparent and opaque questions correctly (OPAC score). These too were summed across the stories to give scores out of
Consistent with other studies, it was found that children improved on false belief tasks with age, with a sharp increase in performance around age four years. Children’s performance on the transparent questions improved with age, but only slightly since even the youngest children were scoring an average of 2.6/3 on the summed score. Children also improved with age on the opacity questions, with the biggest jump in performance taking place from age 4 to 5. The measure of interest was the OPAC score. This score, out of 3, was sensitive to children’s ability to distinguish between the transparent and opaque questions. Children’s OPAC scores improved significantly with age, with the biggest jump in performance from age 4 to 5.

The results demonstrated that even with age removed as a factor, opacity (as measured by the OPAC score) and false belief performance were significantly correlated ($r (117) = .25$, two-tailed, $p < .01$), supporting the view that the two are related (though not equivalent). Further, children found opacity (OPAC) more difficult than the false belief tasks ($t (119) = -5.23$, $p < .001$). The mean score on OPAC was 1.73 (out of 3) and on false belief tasks was 2.21 (out of 3), though both tasks are metarepresentational.

The results were interpreted as indicating that success on opacity tasks was related to metarepresentational ability (as evidenced by their performance on false belief tasks). The results were taken also to support the view that children come to realize that language does not refer to the world directly but does so via one’s representations of it. These results are of interest in that they demonstrate a clear relation between success on referentially opaque contexts and metarepresentational ability, but offer no explanation of why opacity tasks are more difficult.

Another study of interest is that by de Villiers (1994) in which she examined a grammatical aspect of opaque contexts. She pointed out that one property of mental
verbs (e.g., 'know', 'believe') is that they occupy a special grammatical form that is called complementation. A sentence complement is "a sentence-like construction that is embedded as a sister of [the verb]" (O'Grady & Dobrovolsky, 1987, p. 443). It can also be described as "a phrase that appears together with a verb, completing its meaning: ... *I thought HE WAS DEAD.*" (Pinker, 1994, p. 475). Here, the phrase 'he was dead' is the complement of 'thought'. De Villiers investigated complementation as a way of examining the interplay between language and cognition by examining children's ability to deal with sentences that contain clauses that appear the same on the surface, but reveal a difference in grammatical structure when rephrased as a wh-question. For example:

1) "Last night he **called** to organize the party on Saturday"

2) "Last night he **wanted** to organize the party on Saturday"

become:

3) "When did he **call** to organize the party?"

4) "When did he **want** to organize the party?"

The answer to the first question is "last night". The answer to the second question is not so clear. It could be either: (1) he wanted to organize the party 'last night'; or (2) 'Saturday', when the party is to take place (de Villiers, 1994, p. 2).

The difference between these sentences results from the fact that verbs like 'call' take an adjunct (the clause is not selected by the verb), whereas verbs like 'want' take a complement (the clause is selected by the verb). That is, verbs of communicative or mental acts take complements whereas the others take adjuncts. It is these verbs of saying and thinking which link grammatical development to theory of mind.

Russell (1987) makes a similar distinction between verbs that are clearly extensional (transparent) and those that are at least somewhat intensional (opaque) with communication and mental verbs falling into the second category. Examples of
these classes of verbs (or phrases) include: the transparent verb phrase 'stand beside'; the opaque verb phrases 'think that', 'believe that' and 'hope that'; and the communicative phrase 'talking to', which is somewhat opaque (Russell, 1987, p. 293).

Though Russell places communicative verbs in with mental verbs, it might be better to construct a new category for such verbs because some verbs of communication can be opaque if followed by a 'that clause' (e.g., 'say that'), while the same verb might not be opaque in other grammatical contexts (e.g., 'say')1. It is certainly the case that some communication verbs do not take complements (e.g., 'call').

Russell (1987) tested 5- to 7-year-olds to see if they knew that: (1) in an opaque context, the description of an object must be constrained by the knowledge held by the person designated as subject of the sentence; and (2) in a transparent context, there is no such constraint. For example, children were told stories in which a character had only partial knowledge about something that had happened. In one story, a character (George) knows that a thief has stolen his watch, and it is made clear that George does not know anything about the thief, such as the fact that the thief has curly red hair. The participant was then asked questions about what George thinks, such as "Can we say that George was thinking 'I must find the man with curly red hair who stole my watch'". The correct answer of course is 'no'. George does not know that the thief had curly red hair so there is no way he could be said to be thinking about him as the man with curly red hair.

Russell (1987) found that children were better at transparent contexts in general and that the older children performed much better than the younger ones (errors decreasing from around 87.5% to 50%). However, the sentence constructions that were used were fairly complicated and may exaggerate the age at which children can handle

---

1 Perhaps a good name for this category would be 'translucent'. 
them. The questions could be reduced somewhat and still be opaque contexts. Therefore, the key questions used in the tasks designed for the thesis being presented have far less syntactic embedding. Also, the results of Russell's experiment were not examined for their relation to the development of a theory of mind, metalinguistic awareness, general language ability, short term memory, or other kinds of opaque contexts as is done in this thesis.

De Villiers and Fitneva (1996) tested 31 children aged 3 to 6 on both false belief tasks and referential substitution tasks (i.e., opacity tasks). The standard false belief tasks of change in location and unexpected contents were used. Their opacity tasks involved an object that was referred to in three ways: based on its appearance (a silver box), based on its contents (candy), and based on its function (it is a birthday gift). Children were asked about story characters and whether those characters did something to, or thought about, objects under different descriptions. An example of a question with a transparent context from her study is "Did the mom put the silver box/birthday present on the top shelf?" An example of a question with an opaque context is "Did the mom know the silver box/candy was on the table?" (p. 13). De Villiers and Fitneva reported that only the children who passed false belief tasks went on to appropriately constrain referential substitution. De Villiers (1994) advances the view that language development, specifically, the acquisition of complementation, is involved in changes in the child's cognitive development (i.e., the child's grasp of other minds). However, the data (de Villiers & Fitneva, 1996) are equivocal on that point in that many children passed theory of mind (ToM) who continued to fail opacity. If anything, the data may say the opposite; understanding beliefs is a condition for understanding opaque contexts. De Villiers (1994) also maintains that there is a relation between ToM and referential substitution performance. However, when age effects were partialled out (which is necessary since both skills are age related), the results lost
their significance and therefore are not as robust as other findings (Kamawar & Olson, in press) which showed that there was significant relation between ToM and referential substitution, with age partialled out.

Apperly and Robinson (1998) examined the relation between sensitivity to opaque contexts (or, intensional contexts as they refer to them) and theory of mind performance. In the first of two studies reported, 4- to 6-year-olds were administered a standard theory of mind task (the misleading appearance task) and a referential opacity task. The opacity task involved a puppet that knew that X was in the box (e.g., a die) but didn't know that the X was a Y (e.g., the die was an eraser). The participants were to judge whether the puppet knew that the X was a Y (Q1) and whether he knew that there was a Y in the box (Q2). The correct answer for both questions was 'no'. The authors found that children had no problems handling transparent contexts and could say that the puppet could see Y. They found that those children who were able to correctly answer the false belief question were also able to correctly answer Q1 above, but couldn't handle Q2 above (the opaque question).

In order to test the hypothesis that the reason children found Q2 difficult was because of a "lack of knowledge about the pragmatic substitution rules in talk about mental states" (p. 298), they ran a second study. In this study, set up much like the first, participants were to predict where the puppet would look for a Y. Again, the authors found that 5- and 6-year-olds could answer both the false belief question and Q1 above correctly, but still tended to fail the opaque question.

Based on these results, Apperly and Robinson advance the view that in order to deal with reference of multiple perspectives, one must be able to partition the aspects under which the object is represented; that is, one must distinguish among the object's different representations (what I have referred to as partial knowledge/representation). They conclude that their finding that children can pass the false belief test and still not
pass the opacity test is incompatible with the view that theory of mind tests are to be taken as diagnostic of the ability to metarepresent. They argue that if theory of mind tests did in fact measure metarepresentation, then those children who passed ToM tests should also be able to deal with intensional contexts.

I agree that the ability to handle partial representation is more difficult than what is required by a standard false belief task, but I don't agree that false belief tasks are not metarepresentational. Their findings were contrary to Kamawar & Olson (in press) who found that performance on opacity and false belief tasks is related (even with age partialled out) when children were given three of each type of task. I believe that the reason Apperly & Robinson (1998) were not able to discover a relation between false belief and opacity may be that the children in their study were given only one of each type of question, thereby making the study relatively insensitive to individual differences. Kamawar & Olson (in press) proposed that the relation between the tasks is mediated by the development of the child's ability go beyond thinking about reality to thinking about representations; that is, by the development of metarepresentational ability.

4.3 Intentional Opacity

While children of around 4 years of age are able to take intentions into account when judging or describing the behaviour of others (e.g., Astington, 1991; Wimmer, Gruber & Perner, 1984), it is not known how this understanding affects their sensitivity to contexts where the agent's intentions render the context opaque. For example, a child who knows that someone performed an action intentionally (e.g., throwing away a scrap of paper) may believe that the person performed that action intentionally under all descriptions of it (e.g., tossing out a picture drawn by the child).

While the notion of 'intentional opacity' has not been addressed directly, there have been other studies that are relevant to children's understanding of the notion of
intentions, and their ability to distinguish intentional from unintentional behaviour. In fact, most of the studies examining children's understanding of intention do so within the context of children's moral judgments; that is, whether or not children take an agent's intention into account when they make a moral judgment about the agent's actions. Many of these studies propose that young children are not able to use intention cues when evaluating the behaviour of others (Karniol, 1978). However, the claim has been made that there is a problem with those studies based on Piaget's original paradigm since, as Karniol (1978, p. 77) points out, "the acts of damage that are described [for all characters] in the stories are not intended but are the result of carelessness". If this confound was controlled for, Karniol proposed, a different pattern may appear.

Wimmer, Gruber & Perner (1984) examined children's use of the verb 'lying'. This study was conducted to investigate children's moral judgments by telling children stories about different characters who had differing intentions; the participant's task was to lay blame when appropriate. What they found was that until age 7, children base their decisions about whether or not a statement is a lie on observable behaviour and do not use information about the intentions of the agent. They call an untrue statement a lie even if the speaker thought it was true and conversely, they don't call a true statement a lie even when the speaker meant to deceive. This is not to say that children do not understand intention at all because even as young as 4 or 5 years of age they are able to take it into account in their moral judgments, thus demonstrating a concept of intention (e.g., Wimmer et al., 1984; Zelazo, Helwig and Lau, 1996).

Lee (1995) examined, among other things, children's ability to attribute intention using verb forms that referred to either intention in action or to prior intention using a series of illustrated story tasks. An example of one kind of task she used is her fourth task, that looking at inferred intention. In this task, children heard a simple story of two
boys, both of whom are standing near a swimming pool. They are told that one boy 'slips and falls in the water' while the other one 'jumps' into the water. They are then asked 'which boy wanted to get wet?' (p. 36). Here, the participant's job was to infer the intention of the characters in the story and then map the implicitly encoded intention in the verb 'wanted' on to the correct character (the boy who jumped in the pool). Lee found a significant age effect with 81% of the 5-year-olds answering correctly. She also found that for 4-year-olds there was a significant correlation with performance on false belief stories even when the age was controlled. The results for the study as a whole, however, did not find a significant relation between false belief performance and performance on such intention tasks.

This study is directly relevant because it examines children's ability to distinguish between behaviour that shows evidence of prior intention and behaviour that does not. It supports the model being proposed because at least one age group showed a relation between performance on false belief (theory of mind) measures and measures of mapping intentional language on to actions. Such tasks do not measure intentional opacity directly, however.

Another study of interest was conducted by Astington (1986). She examined children's understanding of verbal expressions of intention by having them select from a set of 3 pictures the one that best exemplified the statement uttered by the experimenter (which included an intentional term such as 'will', 'is going to' and 'intended to'). For example, the child would hear "the boy is going to have a swing" and then see three illustrations, one each of: a boy holding a balloon, a boy running toward a swing set, and of a boy swinging on a swing set. The correct solution required the child to choose the picture showing a boy who is going to have a swing (it is in his future). Her main finding was that children's understanding of verbal expressions of intention improves greatly between 5 and 9 years of age.
Astington's (1986) study is relevant in that it examines children's proficiency at matching actions to linguistic descriptions. It is particularly pertinent as it highlights the fact that children's ability to differentiate the intention to act from the act itself develops in the primary school years, the very age group that is being investigated for the purpose of this thesis. While some of the constructions used in this study resulted in opaque contexts (e.g., "Which boy intended to ..."), this study did not examine children's sensitivity to intentional opacity.

Astington (1991) later ran a simpler version of the task outlined above. Instead of choosing from a set of 3 pictures, children were shown only two pictures, for example, one of a boy swinging and one of the same boy running toward the swings. Further, the specific terms used to question the story characters' intentions were also simplified, so instead of using term such as 'will', 'is going to' and 'intended to' (from her 1986 study), she used the terms 'gonna', 'think (s)he'll', 'wants to' and 'would like to'. For example, children were asked, 'which boy thinks he'll swing?'. As a result of these changes, including using less sophisticated expressions of intention, Astington (1991) found that the majority (61%) of the 5-year-olds were able to get all four intention questions correct. As mentioned in reference to Astington (1986), while some constructions may have resulted in opaque contexts, children's sensitivity to intentional opacity was not examined directly in this study.

Kamawar and Astington (April, 1999) found further evidence to support the view that children are able to ascribe intentions to another person before they are able to do so using more sophisticated language; that is, language in which intention is implicitly encoded in the terms used. This was accomplished by comparing children's ability to attribute intention when it is explicitly or implicitly encoded in language. Fifty-nine 4- to 9-year-olds heard three stories, each using a verb that implicitly encoded intention ('stealing', 'lying' and 'hiding'). The key characters in each story differed in
their prior intentions but not in their physical actions. For example, while two characters took a friend's video games home, one knew that she was doing so, and meant to do it, while the other was completely oblivious to the fact that a video game had fallen (accidentally) into his backpack. Questions were asked about story facts and the characters' actions using implicitly/explicitly encoded verbs (e.g., 'Did Adam/Beth steal a game?' and 'Who took a game on purpose/by accident?'). Children were also required to make a moral judgment about each of the characters' behaviour (a five point scale, ranging from 'very bad' to 'very good', with 'not good or bad, just in between' as an option).

To be scored correct on the above questions, the child had to correctly answer both of a given type (e.g., 'Beth stole it' and 'Adam didn't'). Thus, each child received a score out of 3 for each verb type (implicit/explicit). To be considered capable of making an appropriate moral judgment, children had to rate the character that had the intention to 'lie', 'steal', or 'hide' as lower on the moral judgment scale than the unknowing character in the same story.

The researchers found that performance on the moral judgment questions was consistently good across all age groups indicating that even the youngest children understood intention in these contexts. Further, results demonstrated that children's ability to appropriately use verbs that are implicitly encoded for intentionality improves with age, and that this ability comes in after they are able to use explicitly marked verbs. When performance on the verb pairs (explicit/implicit) was contrasted, it was found that the explicit form was always easier than the implicit for each pair of verbs (e.g., 'take on purpose' was easier than 'steal'). Further, it was found that children's performance on the three explicitly marked verbs did not differ significantly, though there were some differences in performance among the implicitly marked verbs ('lie' was easier than 'steal' and 'hide', with the latter verbs not differing from each other).
These results were taken to support the view that children have the concept of intention available to them before they recognize that the meanings of some verbs require that the actor had acted intentionally. So then, it seems that what separates younger children from older children is not whether they have the concept of intention, but rather, whether they have acquired a more sophisticated vocabulary to ascribe intentions.

The results from Astington (1991; 1986) and Kamawar and Astington (1999) support the view that younger children have a concept of intention, but what separates them from older children relies (at least in part) on their understanding of the terms used to ascribe it. While this is not the same as being sensitive to intentionally opaque contexts (i.e., understanding that intentions are directed toward an action/object under a particular description and not necessarily others), it is certainly related. In order to be sensitive to intentionally opaque contexts, not only must children have the concept of intention, they must also be able to use some of the more sophisticated language to ascribe it (e.g., 'mean to'). As such, it ought to be the case that it not until children are in their later elementary school years that they will be able to deal with intentionally opaque contexts containing the sophisticated phrase 'meant to'.

Taken together, these studies show the developmental changes in children's understanding of intention, but they do so without directly assessing the understanding of opacity by means of the standard criterion of substitutability of coreferential terms in the complement clause.

4.4 Linguistic Opacity

Very little work has been done looking at young children's ability to deal with

---

2 Recall that when the description of an event (or object) comes after an expressed intention by means of such words as 'meant to' or 'intended to', it cannot be substituted with a coreferential description without possibly affecting the truth value of the sentence.

3 The two studies presented in this section would have fit equally well in the section on Quotational Opacity.
indirect quotation. Perhaps this is not surprising given that indirect quotation is more often used in writing than in speech (Chafe, 1985), and the age group of interest is still in the early stages of literacy development.

Ely and McCabe (1993) point out that "little is known about the speech that children themselves quote" (p. 671). They point out that the fact that speech can be treated as an event that can be later recalled (i.e., an object of thought), may also be affected by metalinguistic development. This view is certainly consistent with the view being presented in this thesis.

In their first study, the Ely and McCabe (1993) examined the spontaneous narratives from ninety-six 4- to 9-year-olds. Children were interviewed individually and prompted to describe events (e.g., a visit to the doctor) and to describe 'what happened'. Selected portions from each child's conversation with the researcher were analyzed for their reported speech content. Reported speech was defined as any

linguistically marked recounting of a past speech event. This generally required the use of a speech verb (e.g., to say, to tell, to ask), but also could be achieved through contextual and syntactic cues and/or changes in person or tense, such as:

(2) I kept on wanting [to] change pillows and my Mom: 'I want that pillow, here, here ... I want that pillow, Brian.'

Use of the historical present tense was included, as in:

(3) And I said: 'Hi Mom.' And she SAYS: 'Hi.' (Ely & McCabe, 1993, p. 675)

They restricted their analyses to utterances which were reported from real, not imaginary speech, and real, not imaginary, characters. The speech was further categorized as being: (1) direct speech, the type of speech "which in written text is conventionally enclosed in quotation marks" (p. 676); (2) indirect speech (i.e., indirect quotation), the type of speech "marked by the use of a speech verb and a nominal or infinitive clause ... [for example], 'He said that I had ammonia [presumably pneumonia]'" (p. 676); and (3) narrativized speech, which represents a description, or summary, of
speech that had taken place.

The findings of interest were that the use of direct quotation increased with age (though more so in girls than in boys) and that it was used more than the other types (overall) by the 4- to 7-year-olds. However, by the time children were ages 8 and 9 years of age, their use of indirect speech rose to the level of their use of direct speech.⁴ Taken together, these results led Ely and McCabe to conclude that children tend to use direct quotation far more often than indirect quotation.

Hickmann (1993) also reports similar findings from her study examining children's reported speech in narrative discourse. In her study, 4-, 7- and 10-year-olds, as well as some adults, participated in a game in which the task was to tell a 'naïve interlocutor' about what had happened in a particular scenario. The scenarios differed as to whether they were presented as films or as texts (which were read to the children), and whether the children had to narrate the scenarios or act them out (using puppets) for the interlocutor. The scenarios consisted of two characters (anthropomorphized animals) discussing a problem and its solution. For example, in one film version, a giraffe and a donkey were shown having a discussion about a missing penny and what might have happened to it. The text version of the same scenario contained exact reports of the original conversation between the two characters using direct quotations (e.g., 'The donkey said, "Maybe you're right. Friends don't steal. I'm sorry I was mad at you. Now let's go and play."').

The participants' narratives were coded as to whether they contained: (1) unframed direct quotations; (2) framed direct quotations; (3) framed indirect quotations; (4) nonexplicit descriptions; and (5) local mixtures. Unframed direct quotations were those that were not marked in any way as being reported speech, but were direct quotes

⁴ Note that this study compares the relative use of the two ways of reporting speech, not their absolute usage.
nonetheless (e.g., when re-enacting a scenario, a child just has the puppets speak to each other: "You look sad"). Framed direct quotations were those that were explicitly marked as being exactly what a character had said (e.g., 'The raccoon said to the dog, "You look sad"'). Framed indirect quotations were, as the names suggests, indirect quotations that were explicitly marked as being reported speech (e.g., 'The raccoon said that he looked sad'). Nonexplicit descriptions presented information as facts and/or internal states instead of referring to them as speech events (e.g., 'The raccoon thought that the dog looked sad' or 'The dog looked sad'). Finally, local mixtures were those cases that did not fall into any of the other groups; they contained aspects of more than one of the other types (e.g., 'The raccoon said to the dog that he looks sad today'). The above categories were seen as falling into two main types: those that contained explicit metalinguistic boundaries between reporting and reported message and those that did not, with only framed direct and indirect quotations falling into the first type.

Each participant's narrative was further classified according to "the degree to which, and the systematicity with which, the presentation of speech relied upon one or another type" (pp. 71-2). Four kinds of presentation modes were identified: "a re-enacting mode, corresponding to the used of unframed quotations ... ; two reporting modes, corresponding to the use of framed quotations, one variant of which consisted of direct reporting ..., the other of indirect reporting...; [and] a descriptive mode, corresponding to the use of various types of nonexplicit descriptions" (p. 72).

The results of interest from this study were that overall, the youngest children (4-year-olds) tended to use only the re-enactment or descriptive modes, very rarely using framed quotations. The older children (7- and 10-year-olds) showed a strong preference for the direct reporting mode (direct quotes), while the adults used the direct and indirect modes with equal frequency. So, just like Ely and McCabe's (1993) findings, Hickmann's results favour the view that children can deal with direct quotes before
they deal with indirect quotes.

Taken together, these findings, though not about linguistic opacity specifically, are still very interesting. They show that children are more comfortable, and have more experience, with direct quotation than with indirect quotation. Based on these studies, I expect that children will perform better on quotationally opaque contexts (direct quotes) than on linguistically opaque contexts (indirect quotes).

4.5 Quotational Opacity

Torrance, Lee and Olson (1992, April) examined children’s awareness of the distinction between ‘the exact words’ and a paraphrase in two forms of language (narratives and nursery rhymes). In their study, 119 children aged 3 years to 7 years 11 months were asked to judge the correctness of utterances (made by a puppet) which were meant to be either good paraphrases (meaning maintained) or verbatim repetitions of target utterances. After hearing the target utterance, the puppet (a bear) was told to repeat what he had heard either ‘using the same words’ or ‘not using the same words’. The participating children had the job of rewarding the bear with a sticker when he was correct and admonishing the bear when he was wrong by saying ‘no sticker Teddy’. The expectation was that children would perform better when the target utterances were drawn from nursery rhymes, in which the exact wording is very important.

This study found that children were quite competent in distinguishing true from false paraphrases and even the youngest children could recognize identical wording (more so with nursery rhymes). For the narrative condition, children had difficulty rejecting true paraphrases when instructed to accept only verbatim repetitions. For the nursery rhyme condition, children had difficulty accepting true paraphrases when told to attend to meaning. The concept of wording, as distinct from meaning, does not appear to be completely developed until children are 6 to 7 years old. The authors suggest that
this type of metalinguistic awareness arises primarily through exposure to the written word, but once acquired, is applicable to speech.

This study is pertinent to the one at hand as it demonstrates that children, by the age of 7, are able to determine when a repetition is a word for word copy of a target utterance. It provides a baseline of sorts with which to compare the thesis findings regarding the effects of adding the context of "Did Mark say ...?" (a quotationally opaque context) in which the substitution of an alternate description of an object/action can result in an untrue sentence. Thus, it would seem likely that children will not be able to deal appropriately with quotationally opaque contexts until they are at least 7 years of age.

4.6 Metalinguistic Awareness

Robinson, Goelman, and Olson (1983) conducted a series of studies examining children's metalinguistic ability, specifically, their ability to understand the relation between what was said and what was meant. In order to succeed on their tasks, children had to be able to distinguish what the speaker meant from what his or her utterance meant.

In one of their say/mean experiments, 19 Kindergarten children (with a mean age of 5 years 6 months) were required to distinguish what was said from what was meant in their own utterances, as well as those produced by the experimenter. In this study, the child and the experimenter sat on opposite sides of an opaque screen. Each had the same set of cards in front of them that varied along a couple of dimensions (e.g., large/small, red/blue flowers). The child and experimenter would take turns choosing a card from his/her own set and describe the card so that the other player could choose the identical card from his/her own array.

The utterances produced by the experimenter would occasionally be ambiguous. For example, the child might be told to pick up 'the red flower', an expression which
described both 'the big red flower' and 'the small red flower', and communication would fail (i.e., there was a mismatch between the experimenter's and child's cards). When that happened, the child would be asked to make a judgment about what the experimenter had said. The three types of questions all began with "Did I say … " and were followed with one of three types of endings. The first type was a disambiguated version of the original utterance; namely, "Did I say 'the big red flower'?". The second type contained a verbatim copy of the original utterance; namely, "Did I say 'the red flower'?". The final type contained an incorrect version of the original message; namely, "Did I say 'the blue flower'?".

The children were also given the chance to judge what they had said. This was done by having them describe cards for the experimenter to pick up, and when they gave an ambiguous message which resulted in a communication failure, the experimenter would ask them questions of the types described above with 'you' replacing the 'I' (e.g., "Did you say 'the blue flower'?").

Robinson et al. (1983) found that the children (5-year-olds) were very good at detecting incorrect paraphrases and rejecting them as what had been said (they correctly rejected them 81% of the time). Children were also quite good at correctly accepting a verbatim repetition of what the experimenter (or they themselves) had said even though it had resulted in a mismatch (they correctly accepted them 76% of the time). When it came to judging the disambiguated version of the original message, children were not very good. If the disambiguated version of the message was consistent with the intended card, only 40% of the 5-year-olds were able to correctly reject it. Note that phrased another way, 5-year-olds accepted the disambiguated version as what had been said 60% of the time, which is not too different from their acceptance rate of the verbatim version.

The results were interpreted to mean that 5-year-olds have no difficulty rejecting
messages as inadequate if they express an intention that is different from the speaker's (e.g., an incorrect version). However, the authors found that 5-year-olds do have a great deal of trouble distinguishing between two correct, though different, expressions of the same intention (i.e., two expressions referring to the same card). Robinson et al. point out that what children seem to do is select both what was actually said (i.e., verbatim reports), and what should have been said (i.e., disambiguated version), as what was said. The authors take this to support the view that to the child, "the two utterances are indiscriminable ... because they do not, as far as he [or she] can see, express different intentions" (p. 81).

Another way of stating what is happening is that children act as if the two utterances are indiscriminable because they do not recognize their different ways of referring. Children cannot 'see' the expressions in and of themselves; they are only able to 'see through' them to the intended card (referent). This is reminiscent of the definition for metalinguistic ability we saw earlier, which defined it as "[the] ability to make language forms opaque, and attend to them in and for themselves ..." (Cazden, 1974; cited in Bialystok, 1993). So for many of the children in this study, any expression that refers to a given card is equivalent to any other expression with same referent. Stated this way, it becomes clear that the specific type of metalinguistic ability tested for in the say/mean tasks, that of being able to distinguish between what was said from what was referred to, is crucial for dealing with opaque contexts.

In another experiment, Robinson et al. (1983) extended their earlier findings by examining the relation between children's ability to make judgments about what was said (as above) and their ability to make judgments about the adequacy of the message. Thirty-nine 5- to 6- year-olds took part in this next experiment. The ability to make judgments about what was said was tested by having children play the communication game outlined above, but this time, the experimenters required somewhat more on the
part of the child. Not only did the child have to be able to reject incorrect versions, accept verbatim versions, and reject disambiguated versions of the original statement, they also had to be able to provide information as to what was said when they rejected the incorrect and disambiguated versions. For example, if the experimenter said "the big flower" and there was communication failure, when the child was asked by the experimenter "Did I say 'big, red flower'?", the child had to be able to say "no" and then when asked "What did I say?", the child had to be able to say "the big flower". In order to be considered as successful at making judgments about what was said, a child had to be correct on every occasion.

The ability to make adequacy judgments was tested by having children play the same sort of game with an experimenter (i.e., opaque screen, set of cards, etc.), but this time, the questions required the children to make judgments about whether what the experimenter said provided enough information to choose the correct card (the one the experimenter had). In other words, children had to judge whether the expressions used were ambiguous or not. For example, on two of the experimenter's turns to choose the card, she would give an ambiguous message that referred to two cards. After the child chose one of the cards, the experimenter revealed that she had been talking about the other. The child was then asked "Did I tell you enough about my card? Could I have helped you any more?" (p. 82). If the child correctly answered 'no' to the first part, or 'yes' to the second, she was then asked, "What should I have said?". In order to be scored as correct, the child then had to provide the unambiguous description of what should have been said as the answer to the second question. In order to be scored successful on making judgments of message adequacy, a child had to be correct every occasion.

The researchers found that children's ability to make judgments about what was said and about message adequacy did not vary according to whether the message had
been given by the child or the experimenter. As with the previous experiment, children were very good at rejecting the inappropriate versions of the messages. Further, they found that there was a very strong relation between the two kinds of judgments. Of the 27 children who were successful on judging message adequacy, 21 were also successful on judging what was said. Of the 12 who were unsuccessful at making adequacy judgments, 11 were unsuccessful at making judgments about what was said. The authors interpreted this finding to mean that once the younger children knew what the intended referent was, they were not able to hold it in mind separately from the description of if (even when the description was ambiguous).

As with their previous study, Robinson et al. (1983) found that children tended to conflate what was meant with what was said, and therefore only notice that an expression was not the one said when it was in conflict with the card chosen (i.e., 'the blue flower'). They conclude that "children fail to discriminate between two sentences which, as far as they can tell, are both expressions of the same underlying intention" (p. 83). This latter experiment supported the findings of the first, and went further by showing that children's ability to make judgments about what was said was highly related to their ability to evaluate the adequacy of judgments to refer unambiguously.

Again, like the previous study, these results are directly relevant to the thesis at hand in that they demonstrate a way to test for the specific type of metalinguistic ability that I propose is required to deal successfully with referentially opaque contexts. Specifically, say/mean tasks provide one way to test that a child is able "to distinguish between statements that share reference, but do so by using different descriptions" (Section 3.1.3).

More recently, and even more directly related, Robinson, along with Apperly (Apperly & Robinson, 1999; Robinson & Apperly, April, 1999), examined the relation between children's ability to make adequacy judgments and their performance on
referentially opaque contexts. Their reason for examining this relation was their interest in what they see as a theoretical link between the two tasks. They note that an utterance is ambiguous when the referring expression, that is necessarily only a partial representation, is insufficiently narrow in its scope to identify a single intended referent. In such a case, appreciation of ambiguity in the message requires that the message contents be held under their particular terms of description; held not just as referring but as doing so in a particular way. As with intensional [referentially opaque] contexts this requires children to represent the partial relation between the mental or linguistic content and the referent” (p. 8).

Consequently, they expect there to be a relation between performance on the two tasks with the relation being mediated by children's ability to deal with partial referential relations.

Ninety-seven 4- to 6-year-olds participated in their study. Each child took part in a communication game similar to that used in Robinson et al. (1983) and in a partial knowledge task (referential opacity task). The communication game required children to make judgments about message adequacy (as described in Robinson et al., 1983). For example, in this study, there was a set of cards consisting of two red balloons (one round and the other long). The experimenter would say "Heinz says he has chosen the red balloon". After the child held up one of her cards, the puppet held up the other one and the child was asked, "Did Heinz tell us enough?" If the child correctly answered 'no', she would be asked "What should he have told us?" The correct answer would have to contain a disambiguated description of Heinz's card. Children would have to answer both questions correctly to be considered passing. Note that this is slightly different than Robinson et al. (1983); only the experimenter (through a puppet) provided the messages regarding what was to be picked up.

The other type of task that children were given was the partial knowledge task. In a typical partial knowledge task, children and a puppet ("Heinz") would be shown an

---

5 Only the directly relevant portions of this study are being presented.
object that they had to identify, such as a ball. They were then told that the ball was going to be a present for Heinz, but Heinz hasn't been told about it yet. Thus, there are two descriptions (representations) available for a single referent (namely, 'ball' and 'present'). The children were then asked a partial knowledge question "Does Heinz know there's a present in the box?" (I would call this a referentially opaque question). The correct answer is 'no'.

The finding of interest was that even with age partialled out, there was a significant correlation between performance on the message evaluation questions and the hard partial knowledge question (opaque context). Apperly and Robinson's explanation is couched in terms of children's ability to deal with the partial nature of representations (versus being able to differentiate spoken words from intended meaning), and not in terms of children's metalinguistic ability (as is done in this thesis). Partial knowledge designates a representation that is true but incomplete. They argue that children's difficulty on the partial knowledge question and the communication game (adequacy judgment task) arises from the partial nature of the representations involved in both. They draw the conclusion that difficulty with ambiguous utterances (i.e., the message evaluation task) is not a result of the inability to consider the meaning of an utterance as being separate from the speaker's intended meaning (as was done in Robinson et al., 1983). Rather, they argue that what children have trouble doing is constructing a distinct 'mental space' that specifies the partial content of the belief (as in opacity tasks) or the utterance (as in the message evaluation portion of the say/mean task).

Apperly and Robinson's (1999; Robinson & Apperly, April, 1999) conclusion regarding what is required for successful performance on the intensional (opacity) task is consistent with what I have described (in Section 3.1.2) as the criterion of being able to deal with partial knowledge. However, I would characterize the ability to deal
appropriately with the message evaluation task in terms of metalinguistic ability (as defined earlier). The authors do allude to the metalinguistic nature of the latter task when discussing the crucial insight behind being able to make a distinction between message meaning and intended meaning. They note that "if children do not represent a linguistic expression under its particular terms they lack the capacity to reflect on the aspects of meaning arising therefrom" (Apperly & Robinson, 1999, p. 18). However, they do not discuss performance on the message evaluation task in terms of metalinguistic awareness explicitly.

This study is particularly relevant to this thesis in that it demonstrates the there is a clear relation, even with age removed as a factor, between performance on a referential opacity task and a task measuring metalinguistic awareness. Thus, it would seem that success on opacity tasks is strongly related to success on a task that requires metalinguistic awareness.

As we have seen from Kamawar and Olson (in press), there is a relation between performance on referential opacity and metarepresentational awareness (as measured by false belief). Apperly and Robinson (1999; Robinson & Apperly, April, 1999) demonstrate that performance on referentially opaque contexts is related to metalinguistic awareness. However, an important question not yet answered is whether both metarepresentational ability and metalinguistic ability are related to successful performance on referential opacity tasks. This leads directly to the very question addressed by Study Two.
Chapter 5: Study 1

5.1 Purpose

The primary purpose of this study was to answer the questions: (1) Is it the case that once a child is able to perform successfully on one type of opaque context, she is then able to perform successfully on all other types of opacity?; (2) If the answer to Question 1 is 'no', then, even if the types differ in difficulty, is performance across all of them still correlated even when factors such as age, digit span and vocabulary score are partialled out?; and (3) Is performance on opacity tasks similar for both object and action referents?

These questions were answered by comparing children's performance across a wide range of tasks. Specifically, I examined: (1) children's performance on referentially, linguistically, and intentionally opaque contexts (three opaque contexts); (2) children's ability to deal with two types of referent for each type of opacity (i.e., object and action referents were measured for each type of opaque context); (3) children's performance on tasks measuring their metarepresentational ability (false belief tasks) (4) children's general language ability (measured using the vocabulary subtest from the Stanford-Binet test of Intelligence); and (5) children's short-term memory (measured using the forward and backward digit-span subtest from the Stanford-Binet test of Intelligence).

5.2 Hypotheses

The hypotheses were as follows:

Hypothesis 1. The first hypothesis is that during their primary years (i.e.,
Kindergarten through Grade 3), children will advance from allowing substitution in all contexts to beginning to differentiate those contexts that allow substitution from those that do not. In other words, children's performance on opacity tasks will improve as they get older.

**Hypothesis 2.** The second hypothesis is that sensitivity to opaque contexts based on multiple representations of an object will be easier than (or the same as) those based on multiple representations of an action. This prediction is based on the belief that children have more experience talking explicitly about an object with more than one name/description (e.g., 'dog' = 'Fluffy') than they do with talking about an action under various descriptions (e.g., 'turning off the light' = 'scaring away the raccoon'). For example, children at quite a young age have experience with someone calling their mothers or fathers by their first names, so that 'mom' is known to some people as 'Sue'. They do not, I believe, have as much experience with actions being described explicitly under multiple descriptions. Often, actions are described in terms of their effects instead of as being seen as multiple descriptions of the same physical event.

**Hypothesis 3.** The third hypothesis is a question more than an expectation: Is performance consistent across the three types of opaque contexts, or are they ordered by difficulty?

**Hypothesis 4.** If there is a difference in difficulty among the three opacity types examined in this study (i.e., linguistically, and intentionally opaque contexts), the fourth hypothesis is that children will succeed on referentially opaque contexts before they become sensitive to intentionally opaque contexts. In other words, referentially opaque contexts are taken as 'basic' and if there are differences among types of opacity, it is expected that referentially opaque contexts will be the easiest. The reason for this expectation is that the other types of opacity have as a minimum the same requirements
as referentially opaque contexts.

**Hypothesis 5.** If there is a difference in difficulty among the opacity types, the fifth hypothesis is that children will become sensitive to referentially opaque contexts before they become sensitive to linguistically opaque contexts. The argument here is the same as for Hypothesis 4.

**Hypothesis 6.** The sixth hypothesis is that once age, vocabulary and digit span are partialled out, performance on false belief tasks will still be significantly correlated to performance on referentially, intentionally and linguistically opaque contexts. In other words, there will be a relation among the tasks that require metarepresentational ability.

**Hypothesis 7.** The seventh hypothesis is that all opacity measures (referential opacity, intentional opacity and linguistic opacity) will be related to general language ability (as measured by the vocabulary test) after age and digit span are partialled out. This is expected because all types of opacity require that children be able to deal with fairly sophisticated language constructions.

**Hypothesis 8.** The eighth, and final, hypothesis is that performance on the three types of opacity are related to one another even after age, vocabulary, and digit span have been partialled out. This is expected because all three types of opacity share common requirements.

### 5.3 Participants

Parental permission forms were distributed to children in Junior Kindergarten, Grade 1 and Grade 3 at the participating Toronto schools¹ and daycares. All children whose parents returned the forms were tested, though four children were excluded

---

¹ Permission was granted by the Scarborough Board of Education to send letters and consent forms out with the children in some of their schools. A copy of this letter can be found in Appendix A.
because they failed to correctly answer the key comprehension questions for the opacity tasks (details below). The children came from primarily lower-middle to middle income areas. Some participants did not speak English as a first language, but were judged by their classroom teachers to be competent speakers of English ('at or above grade level'). The sample was ethnically diverse, with no single group predominating.

Of the 73² children who participated, there were 23 participants classified as 4 years (11 boys and 12 girls, age range 4 years 1 month to 5 years 3 months, mean 4 years 9 months), 25 participants classified as 6 years (12 boys and 13 girls, age range 6 years 2 months to 7 years 2 months, mean 6 years 7 months) and 25 participants classified as 8 years (13 boys and 12 girls, age range 7 years 9 months to 9 years 6 months, mean 8 years 7 months).

5.4 Materials

5.4.1 Opacity Tasks. The purpose of these tasks was to provide a measure of each child's ability to deal with the three types of opaque contexts of interest (referential, linguistic and intentional). There were six colourfully illustrated stories about two friends (Mark and Cathy) with a few other characters in supporting roles (the characters' parents). The number of story characters was kept to a minimum in an effort to reduce the memory strain on the participants. The six scenarios were parallel in design. In each, there was either an object or an action that could be described in two distinct ways (the referent). There was also a story character in each (Mark or Cathy) who knew the object/action under only one description. In each story, the character with partial knowledge of the referent made some statement about it using the description under which he (or she) was aware of the referent. Further, it was made clear that the same character had an intention directed toward the object/action under

² This does not include the four children who were excluded.
one description of it, but not under the other (see Appendix B for the exact scenarios and illustrations).

Three of the six stories contained object referents with two distinct descriptions (and representations) such as a 'yellow ball' that was also the 'big ball', a 'police officer' who was also 'Cathy's Dad' and a 'doctor' who was also 'Mark's mom'. The other three stories contained action referents that could be referred to in two distinct ways, such as 'turning on the light' and 'scaring away the raccoon', 'pouring water out the window' and 'watering the plant', and finally, 'shaking the blanket clean [of crumbs]' and 'feeding the ants'.

Accompanying each story were colourful illustrations depicting key points. The stories were read to each child after he or she was told: “I’m going to tell you some stories about two friends, Mark and Cathy. As we go along, I’m going to ask you some questions about them, and what they did. So listen carefully, okay?” Following each story were comprehension questions as well as those testing for sensitivity to referential opacity (opaque by virtue of the mental verb 'know'), intentional opacity (opaque by virtue of the phrase 'mean to'), and linguistic opacity (opaque by virtue of using an indirect quote). There were two questions for each type of opacity. One asked about the story characters' thoughts (intentions, utterances) using the known description of the referent (known to both character and participant), while the other used the unknown description (known to the participating child though not the story character). An example of a story with an object referent, and one with an action referent, the questions, and what counted as correct answers are as follows:

**Story (Object Referent):**

One day, after Cathy finished talking to her friend Jenny on the phone, she got ready to have her friend Mark come over to play soccer. Cathy got out all of her toy balls and put them on her bed. She had a new yellow ball, it was the biggest one, and she also had a red one and a green one. The light bulb burnt out and she was in the dark. She didn’t have
time to ask her mom to change the bulb because just then, the doorbell rang. She went downstairs to let Mark in.

Cathy asked Mark to get a ball from her room. She told him that all of the balls were on the bed because she knew that they would be playing soccer. She said, "Get any ball you want except for the yellow ball because I don't want to get it dirty". As Mark goes up to Cathy's room he says, "I don't care what colour it is. I'm going to take the big ball." He wanted the big ball because it's best for playing soccer.

Upstairs in Cathy's room, Mark can't see anything because the light is out. He finds his way to the bed, he feels around until he finds the biggest one. He can't tell that it's the yellow ball. He picks it up and starts to leave Cathy's room. [note: questions are asked while Mark is still in the dark]

Questions:

Comprehension:
Did Mark take the big ball? [correct answer: yes] - transparent context
Did he take the yellow ball? [yes] - transparent context
Did he know that Cathy was talking on the phone before he came over? [no]

Referential Opacity:
Did he know that he took the big ball? [yes]
Did he know that he took the yellow ball? [no]

Intentional Opacity:
Did he mean to take the big ball? [yes]
Did he mean to take the yellow ball? [no]

Linguistic Opacity:
Remember what Mark said to Cathy. He said, "I'm going to take ..."
Did he say that he was going to take the yellow ball? [no]
Did he say that he was going to take the big ball? [yes]

Story (Action Referent):

One night, Mark was sitting in the kitchen having a snack. The back door of the house was right by the kitchen. Outside, right by the back door, there was a raccoon in the dark. It was very, very quiet so Mark didn't hear it.

Mark's dad was upstairs wrapping a surprise gift for Mark. He called down, "Mark, please turn on the light by the back door."

Mark got up and flicked the light switch and the light just above the back door turned on. Mark said to his dad, "Okay Dad, I turned on the light". And guess what? - That light scared away the raccoon. Mark didn't see that the light scared the raccoon and it ran away so quietly Mark

---

3 Note: The questions did not appear in this order.
4 Either 'Mark' or 'he' was used, depending on the order in which the questions appeared.
couldn't tell that a raccoon was ever there.

Questions:

Comprehension:
Did Mark turn on the light? [correct answer: yes] – transparent context
Did he scare away the raccoon? [yes] – transparent context
Did he know that his dad was wrapping a surprise gift? [no]

Referential Opacity:
Did he know that he turned on the light? [yes]
Did he know that he scared away the raccoon? [no]

Intentional Opacity:
Did Mark mean to turn on the light? [yes]
Did he mean to scare away the raccoon? [no]

Linguistic Opacity:
Remember what Mark said to his dad. He said, "Okay Dad, I ..."
Did he say that he turned on the light? [yes]
Did he say that he scared away the raccoon? [no]

The comprehension questions were always asked before the opacity questions (though not always in the order they appear in above). In order to proceed to the opacity questions, each child was required to answer the transparent questions correctly. If a child was not able to do so for a given story, then the key points of the story were repeated and the child was asked them again. If a child was still not able to answer those two questions correctly, she was given some stickers and returned to her classroom. This was done to ensure that children were aware of the critical points of the story; for example, that Mark took the big ball which meant he did in fact take the yellow ball. Without this understanding, the child's answers to the opaque questions would not make sense. For example, if a participating child did not know that Mark had taken the 'yellow ball', her answer to the question "Did Mark mean to take the yellow ball?" would be difficult to interpret. The purpose of the third comprehension question was to have a question regarding the story content that had an embedded mental verb.

Note: The questions did not appear in this order.

Either 'Mark' or 'he' was used, depending on the order in which the questions appeared.

Four children were excluded for failing to correctly answer the transparent questions.
for which the correct answer was 'yes'. This purpose of such a question was to make sure that children did not think that every question with an embedded mental verb required a 'no' answer.  

In order to be scored correct on a given story, for a given type of opaque context, children had to get both questions for that type of opacity correct. For example, in the first story above, children had to be able to say both that Mark knew that he took the big ball ('yes'), and that he did not know that he took the yellow ball ('no'). Thus, in order to be given credit for being correct on a particular type of opacity (for a given story), children had to allow that the story character had a particular belief about (intention toward, said something about) the object under the known description, but not under another, though equally legitimate, unknown description of the same referent. I believe that this scoring method provides a superior measure of children's ability to deal with opaque contexts (as opposed to just being able to answer 'no' to the question with the substituted description alone). In order to demonstrate that a child really has a grasp of opacity, she should both accept that a character knew (or had an intention toward, or said something about) a referent under the known description, and reject that the character knew (or had an intention toward, or said something about) the referent under another, unknown, description. Therefore, this scoring method is sensitive to the participant's ability to distinguish between the description the character had knowledge of from the one that was unknown to her.

Children's scores on the opacity questions were summed to yield composite scores. Thus, each child earned a score out of 3 for each type of opacity with object referents, a

---

8 When designing this study, I erroneously constructed a question requiring a 'no' answer instead of a 'yes' one; this error was corrected for Study 2. However, I do not believe that it is a problem for this study because the scoring system required that children be able to appropriately make both 'yes' and 'no' responses to questions containing opaque contexts.
score out of 3 for each type of opacity with action referents, a score out of 6 for each type of opacity (summed across referent types) and a score out of 9 for each type of referent (summed across opacity types). These summed scores were used in all analyses, except where otherwise noted.

Having three questions for each type of opacity (by each type of referent) enabled me to obtain a more sensitive measure of each child's ability. Composite scores (out of 3, 6 or 9) can be used in some analyses (e.g., correlations) that aren't legitimate when there is only one question of each type. It was therefore possible to determine the relation among performance on the various types of opacity and the other measures of interest. It is to those measures that I will now turn.

5.4.2 Memory measure. The memory measure that was used was the digit span subtest of the Stanford-Binet Test of intelligence. Children were required to remember a series of single digit numbers and repeat them back to the tester in either the order in which they heard the numbers (digit forward) or in the opposite order to that in which they heard them (digit backward). Children were administered the task until they reaching ceiling (3 or more wrong out of 4 on two consecutive pairs of items). Their raw score (the total of forward and backward versions) was used in subsequent analyses.

The information provided by this measure was used to determine the extent to which success (or lack thereof) on the opacity tasks was due to memory constraints on the part of the child. Children's scores on this measure were partialled out of other correlations so that it would be clearer whether the relation between performance on opacity tasks and other tasks was due to something as simple as their ability to remember stories, or something more meaningful.

5.4.3 Language measure. The language measure was the vocabulary subtest from the Stanford-Binet Test of Intelligence. This is a standardized measure that requires
children to label figures and define terms. For example, one item was a drawing of a lamb, and the child had label it as a 'lamb' or a 'sheep'. Another item required the child to say what the word 'dollar' means. Acceptable answers included '100 cents' or 'it's money'. The task was administered to each child until she or he reached ceiling (3 or more incorrect answers out of 4, on two consecutive pairs of items) and raw scores were used in the analyses.

The purpose of this measure was to allow me to determine the extent to which general vocabulary is related to performance on the other tasks. Further, scores from this task can be partialled out to see what types of relations remain among other measures.

5.4.4 False Belief Tasks. A number of false belief tasks were administered, together measuring children's ability to recognize false beliefs in themselves and in others. Children were administered the following false belief tasks: (1) an unexpected contents task using a Crayola box filled with stickers (a question for self and a question for other); (2) an appearance/reality task using a rock that looks like a sponge (a question for self, a question for other, and one about appearance vs. reality); and (3) a change of location task in which one child moves another's toy from a red container to a yellow one while the first child is away (question for other). For example, in the 'unexpected contents task', children were asked what was inside a Crayola crayon box (which was closed in front of them). Children tended to respond 'crayons' (answers like 'colours' were accepted too). The box was then opened and children were shown that the box contained stickers. The stickers were put back in the box, the box was closed and the children were asked what they had thought was in the box before it was opened and what someone else, who hadn't seen inside the box would think was in it (not always in this order). The correct answer for both questions was 'crayons' (or whatever the child
had answered to the first question).

Children were asked a total of 6 questions across the 3 tasks (i.e., self/other/appearance-reality). Their scores on these questions were summed for a score out of 6 points (which was used in all analyses unless otherwise noted). They were also administered a true belief task to ensure that they could remember an earlier, true, belief of theirs (see Appendix C for the exact tasks). The purpose of this task was to ensure that any errors children made on false belief questions could be attributed to their lack of false belief understanding and not due to a failure to recall their earlier utterances.

5.5 Procedure and Design

Each child was interviewed individually in two sessions separated by at least one day, and not more than five. The procedure was the same for all participants; they were introduced to the researcher, given a brief introduction as to what was happening (they were informed that they were going to hear stories and by doing that, would be helping me with my 'homework') and told that they if they wanted to, they could go back to class, and that no one would be upset if they did not want to hear any 'stories'. At the end of each session, each child was given some stickers or a colourful pencil (their choice) and then returned to the classroom.

During one of the two sessions, children received three of the opacity stories (and questions), the false belief and the digit span tasks. During the other session, children received the other three opacity stories (and questions), and the vocabulary task. Each session lasted approximately 15 minutes. The two sessions, the key questions within the opacity and false belief tasks, and the measures were ordered differently across participants.

Four children were excluded from participating in this study for not being able to
correctly answer transparent context questions on their second attempt. Of the 73 children who participated, the number who were correct on the first attempt across the six stories were: 61 on all stories; 10 on five of the stories; 1 on four of the stories; and 1 on three of the stories (mean number correct on the first attempt is 5.79/6). The results reported below are for the 73 children who, by at least their second attempt, were able to correctly answer the transparent context questions.

5.6 Results

5.6.1 Methods of Analysis. The data were entered on a Macintosh computer using SPSS (Release 6.0). The primary statistical tests used were analysis of variance, Chi-Square test for independence, and partial correlations. Participants' performance was analyzed using their raw/summed scores on each task/measure unless otherwise noted (e.g., categorical scores were used for the Chi-square analyses). Children's ages were measured in months for correlational purposes, and coded into 3 groups (4, 6 and 8 years) for use in ANOVAs.

A repeated measures ANOVA\(^9\) was performed with Opacity (3 levels) and Referent (2 levels) as within subject factors, and with the sex of the participant (2 levels), session order (2 levels) and the order of the key questions (3 levels) as between subject factors. The results showed no main effects for sex, session order and form (order of key questions). Further, no interaction effects were found for these variables. Therefore, orders, forms and sexes were combined for further analysis.

5.6.2 ANOVA Results (Hypotheses 1 – 3) The main analysis for this first study was a repeated measures ANOVA with Referent and Opacity entered as the within

---

\(^9\) All and only significant results will be reported unless a non-significant result directly bears upon a Hypothesis or question of interest.

\(^{10}\) In regard to the assumption of homogeneity of variance, ANOVAs are considered conditionally robust if the group sizes (at each level) are equal or approximately equal (Stevens, 1996), as is the case in this study.
subject factors. The Opacity factor had three levels (referential, linguistic and intentional opacity) and the Referent factor had two levels (object and action). Age (coded into 4, 6 or 8 years) and ESL status (non-ESL or ESL) were entered as the between subject factors (see Table 1 for a summary of performance on the within subject factors summed across age as well as by age group).
Table 1.

Descriptive Statistics for Referent, Opacity and Referent by Opacity Scores for Study 1.

<table>
<thead>
<tr>
<th>Measure</th>
<th>All Participants (N=73)</th>
<th>4-year-olds (n=23)</th>
<th>6-year-olds (n=25)</th>
<th>8-year-olds (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Object Referent a</td>
<td>4.81</td>
<td>2.82</td>
<td>1.78</td>
<td>1.51</td>
</tr>
<tr>
<td>Action Referent a</td>
<td>5.89</td>
<td>3.06</td>
<td>2.70</td>
<td>2.72</td>
</tr>
<tr>
<td>Intentional Opacity b</td>
<td>2.82</td>
<td>2.17</td>
<td>.87</td>
<td>1.18</td>
</tr>
<tr>
<td>Linguistic Opacity b</td>
<td>3.53</td>
<td>1.96</td>
<td>1.61</td>
<td>1.56</td>
</tr>
<tr>
<td>Referential Opacity b</td>
<td>4.34</td>
<td>2.08</td>
<td>2.00</td>
<td>1.93</td>
</tr>
<tr>
<td>Object – Intentional c</td>
<td>1.15</td>
<td>1.09</td>
<td>.30</td>
<td>.47</td>
</tr>
<tr>
<td>Object – Linguistic c</td>
<td>1.60</td>
<td>1.12</td>
<td>.65</td>
<td>.83</td>
</tr>
<tr>
<td>Object – Referential c</td>
<td>2.05</td>
<td>1.13</td>
<td>.83</td>
<td>.98</td>
</tr>
<tr>
<td>Action – Intentional c</td>
<td>1.67</td>
<td>1.30</td>
<td>.57</td>
<td>.95</td>
</tr>
<tr>
<td>Action – Linguistic c</td>
<td>1.93</td>
<td>1.16</td>
<td>.96</td>
<td>1.26</td>
</tr>
<tr>
<td>Action – Referential c</td>
<td>2.29</td>
<td>1.09</td>
<td>1.17</td>
<td>1.23</td>
</tr>
</tbody>
</table>

a out of 9.  b out of 6.  c out of 3.
In order to address some of the research questions, specifically those in Hypotheses 1, 2 and 3, the ANOVA was performed using a planned Helmert contrast for factors with more than 2 levels (i.e., Opacity and Age). A Helmert contrast compares performance on each level in the factor to the averaged performance on the remaining levels (the experimenter determines the orders of the levels within the factor). In the case of Opacity, the first Helmert contrast was between performance on referential opacity and the averaged performance on linguistic and intentional opacity. The second Helmert contrast compared performance on linguistic opacity with performance on intentional opacity. In the case of Age, the first contrast was between the youngest age group (4-year-olds) and the averaged performance of the two oldest age groups (6- and 8-year-olds), and the second contrast between the two oldest age groups (6- versus 8-year-olds).

Age. The results from this ANOVA established that children's performance improves significantly with age (F (2, 67) = 24.25, p < .001). The planned contrast revealed that there is a significant difference between the youngest group and the average of the two oldest groups (F (1, 67) = 24.94, p < .001), and between the two oldest groups (F (1, 67) = 23.56, p < .001. The mean score on opacity (across all types and referents) was 4.48/18 for the 4-year-olds, 11.4/18 for the 6-year-olds, and 15.72/18 for the 8-year-olds.

ESL Status. There was no main effect for ESL status.

Referent. A main effect for Referent was found (F (1, 67) = 21.12, p < .001) with children performing better on contexts containing action referents than on contexts containing object referents (see Table 1; see Figure 1), contrary to expectations.

---

11 While a different type of contrast (e.g., Repeated contrast), would have been preferred, such a contrast would have violated the assumption of orthogonality (independence). Therefore, Helmert contrasts have been used.
Figure 1. Mean scores (+SE) on object and action referents by age.
Opacity. There was a main effect for Opacity ($F(2, 67) = 15.78, p < .001$) indicating that performance was not the same across all types of opaque contexts. The first Helmert contrast revealed a difference between performance on referential opacity and the average of linguistic and intentional opacity ($F(1, 67) = 25.29, p < .001$, see Table 1 for means), and the second Helmert contrast revealed a difference between linguistic and intentional opacity ($F(1, 67) = 11.29, p < .01$). In order to clarify what was happening in the first Helmert contrast, a post-hoc paired t-test was performed. It revealed that referentially opaque contexts were easier than linguistically opaque contexts ($t(72) = -4.96, p < .001$), which in turn were easier than intentionally opaque contexts (determined by second Helmert contrast above; see Figure 2).

Two-Way Interactions. An interaction was found between Age and Opacity ($F(2, 67) = 2.95, p < .05$) with the effect being specific to the first Helmert contrast (referential opacity and the average of linguistic and intentional opacity) and only between the two older age groups ($F(1, 67) = 9.60, p < .01$). However, given that the oldest group (the 8-year-olds) was virtually at ceiling on the referential opacity task (see Figure 3), not too much can be said of this interaction. Given that there is a ceiling effect for the oldest group, only the two youngest groups' scores should be examined to see if this two-way interaction holds up. As stated above, however, the interaction was specific only to the second Helmert contrast for Age and did not hold for the 4- and 6-year-olds. Therefore, it would appear that the interaction is an artifact of the testing measure and that this two-way interaction is due to the nearly perfect performance for the 8-year-olds.
Figure 2. Mean scores (+SE) on referential, linguistic and intentional opacity by age.
Figure 3. Mean scores (+SE) on referential, linguistic and intentional opacity by age.
There was also an interaction effect between Opacity and ESL status ($F (2,67) = 3.34, p < .05$), with the effect being specific to the contrast between linguistic and intentional opacity. Those children who spoke English as a second language had a larger difference in performance between these two types of opacity than the non-ESL group, with their intentional opacity scores being higher (see Figure 4). This finding needs to be qualified because this study was not designed to examine differences between ESL and non-ESL children, and therefore, there are not equal numbers of each at each level. The mean age of the ESL children is considerably higher than the mean age of the non-ESL children (7 years 9 months, and 5 years 11 months respectively). Therefore, it seems possible that this finding has less to do with ESL status than it has to do with age effects$^{12}$.

---

$^{12}$ Note: I also performed an ANCOVA with ESL status as a covariate because the effect of ESL status was not a research question. The results from the ANCOVA, except for those interactions involving ESL status, were the same as those in presented analysis. However, I chose to report this analysis instead of the ANCOVA because even though I cannot say too much about any interaction involving ESL status given its connection with age, it is possible that such interactions are hinting at something interesting that needs to be explored in future work.
Figure 4. Mean scores (+SE) on linguistic and intentional opacity by ESL status.
**Three-Way Interactions.** A three-way interaction was found among ESL status, Referent and Age ($F (2,67) = 3.55, p < .05$). The planned Helmert contrast revealed this interaction to be between the 4-year-olds and the average performance of the 6- and 8-year-olds ($F (1,67) = 4.94, p < .05$). However, given that there is only one person in one of the cells (only one 4-year-old was ESL), and therefore the cell sizes are very uneven, little can be said of this interaction.

Another three-way interaction was found among ESL status, Referent and Opacity ($F (2,67) = 4.22, p < .05$). The Helmert contrast revealed that this interaction was specific to linguistic and intentional opacity ($F (1,67) = 8.19, p < .01$). This three-way interaction was the result of there being a difference between each of the two-way interactions between ESL status and Referent for the two types of opacity. By producing errorplots of the two, two-way interactions, it becomes clearer why there is a significant three-way interaction. Figure 5a reveals that for intentional opacity, there is a significant difference between referent types for the non-ESL children (paired t-test: $t (41) = 4.04, p < .001$), but not for the ESL children (paired t-test: $t (30) = 1.94, \text{ns}$). Figure 5b reveals that for linguistic opacity, the pattern is the opposite: there is a significant difference between Referent types for the ESL children ($t (30) = 2.09, p < .05$), but not for the non-ESL children ($t (41) = 1.43, \text{ns}$). For both, it appears that the ESL children are performing better than the non-ESL children. It is this overall pattern of differences between the two-way interactions that results in the three-way interaction. Again, this finding needs to be qualified given the difference in ages between the non-ESL and ESL children.
Figure 5a. Mean scores and 95% confidence intervals for intentional opacity by ESL status.
Figure 5b. Mean scores and 95% confidence intervals for linguistic opacity by ESL status.
**Four-Way Interactions.** One four-way interaction was found among ESL status, Referent type, Age and Opacity ($F(4,134) = 3.24, \ p < .05$). The planned contrasts revealed that the interaction was taking place between the two oldest age groups (the 6- and 8-year-olds), and was specific to linguistic and intentional opacity ($F(1,67) = 5.64, \ p < .05$). The nature of this interaction becomes a little clearer when graphed (see Figures 6a and 6b). As can be seen in Figure 6a, among the 6-year-olds, there is little difference in performance between ESL and non-ESL children for linguistic opacity for either action referents ($F(1,23) = .8, \ p = .38$) or object referents ($F(1,23) = 1.48, \ p = .24$), nor is there any difference for intentional opacity with object referents ($F(1,23) = .05, \ p = .83$). However, there is a great difference between the two subgroups when it comes to intentional opacity for action referents ($F(1,23) = 18.97, \ p < .001$). Figure 6b reveals that for the 8-year-olds, there is a different pattern. There is no difference in performance between ESL and non-ESL children on any of the scores (all $p > .12$).
Figure 6a: Interaction Among Opacity, ESL Status and Referent Type
(Age = 6 years)
Figure 6b: Interaction Among Opacity, ESL Status and Referent Type
(Age = 8 years)
Interesting non-significant result. It is worth noting that there was no significant interaction between Referent and Opacity which indicates that the referent effect (that of object being harder than action) is generally consistent for all types of opacity and its corollary that the opacity effect is consistent for both types of referent. This statement does, however, need to be qualified because there may be interactions between opacity and referent types in certain situations, and this is what the 3- and 4-way interactions are hinting at. However, on the whole, the patterns are so striking that it seems that there is a difference for the two types of referent for each type of opacity, and a difference between the types of opacity for both types of referent.

5.6.3 Chi-Square Analyses (Hypotheses 4 and 5). Given that there were differences in difficulty among the types of opacity, Chi-square tests for independence were performed to test Hypotheses 4 and 5. Summed scores were converted into pass/fail scores (i.e., summed scores were converted into dichotomous data). A pass was considered 5 or more out of 6 (on opacity scores summed across referent types) and 2 or more out of 3 (on opacity scores by referent types). Analyses using a stricter pass criterion (i.e., 6/6 and 3/3) yielded the same results unless otherwise specified.

A comparison of performance on referential and intentional opacity (collapsed across Referent) revealed that sensitivity to both types of opacity is related ($\chi^2(1, N=73) = 10.56, p<.01$). Of the children who failed referential opacity, very few passed intentional opacity. Specifically, of the 24 children who failed referential opacity, only 1 passed intentional opacity. Passing referential opacity, however, did not guarantee success on intentional opacity. Of the 49 children who passed referential opacity, only 20 passed intentional opacity (See Table 2 in Appendix D).

This relation persisted for both object referents ($\chi^2(1, N=73) = 16.40, p < .001$) and action referents ($\chi^2(1, N=73) = 25.58, p < .001$). When dealing with object referents,
of the 22 who failed referential opacity, none passed intentional opacity. Passing referential opacity did not ensure success on intentional opacity (only 25 of the 51 who passed referential opacity did the same on intentional opacity, see Table 3 in Appendix D). When dealing with action referents, of the 15 children who failed referential opacity, none passed intentional opacity. Once again, passing referential opacity did not guarantee success on intentional opacity, as only 42 of the 58 who passed referential opacity passed intentional opacity (See Table 4 in Appendix D).

A similar pattern was found between performance on referential opacity and linguistic opacity summed across referent types ($\chi^2 (1, N=73) = 21.5, p < .001$). While failing referential opacity tended to result in failing linguistic opacity, passing did not ensure success. Of the 24 children who failed referential opacity, only 1 passed linguistic opacity. Only 30 of the 49 who passed referential opacity succeeded on linguistic opacity (See Table 5 in Appendix D). This is consistent with the previously reported result that a paired t-test revealed that children found referentially opaque contexts easier than linguistically opaque ones.

This pattern of results also persisted for both object referents ($\chi^2 (1, N=73) = 18.62, p < .0001$) and action referents ($\chi^2 (1, N=73) = 18.30, p < .001$). When dealing with object referents, of the 22 children who failed referential opacity, only 3 passed linguistic opacity. Of the 51 who passed referential opacity, 35 passed linguistic opacity (see Table 6 in Appendix D). Of the 15 children who failed referential opacity with action referents, 4 passed linguistic opacity. Of the 58 who passed referential opacity, 48 passed linguistic opacity (see Table 7 in Appendix D).

5.6.4 False Belief Tasks. Performance was consistent with other work in this area (see Figure 7). Given that the two older groups of children were virtually at ceiling (4-

---

13 If the stricter pass criterion is used, the effect is only marginally significant ($\chi^2 (1, N=73) = 3.60, p = .058$).
year-olds mean = 4.04/6; 6-year-olds mean = 5.76, and the 8-year-olds 5.96) the scores
were converted into categorical data and Chi-square tests for independence were
performed to determine whether there was an age effect. Summed scores were
converted into pass/fail scores (i.e., summed scores were converted into dichotomous
data). A pass was considered 5 or more out of 6. The results from this analysis
revealed an age effect ($\chi^2 (2, N=73) = 29.84, p<.0001$; see Table 8 in Appendix D)
thereby supporting the view that children's performance on false belief tasks improves
with age.

5.6.5 Correlational Analyses (Hypotheses 6 -8). When performing correlations, age
was measured in months, raw scores were used for vocabulary and digit span, and
summed scores were used for all other measures (i.e., for opacity, referent and opacity
by referent scores). Table 9 (see Appendix E) provides a correlation matrix for all
measures/scores with nothing partialled out. As can be seen, every measure/score is
significantly correlated with every other measure/score ($p < .001$, with one exception
where $p < .01$). Partial correlations were performed to address specific questions (see
Table 10 for a summary of the partial correlations).

---

14 Analyses using a stricter pass criterion (i.e., 6/6) yielded the same results.
Table 10.

Partial correlation matrix: Age, vocabulary and digit span performance partialled out.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ref. Opacity</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Lin. Opacity</td>
<td>.31**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Int. Opacity</td>
<td>.23</td>
<td>.31*</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Object Ref.</td>
<td>.45***</td>
<td>.63***</td>
<td>.73***</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Action Ref.</td>
<td>.73***</td>
<td>.57***</td>
<td>.57***</td>
<td>.45***</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Obj. - Ref.</td>
<td>.84***</td>
<td>.33**</td>
<td>.27*</td>
<td>.57***</td>
<td>.54***</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Obj. - Lin.</td>
<td>.09</td>
<td>.70***</td>
<td>.33**</td>
<td>.69***</td>
<td>.18</td>
<td>.08</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Obj. - Int.</td>
<td>.04</td>
<td>.26*</td>
<td>.82***</td>
<td>.75***</td>
<td>.21</td>
<td>.15</td>
<td>.29*</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Act. - Ref.</td>
<td>.86***</td>
<td>.21</td>
<td>.13</td>
<td>.20</td>
<td>.70***</td>
<td>.45***</td>
<td>.08</td>
<td>-.08</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Act. - Lin.</td>
<td>.35**</td>
<td>.72***</td>
<td>.11</td>
<td>.23</td>
<td>.63***</td>
<td>.39**</td>
<td>.20</td>
<td>.08</td>
<td>.21</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Act. - Int.</td>
<td>.34**</td>
<td>.25*</td>
<td>.84***</td>
<td>.47***</td>
<td>.72***</td>
<td>.30*</td>
<td>.25*</td>
<td>.39**</td>
<td>.28*</td>
<td>.11</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>12. FB</td>
<td>.51***</td>
<td>.15</td>
<td>.17</td>
<td>.23</td>
<td>.40**</td>
<td>.44***</td>
<td>.02</td>
<td>.05</td>
<td>.42***</td>
<td>.19</td>
<td>.23</td>
<td>--</td>
</tr>
</tbody>
</table>

Note. All correlations have N = 73.  
* p < .05  ** p < .01  *** p < .001
Even with age, vocabulary and digit span partialled out, there was a significant relation between false belief and referential opacity performance (r = .51, two-tailed, p < .001), but contrary to expectations, the same could not be said for either intentional or linguistic opacity.

Another partial correlation was performed examining the relation between each type of opacity and vocabulary scores with age and digit span partialled out. It was found that vocabulary was significantly correlated to referential opacity (r = .36, two-tailed, p < .01), to linguistic opacity (r = .35, two-tailed, p < .01) and to intentional opacity (r = .24, two-tailed, p < .05).

A final correlation was computed examining the relations among performance for all opacity and referent types after age, vocabulary and digit span had been partialled out. Performance on intentional and linguistic opacity were correlated (r = .31, two-tailed, p < .05). Referential opacity was significantly related to linguistic opacity (r = .31, two-tailed, p < .01) and marginally related to intentional opacity (r = .23, two-tailed, p = .051). However, given that Hypothesis 8 had predicted a positive correlation among the different types of opacity, we can also examine the results of a 1-tailed partial correlation between referential and intentional opacity which is significant (r = .23, one-tailed, p < .05). Performance on contexts with action referents was significantly related to those with object referents (r = .45, p < .001), even when age, digit span, and vocabulary were partialled out.
5.7 Discussion

For the most part, the results confirmed the hypotheses with a few notable and interesting exceptions. The results will be discussed within the framework provided by the Hypotheses and Results sections above.

5.7.1 Age. Not surprisingly, the hypothesis that children's sensitivity to opaque contexts improves during their primary school years was supported by the results, which showed a substantial increase in performance between ages 4 and 8 years. There is not too much to say about this finding other than that it is consistent with a great deal of research showing that children's performance on a wide range of tasks improves as they get older.

5.7.2 Referent. The hypothesis that children's performance on opaque contexts with object referents would be better than on than those containing action referents was not supported by the results; in fact, children found opaque contexts with action referents to be easier than those with object referents consistently for each type of opacity. I predicted that object referents would be easier because I believed that children had more experience explicitly talking about an object with multiple descriptions (e.g., 'my dog' = 'Fido') than they did talking about an action with multiple descriptions. While it may be true that children have more of this type of experience with objects than with actions, that experience might be overshadowed by other differences between the two referent types.

It may be the case that children were more successful with action referents because they treat multiple descriptions of an object as referring to the same referent while they treat multiple descriptions of an action as each referring to a separate action. For example, when children fail opacity tasks by saying something like 'Mark meant to take the yellow ball' and 'Mark meant to take the big ball', it is because they are treating
both descriptions as referring to the *same* object in the *same* way (no understanding of sense). However, when children come across multiple descriptions of an action, it may be the case that they sometimes treat the two descriptions as referring to *two different* actions; in other words, they act as if 'Mark turned on the light' and 'Mark scared away a raccoon' are descriptions of two actions that Mark performed, and not as two descriptions of one action that Mark performed. Notice that this explanation is consistent with the fact that all participating children were able to say that 'Mark took the big ball', 'Mark took the yellow ball', 'Mark turned on the light' and 'Mark scared away the raccoon'. For both kinds of referents it is relatively easy to answer 'yes' to these questions even if children are making the kind of error I have outlined above.

The difference between the two becomes apparent when we look at opaque contexts. If children are behaving as if Mark performed two actions, then opaque contexts become easier; the questions can be seen as being about two different events. Using intentional opacity as an example, I'll outline what might be happening on the part of the child. If the child treats two descriptions as referring to two separate events, it is then easy to 'tag' one event (i.e., 'turning on the light') as intentional and the other (i.e., 'scaring away the raccoon') as unintentional. Note that this is easier than tagging the descriptions as picking out intentional/unintentional aspects of the same action.

In the case of object referents, the story is somewhat different. It is not likely that children represent the descriptions 'yellow ball' and 'big ball' as referring to two different objects. Objects take up physical space, and unlike actions, are tangible. The child can see from the illustrations that the 'big ball' is the 'yellow ball'. Therefore, children are unlikely to make the error of thinking that the two descriptions pick out two referents. However, I propose that for actions, it is easier to make just such an error. In brief, it may be that some children hold the mistaken belief that multiple descriptions
refer to different actions, and that this error results in making opaque questions easier by allowing children to treat them as questions about different events.

Another alternative, motivated more by methodological issues than by conceptual ones, is that the stories with action referents were somehow easier than the stories with object referents. In other words, the effect may not be because of a difference between object and action referents directly, but as a result of a difference between stories with object referents and those with actions. In examining the text of the stories, I could not find any differences that would account for the pattern described above. In examining the pictures, however, it may be noticed that while the pictures depict the ball as both large and yellow, in the action stories, the turning on of the light and the scaring away of the raccoon occur in different pictures, perhaps facilitating the distinction.

Unfortunately, it is not possible to choose between these two alternatives by examining the data more closely. In order to determine which alternative is correct, it is necessary to conduct a follow-up study to ascertain whether children recognize that the two descriptions of a referent both refer to the same referent. For example, for object referents, we could ask something like "Is the 'big ball' the 'yellow ball'?" and for action referents, we could ask something like "Does 'turning on the light' 'scare away the raccoon'?". However, once the questions are phrased this way, the difference between the two types of referents becomes clearer and it seems that the first explanation is more likely. While the description 'turning on the light' refers to the same action described by 'scaring away the raccoon', talking about them often includes phrases like 'causes' or 'results in' connecting the two descriptions; for example, 'turning on the light' resulted in 'scaring away the raccoon'. Stated this way, they sound like two separate actions. This is not case for object referents: it would be strange to say 'taking the big ball
resulted in taking the yellow ball'. When phrased that way, it too makes it seem as if two things were taken. Since the story is clear that only one ball was taken, this interpretation is blocked. This supports the view that action referents work somewhat differently than object referents. It is therefore possible that because descriptions of action referents are more likely to erroneously be thought of as referring to different actions that children do better with them.

5.7.3 Opacity. The third hypothesis was phrased in the form of a question: "Is performance consistent across the three types of opaque contexts, or are they ordered by difficulty?" The results are quite striking in their support of the view that sensitivity to opaque contexts is clearly not an all-or-nothing ability. It is specific to different types of opaque contexts, and is ordered by difficulty, with referential opacity as the easiest, intentional opacity as the hardest, and linguistic opacity between the two. The lack of an interaction between Opacity and Referent indicates that this effect is generally consistent across referent types\(^\text{15}\) (see sections 6.7.4 and 6.7.5 below for a discussion of the differences among opacity types).

Even though performance on the three types of opaque contexts differed significantly from each other, they all were inter-related even when the effects of age, vocabulary and digit span were statistically removed. This supports the view that while not identical, all three types of opaque contexts have something in common. I propose that this inter-correlation among all the types of opacity is indicative of their shared requirements. Success on opaque contexts requires many things, such as: (1) the ability to allow multiple representations/descriptions of a referent; (2) the ability to hold the representations/descriptions separately and see them as distinct ways of picking out a referent; (3) the realization that access to an object under one description does not

\(^{15}\) As mentioned earlier, this statement needs to be qualified somewhat. There may be particular situations where this is not the case, as hinted by 3- and 4-way interactions.
ensure access to it under all (partial knowledge); and (4) to 'see' the words used to refer and not simply see 'through' them to the referent. While there are other requirements that differ from opaque context to opaque context (which accounts for the differences among them), there is nonetheless some sort of unity among these three types of opacity.

An interaction was also found between opacity (specific to linguistic and intentional) and ESL status (shown in Figure 4). On the face of it, this result seems to indicate that ESL children performed better than the non-ESL children on intentional and linguistic opacity. However, I need to point out once again that the difference in ages between the ESL and non-ESL groups makes this finding difficult to interpret. The higher performance by ESL children may be explained by the fact that the ESL children are, on average, older, but it is not possible to determine this with certainty. There may be something interesting about the fact that for ESL children, there was a greater difference between linguistic and intentional opacity than there was for the non-ESL children), but unfortunately, we cannot know for sure. In order to clarify whether this pattern of results was due to ESL status, or simply an artifact of the particular sample, this study needs to be replicated using roughly equal numbers of ESL and non-ESL children at each age level.

5.7.4 Three-Way Interactions. A three-way interaction was found among ESL status and referent type, between the youngest age group and the averaged performance of the two oldest age groups (i.e., 4-year-olds versus 6- and 8-year-olds). Given that there was only child who fell into the cell containing 4-year-old ESL children, I do not believe any generalizations can be made about this interaction.

A second three-way interaction was found between ESL status and referent type,
specific to intentional and linguistic opacity. The interaction revealed that when dealing with intentionally opaque contexts, non-ESL children’s performance was better with action referents than with object referents, but that there was no difference between the two types of referents when dealing linguistically opaque contexts. The opposite pattern was found for ESL children who did not show a difference for action and object referent in intentionally opaque contexts, but when dealing linguistically opaque contexts, found action referents easier. Also, the ESL children’s level of performance seemed higher than the non-ESL children’s performance for both types of opacity. Given that there are unequal numbers of children at each level of ESL status, and that the average ages of each level of ESL status differ, the implications of this interaction are not at all clear. As before, a replication of this study, carefully controlling the numbers and ages of ESL/non-ESL children, is necessary to clarify whether this interaction is real, or just an artifact of this particular sample.

5.7.5 Four-Way Interactions. A four-way interaction was found among ESL status and referent type, specific to 6- and 8-year-olds dealing with linguistic and intentional opacity. Figures 6a and 6b presented this interaction graphically, showing that the interaction was due the fact that among the 6-year-olds, non-ESL and ESL children differ in their ability to deal with intentional opacity with action referents (ESL children have a lot more difficulty with them). It appears that the 6-year-old non-ESL children deviate from the usual pattern of finding intentional opacity (action referent) the most difficult. However, by the time children are 8 years old, no significant differences were found between the non-ESL and ESL children. As with the previous interaction (and like many 4-way interactions), the implications of this result are not clear.

16 Here I am talking about metalinguistic awareness. While this was not addressed in Study 1, it is still relevant though not supported by data.
5.7.6 *Referential and Intentional Opacity*. Support was found for Hypothesis 4 which predicted that if there was a difference among the three types of opacity, then children would succeed on referentially opaque contexts prior to intentionally opaque ones. This was predicted because referentially opaque contexts are considered primary; in other words, the cognitive requirements for intentionally opaque contexts include the same requirements as referentially opaque contexts, plus the added requirement of being able to understand a sophisticated way of describing intention (e.g., 'meant to'). Given that the same six stories were used to evaluate children's proficiency with both types of opacity, it seems fairly clear that the difference in performance between the two contexts can be attributed to the added difficulty of having to deal with a complex expression of intention. While younger children may be able to understand the concept of intention, it is not until they are older that they can deal with more advanced ways of talking about it (Astington, 1991; 1986; Kamawar & Astington 1999, April). It is for this reason children found intentional opacity more difficult than referential opacity.

5.7.7 *Referential and Linguistic Opacity*. Hypothesis 5, which predicted that sensitivity to referentially opaque contexts would precede sensitivity to linguistically opaque contexts, was also supported by the results. Remember that in indirect discourse, substitution of coreferential terms is not allowed if the speaker of the original utterance has no knowledge of the given referent under the substituted description or name. While successful performance on both referential and linguistic opacity requires the child to attend to the way in which a referent is represented (i.e., which description the character is aware of), linguistic opacity has the additional requirement of having to remember precisely what was said in order to compare the actual utterance with the paraphrased utterance. The child's task is not the relatively more straightforward one of comparing two utterances to see if they are identical, where *any* change results in a false
quotation (as in quotational opacity). Here, the task is more complicated. The child must remember the original utterance, and then she must make a judgment about a paraphrase of that utterance. If 'Mark said "I am going to take the big ball"' is true, and given that Mark doesn’t know that the ball is yellow, is it true then to say that 'Mark said that he was going to take the yellow ball'? Strictly speaking, no, it is not. However, for the young child with limited experience with indirect quotes (Ely & McCabe, 1993; Hickmann, 1993), this may not be clear at all.

5.7.8 Relation between False Belief and Opacity. The findings from this study only moderately support the sixth hypothesis which predicted that once vocabulary, digit span and age were partialled out, performance on false belief tasks would still be significantly correlated to performance on referentially, intentionally and linguistically opaque contexts.

The results supported the first part of the hypothesis by demonstrating that false belief and referential opacity were significantly correlated, replicating Kamawar and Olson (in press) and supporting the view that the two tasks share a common requirement (i.e., metarepresentational ability).

The results, however, do not support the hypothesis in that there was no a significant correlation between false belief and intentional and linguistic opacity (when age, digit span and vocabulary were partialled out). It is worth noting that this does not necessarily mean that metarepresentational ability is not a requirement for success on intentional and linguistic opacity. It may be the case that the additional abilities required for these two types of opacity are acquired so much later than metarepresentational understanding that a direct relation is difficult to see. In other words, the relation may be masked by the fact that by the time children start to become somewhat proficient with intentional and linguistic opacity, they are already at ceiling
in their false belief performance.

5.7.9 **Vocabulary and Opacity.** The results supported the seventh hypothesis which predicted a relation between all types of opacity and general language ability (as measured by the vocabulary test), even when age and digit span were partialled out. This result can be taken as support for view that success on opaque contexts requires the ability to deal with fairly sophisticated language constructions - constructions that are complex because they contain complement structures and/or embedded clauses (e.g., de Villiers, 1994). However, since vocabulary tests do not measure such syntactic properties, this result can be considered as support of the above view only in very broad terms.

5.8 **Summary and Conclusion**

Taken as a whole, the results are supportive of the Hypotheses presented in Section 5.2 above. As a way of summarizing the main findings, I will revisit the three questions that provided the framework for this study, this time providing answers for them.

The first question was "Is it the case that once a child is able to perform successfully on one type of opaque context, she is then able to perform successfully on all other types of opacity?" The data are clear in that the answer is 'no'. Children's performance on the three types of opacity (as measured in this study) differs by level of difficulty. For the reasons outlined above, we can see why children find referential opacity easier than both linguistic and intentional opacity. Referential opacity was predicted to be the most 'basic' version of the three and was expected to be acquired before the other two. The results bore out this prediction. Generally, before children become sensitive to either linguistic or intentional opacity, they have to become sensitive to referentially opaque contexts. These results support the view that dealing
with opaque contexts is not an all-or-nothing ability; opaque contexts are graded in terms of their requirements, and therefore differ in terms of difficulty.

The second question was "If the answer to Question 1 is 'no', then, even if the types differ in difficulty, is performance across all of them still related even when factors such as age, digit span and vocabulary score are partialled out?". Partial correlations reveal the three types of opacity to be related to one another even after the effects of factors such as age, memory span and general language ability have been removed. This supports the view that success on these three opaque contexts rests upon some common abilities (e.g., metarepresentation).

The third, and final, question was "Is performance on opacity tasks similar for both object and action referents?" Again, the data are clear that children found action referents significantly easier than object referents. While this finding went against original expectations, it made sense when the precise nature of the two types of referents was examined more closely. This finding sheds light on one area of opacity that, to my knowledge, has not yet been addressed by developmental research.

These results are of interest to anyone concerned with children's understanding of: (a) the representational nature of language; and (b) the fact that the language used to describe intentions, thoughts and utterances reflects particular representations of the world and not others. Overall, these results mark a clear and important transition from thinking about things, to thinking about the representation of things in language.
Chapter 6: Study 2

6.1 Purpose

The primary purpose of this Study was to answer the questions: (1) In addition to metarepresentational ability, does metalinguistic awareness predict performance on referentially opaque contexts?; and (2) What does the developmental trajectory for performance on quotationally opaque contexts look like and how is it related to referential opacity?

Referential opacity was chosen to be the main focus of this second study because it is the first type of opacity that children are able to succeed on (as seen from Study 1), and because it is the most basic form of opaque context of the three types examined thus far.1 Further, it is also the type of opacity that is most talked about in the developmental literature (e.g., Apperly & Robinson, 1998; 1999; de Villiers & Fitneva, 1996), and therefore will be of the most interest. Given that opacity tasks, in addition to being metarepresentational tasks, are also metalinguistic tasks, it is expected that metalinguistic awareness will play an important role in children's sensitivity to opaque contexts.

While some studies have found that children's performance on referentially opaque contexts is related to their performance on false belief tasks (de Villiers & Fitneva, 1996; Kamawar, 1996), and one study has discussed understanding of referentially opaque contexts in terms of the ability to deal with partial referential relations (Apperly & Robinson, 1999), none have examined performance on referentially

---

1 Recall that intentional opacity was a variant of referential opacity and that linguistic opacity was a variant of quotationally opacity (Section 2.2).
opaque contexts as requiring *both* metarepresentational and metalinguistic ability. Thus, the main goal of Study 2 was to determine whether children's performance on tasks of referential opacity is predicted by their metalinguistic awareness in addition to their metarepresentational ability.

The second purpose of Study 2 was to examine children's ability to deal with quotationally opaque contexts (opaque by virtue of containing a direct quotation). Recall that quotationally opaque contexts are opaque by virtue of containing direct quotations (see p. 19, this thesis). Also recall that this type of opacity is a 'primary' or 'basic' type of opacity (i.e., one of the three laid out by philosophers of language) of which linguistic opacity is a variant (see p. 8, this thesis).

The research questions were answered by comparing children's performance across a wide range of tasks. Specifically, I examined: (1) children's performance on referentially and quotationally opaque contexts (object referents); (2) children's performance on tasks measuring their metarepresentational ability (false belief tasks); (3) children's general language ability (measured using the vocabulary subtest from the Stanford-Binet Test of Intelligence); (4) children's short-term memory (measured using the forward and backward digit-span subtest from the Stanford-Binet Test of Intelligence); and (5) three measures of metalinguistic awareness (the say/refer, and two say/mean tasks).

6.2 Hypotheses

The hypotheses are as follows:

**Hypothesis 1.** The first hypothesis is that during their early years of elementary school, children will advance from allowing substitution of coreferential expressions in all contexts to beginning to be able to differentiate those contexts that allow substitution from those that do not. In other words, they will get better with age on both opaque contexts.
Hypothesis 2. The second hypothesis is not really a hypothesis, but a question for which I would like to find an answer: Is performance similar for both referentially and quotationally opaque contexts, or are they ordered by difficulty?

Hypothesis 3. The third hypothesis is that if there is a difference in difficulty between the two types of opacity, children would succeed on quotationally opaque contexts before they become able to deal with referentially opaque contexts. The reason for this expectation is that in order to succeed on referentially opaque contexts one must be able, at a minimum, to metarepresent a mental state, while no such ability is required for success on quotational opacity.

Hypothesis 4. The fourth hypothesis is that both metarepresentational ability (as measured by false belief task performance) and metalinguistic ability will predict performance on referentially opaque contexts (in a hierarchical regression) even after variability from factors such as age, vocabulary and digit span have been taken out. The reasons for this expectation are laid out in Chapter 3.

Hypothesis 5. The fifth hypothesis is that performance on quotationally opaque contexts will not be related to performance on false belief tasks, but will be related to metalinguistic ability (after age, digit span and vocabulary have been partialled out).

Hypothesis 6. The sixth hypothesis is that there will be a relation between performance on referentially and quotationally opaque contexts with age, digit span and vocabulary performance partialled out.

6.3 Participants

Parental permission forms and letters were distributed to children in Junior Kindergarten, Senior Kindergarten and Grade 1 at participating Toronto schools and daycares. All children whose parents returned the forms were tested, though three children were excluded because they failed to correctly answer the key comprehension questions for the opacity tasks (details below). The children came from primarily lower-
middle to middle income areas. Some participants did not speak English as a first language, but were judged by their classroom teachers to be competent speakers of English. The sample was racially and ethnically diverse, with no single group predominating.

Of the 64 participants who participated, there were 22 participants aged 5 years (14 boys and 8 girls, age range 4 years 5 months to 5 years 6 months, mean 5 years), 21 participants aged 6 years (9 boys and 12 girls, age range 5 years 7 months to 6 years 7 months, mean 5 years 11 months) and 21 participants aged 7 years (11 boys and 10 girls, age range 6 years 8 months to 7 years 5 months, mean 7 years).

6.4 Materials

6.4.1 Opacity Tasks. The illustrated stories used in this study were similar to those in Study 1. However, this time, only two types of opaque questions were asked (referential and quotational) and only object referents were used. As before, each story had a key character who knew the object under only one description (partial knowledge), and made a statement about the referent using the alternative description known to the child subject but not the story character (see Appendix F for the exact scenarios and illustrations).

As with Study 1, the object referents had two distinct descriptions (and representations); namely, a 'yellow ball' that was also the 'big ball', a 'police officer' who was also 'Cathy's Dad' and a 'doctor' who was also 'Mark's mom'. Accompanying each story were colourful illustrations depicting key points. The stories were read to each child after he or she was told: "I'm going to tell you some stories about two friends, Mark and Cathy. As we go along, I'm going to ask you some questions about them, and what they did. So listen carefully, okay?". Following each story were comprehension

---

2 This does not include the three children who were excluded.
questions as well as those testing for sensitivity to referential opacity and quotational opacity. An example of a story, the questions, and what were taken to be correct answers are as follows:

**Story:**

One day, Cathy got out all of her toy balls and put them on her bed. She had a new yellow ball, it was the biggest one, and she also had a red one and a green one. The light bulb burnt out and she was in the dark. She didn’t have time to ask her mom to change the bulb because just then, the doorbell rang. She went downstairs to let Mark in.

Cathy asked Mark to get a ball from her room. She told him that all of the balls were on the bed because she knew that they would be playing soccer. She said, “Get any ball you want except for the yellow ball because I don’t want to get it dirty”. As Mark goes up to Cathy’s room he says, “I don’t care what colour it is. I’m going to take the big ball.” He wanted the big ball because it’s best for playing soccer.

Upstairs in Cathy’s room, Mark can’t see anything because the light is out. He finds his way to the bed, he feels around until he finds the biggest one [dramatize]. He can’t tell that it’s the yellow ball. He picks it up and starts to leave Cathy’s room. [note: the questions are asked while Mark is still in the dark]

**Questions:**

**Comprehension:**
Did Mark take the big ball? [correct answer: yes] – transparent context
Did he take the yellow ball? [yes] – transparent context
Did Cathy know that her room was dark? [yes]

**Referential Opacity:**
Did Mark know that he took the big ball? [yes]
Did he know that he took the yellow ball? [no]

**Quotational Opacity:**
Remember what Mark said to Cathy: He said, “I’m going to take ...”
Did Mark say “I’m going to take the big ball”? [yes]
Did he say “I’m going to take the yellow ball”? [no]

As with Study 1, the comprehension questions were always asked before the opacity questions (though not always in the order they appear in above). In order to proceed to the opacity questions, each child was required to answer the transparent

---

3 The same sample story that was presented in Study 1 is shown here to make clear the changes from the last study to this one.
4 Note: The questions did not appear in this order.
questions correctly. If a child was not able to do so for a given story, then the key points
of the story were repeated and the child was asked them again. If a child was still not
able to answer those two questions correctly, she was given some stickers and returned
to her classroom. The reason for excluding children who could not correctly answer the
transparent questions was the same as in Study 1; namely, this was done to ensure that
children were aware that in taking the big ball, Mark took the yellow ball. Without this
understanding, the child's answers to the opaque questions would not make sense. For
example, if a participating child did not know that Mark had taken the 'yellow ball', her
answer to the question "Did Mark know that he took the yellow ball?" would be
difficult to interpret. The purpose of the third comprehension question was to have a
question regarding the story content that had an embedded mental verb for which the
correct answer was 'yes'. The purpose of such a question was to make sure that children
didn't think that every question with an embedded mental verb required a 'no' answer.6
Also as mentioned in Chapter Five, this question is not very important because of the
way in which children's answers are scored.

As with Study 1, in order to be scored correct on a given story, for a given type of
opaque context, children had to get both questions for that type of opacity correct. For
example, in the story above, children had to be able to say both that Mark knew that he
took the big ball ('yes'), and that he didn't know that he took the yellow ball ('no'). Thus,
in order to be given credit for being correct on a particular type of opacity (for a given
story), children had to allow that the story character had a particular belief about (said
something about) the object under the known description, but not under another,
equally legitimate, though unknown description of the same referent. I believe that this
scoring method provides a superior measure of children's ability to deal with opaque

---

5 Three children were excluded for failing to correctly answer the transparent questions.
6 Note: this question has been corrected for the error that was mentioned in Chapter 5.
contexts than just being able to answer 'no' to the question with the substituted description alone; it is sensitive to the participant's ability to distinguish between the description the character had knowledge of from the one that was unknown to him or her.

Children's scores on the opacity questions were summed to yield composite scores. Thus, each child earned a score out of 3 for each type of opacity. These summed scores were used in all analyses.

The reason for requiring children to answer three questions for each type of opacity was to provide a fairly sensitive measure of each child's ability; composite scores can be used in some analyses where dichotomous data cannot (e.g., correlations). It was therefore possible to determine the relation among performance on both types of opacity and the other measures of interest. It is to those measures that I will now turn.

6.4.2 Memory measure. The memory measure was the same as the one used in Study 1 (the digit span subtest of Stanford-Binet Test of Intelligence). Children were required to remember a series of single digit numbers and repeat them back to the tester in either the order that they heard them (digit forward), or in the order opposite to the way in which they heard them (digit backward). Children were administered items until they reaching ceiling (3 or more wrong out of 4 on two consecutive pairs of items). Children's raw scores, which consisted of the sum of their digit forward and digit backward scores, were used in subsequent analyses.

This measure provided an indication of each child's memory span which was then partialled out of correlational analyses (to see what results remained). Therefore, it could be determined whether the results were due to something more than just being able to remember the stories.

6.4.3 Language measure. The language measure was the same as that used in Study 1 (the vocabulary subtest from the Stanford-Binet Test of Intelligence). Recall that this is
a standardized measure that requires children to name objects and define terms. For example, one item consists of a picture of a lamb, and the child has to be able to label it as a lamb or a sheep. The task was administered to each child until she reached ceiling (3 or more incorrect answers out of 4 on two consecutive pairs of items) and their raw score was used in the analyses.

As with Study 1, this measure was used to determine the extent to which vocabulary performance is related to performance on the other tasks, and so that performance on this task could be partialled out to see what relations remain among the key tasks.

6.4.4 False Belief Tasks. The false belief tasks used in this study are the same as those used in Study 1, specifically: (1) an unexpected contents task using a Crayola box filled with stickers (self/other); (2) a change of location task in which one story character moves another's toy from a red container to a yellow one while the first character is away (other); and (3) an appearance/reality task using a rock that looks like a sponge (self/other/appearance-reality) for a possible total of 6 points (each child’s score out of 6 was used as her False belief score in all analyses). Children were also administered a true belief task (which all of them passed). See Appendix B for the exact tasks.

6.4.5 Say/Refer Task. It is fairly clear that children have little or no trouble calling one thing by two names/descriptions (e.g., Apperly & Robinson, 1998; Kamawar & Olson, 1999); it is making judgments about those names/descriptions that is hard (David Olson, personal communication). It is this very type of metalinguistic ability that this metalinguistic task was designed to measure; in other words, this task measured children's ability to differentiate between what was said and what was referred to by some name or description.

In this task, participants watched two characters (Mark and Anna) play a
matching game in which the characters sat on opposite sides of an opaque wall where they could not 'see' the other's cards. Both characters had identical sets of picture cards laid out in front of them. Participants were positioned such that they could see both characters and their cards. In the game, each character chose a card and placed it next to him/herself (with a little help from the experimenter). They would then say out loud which card they had chosen (again, with help from the experimenter). The participating child's tasks were: (1) to say whether the two characters had picked the same card; and (2) to make a judgment as to whether a third utterance (given by the experimenter) was identical to the one made by Mark, Anna, both of them, or nobody (given as a forced choice)\(^7\).

Children were first given warm-up trials to ensure that they understood the game and were willing to make each type of judgment. The pictures on the cards that were used for the warm-up portion of this task were a car, a tree, a house and a heart. Before beginning the warm-up trials, each child had to be able to name the depicted items. Notice that for this portion of the task, the name of a depicted item unambiguously identifies a single card (e.g., 'house').

The first two warm-up trials only required children to say whether the cards chosen by the characters were the same or not. For example, in one of the trials, Mark and Anna both picked the tree, placed the picture beside themselves, and said, "I have a tree", and the child was asked, "Do they have the same one?" (correct answer: 'yes'). In the other of the first two trials, the characters had different cards and made different statements (e.g., one picked a car and the other a heart). The child was asked again, "Do they have the same one?" (correct answer: no). Following these two trials, each participant was given three more warm-up trials in which they had to make judgments

\(^7\) The order of the items in the forced choice question varied across participants.
about the characters’ utterances about the cards, in addition to saying whether the cards were the same. For example, participants would see the characters perform the described actions and then hear the following:

Experimenter: "Anna and Mark have each picked a picture and put it in front of them. Now they are going to tell each other which one they have.”
Anna: "I have a tree”
Mark: "I have a heart”
Qa: "Do they have the same one?” [correct answer: no]
Qb: "Who said 'I have a tree'? - was it just Anna, just Mark, Anna and Mark or nobody?” [correct answer: Anna]

If a child gives an incorrect answer, she hears: “Now remember, Anna said, ‘I have a tree’ and Mark said, ‘I have a heart’”
Qb: So, who said ‘I have a tree?’ - was it just Anna, just Mark, Anna and Mark, or nobody? [correct answer: Anna]

The correct answer for Qa was 'no' and for Qb (and Qb’) was 'Anna'. In another of the warm-up trials, the characters chose the same card, and the target utterance matched what both of them had said. So, the correct answers for that trial were 'yes' for Qa and 'Anna & Mark' for Qb. For the third trial, the characters had different cards and the target utterance matched neither of their statements. Therefore, the correct answers were 'no' for Qa, and 'nobody' for Qb. If children answered incorrectly to Qb for a given trial, they heard Qb’ which repeated the characters’ utterances and asked for the answer again. If they were incorrect on a second try for any of the warm-up trials, they were given a final chance with three more warm-up trials (one of each type, as above). If they failed to answer the second set of warm-up trials correctly, they were to be excluded from the study (no children were excluded for this reason).

Notice that for these warm-up trials, the utterance judgment questions (Qb) can be answered correctly by appealing to the chosen card (the referent) without having to consider the descriptions used by the characters. In other words, when asked who said "I have an X”, the child, instead of making a judgment about the characters' utterances,

---

8 Though named "Qa" and "Qb", these key questions did not always appear in this order.
could simply answer with the name of the character(s) who had a card with an 'X' on it. However, the warm-up trials were only designed to ensure that children were both willing and able to give each kind of response to judgment questions. Thus, across the warm-up trials as a whole, children answered questions about whether the two characters' cards matched with both 'yes' and 'no' responses, and answered utterance judgment questions with one response of each type (i.e., just one character, both, and neither). After completing the warm-up trials, children moved on to the test trials.

The cards used for the test phase were: (1) a small, blue flower; (2) a big, red flower; (3) a small, black balloon; (4) a big, yellow balloon; (5) a small, green star; and (6) a big, orange star. Notice that for the test trials, the cards could be described along two dimensions (size and colour) for each kind of thing depicted; for example, the stars could be contrasted by saying whether they were big/small, and by whether they were green/orange. Therefore, simply naming the item depicted would not unambiguously select a card; for example, 'star' referred to both the 'big/orange star' and the 'small/green star'. Further, there were two unambiguous descriptions available for each card; for example, the descriptors 'big star' and 'orange star' both referred to a single card. In other words, each card had two partial, but unambiguous, descriptors.

Prior to beginning the test trials, each child had to be able to name the items depicted, the colours of the items, and be able to point to the big/small example of each item (e.g., "Point to the big star/flower/balloon"). This was done to make sure that errors could not be attributed to a lack of knowledge regarding the colours, sizes or types of items depicted.

The test trials were similar in method to the warm-up trials. Each character picked a card and placed it next to him/herself prior to announcing what card had been chosen. After that was done, children answered questions about the cards and the characters' utterances.
The test trials differed from the warm-up trials as to the types of descriptions used (i.e., in the relation among the various referring utterances and their referents). Unlike the warm-up trials, the test trials had characters refer to the same card using different descriptions. Therefore, in order to succeed on the utterance judgment questions, children could not just look at the cards that were next to the characters; they had to be able to make comparisons between the descriptors contained in the utterances. Another difference between the warm-up and test trials was that for the test trials, children were not given second chances to answer the question; no feedback was given.

There were 12 items in the test phase (see Appendix G for exact trials). In three of them, the characters did not have the same cards as each other (i.e., the answer to Qa was 'yes' for 9/12 items and 'no' for 3/12 items). For four of them, the correct answer to Qb was one of the two characters (two 'Anna' and two 'Mark'). For example:

Mark: "I have an orange star"
Anna: "I have a big star"
Qa: "Do they have the same one?" [correct answer: yes]
Qb: Who said "I have an orange star"? - was it just Anna, just Mark, Anna and Mark, or nobody? [correct answer: Mark]

In three of the 12 trials, the correct answer to Qb was 'both'. For example:

Anna: "I have a small balloon"
Mark: "I have a small balloon"
Qa: "Do they have the same one?" [correct answer: yes]
Qb: Who said "I have a small balloon"? - was it just Anna, just Mark, Anna and Mark, or nobody? [correct answer: Anna and Mark]

For five of the 12 trials, the correct answer was 'nobody'. For two of these trials, the target utterance (the third one) unambiguously referred to the same card as both characters, but did so using more information. For example:

Anna: "I have a big star"
Mark: "I have an orange star"
Qa: "Do they have the same one?" [correct answer: yes]
Qb: Who said, "I have a big, orange star"? - was it just Anna, just Mark, Anna and Mark, or nobody? [correct answer: nobody]
In another two of the five trials in which the correct answer to Qb was 'nobody', the target utterance unambiguously referred to the same card as the characters' utterances, but did so using less information. For example:

Mark: "I have a big, yellow balloon"
Anna: "I have a big, yellow balloon"
Qa: "Do they have the same one?" [correct answer: yes]
Qb: Who said, "I have a yellow balloon"—was it just Anna, just Mark, Anna and Mark, or nobody? [correct answer: nobody]

In the fifth 'nobody' trial, the two characters had different cards and the target utterance was coreferential (but not the same as) only one of the characters' utterances.

Of the 12 test trials, there were 5 for which correct answers did not necessarily indicate that the participant was paying attention to the exact referring description used. In 3 of them, both Mark and Anna had the same card and said the same thing. So, in answering Qb, the child who said that they had both said the same thing may have been attending to the utterances, but she may also have answered correctly by just paying attention to the cards next to the characters. In the other 2 of the 5, the target utterance matched only one of the characters, but since Mark and Anna had different cards, correct answers did not offer clear indication that the participants were attending to the utterances and not the cards (see Appendix G for all trials). Therefore, children earned a score out of 7 for their performance on the test trials that could only be answered correctly by attending to the descriptors contained in the utterances (this score will be used in all analyses).

Children’s scores on the utterance judgment questions (type Qb) were taken to reflect their ability to pay attention to the descriptions used to refer to the cards independently of their referents. In other words, children’s scores on this task reflect their ability to distinguish between which utterances said the same thing from those that said something about the same thing; in other words, this measure provided an indication
of children's metalinguistic awareness.

Though similar, this measure should not be considered simply to be a new kind of opacity task. Unlike opacity tasks, this measure does not require: (1) truth/falsity judgments; (2) the ability to deal with partial knowledge (both of the characters have access to the full descriptions of each referent); and (3) the ability to deal with mental verbs (only the transparent verbs 'pick' and 'have' are used). However, say/refer and opacity tasks are similar in that they require children to distinguish between what was referred to from the way in which it was referred. In other words, the say/refer task taps into the type of metalinguistic awareness necessary for success on opaque contexts by measuring the ability to treat a name/description as an entity which can be judged separately from that to which it refers. Such an ability is fundamental for dealing with opacity.

6.4.6 Say/Mean Task. The other metalinguistic tasks used in this study were the say/mean tasks, which were closely modeled on those used by Robinson, Goelman & Olson (1983) and Apperly and Robinson (1999). Like the say/refer task, these tasks measure children's ability to distinguish between a referent (what was meant) and a description of the referent (what was said). However, unlike the say/refer task, ambiguous descriptors are used. These tasks were originally designed to indicate metalinguistic knowledge of the concepts of say and mean. As mentioned in Chapter Three, the say/mean tasks are measures of metalinguistic ability because they require the child to make comparisons between ambiguous and unambiguous referring expressions. In order to succeed on the say/mean tasks, children must be able to treat descriptions as objects in their own right, and base their judgments/evaluations of ambiguity on them.

In this task, children played a 'matching game' with a puppet ("Leo" – operated by the experimenter) in which the participant and the puppet sat on opposite sides of
an opaque partition. Both the participant and the puppet had before them identical sets of 8 cards: a big red flower; a big blue flower; a small red flower; a small blue flower; a big yellow star; a big black star; a small yellow star; and a small black star. Before beginning, each child was required to name the item depicted on each card and its colour, and to be able to point to the big/small example of each item (e.g., "Point to the big star/flower"). It was explained that the child's task was to select from her array the card described by Leo, to "hold it up high", to compare her card with the one Leo held up, and to say whether or not she had chosen the 'right one'.

Before beginning the test trials, each child had two warm-up trials (with feedback) to make sure they understood the rules of the game. In both of these trials, Leo provided an unambiguous description of a card, the child held it up, and then made a judgment regarding whether she had chosen the correct one. All children found these trials very easy.

On some of the items in the test phase, the description given by Leo unambiguously described a card (e.g., 'the big, red flower') and other times it did not (e.g., 'the red flower'). The ambiguous messages were interspersed among unambiguous messages. On the occasions when Leo provided an ambiguous message (e.g., 'the red flower'), Leo always selected the card that the child did not select so as to result in a mismatch between his and the child's cards (this mismatch is known as a 'communication failure'); for example, if the child chose the 'big, red flower', Leo chose the 'small, red flower'. After the child judged that she had chosen the wrong card, she was asked a question about what Leo had said. For 3 of the communication failures, children were required to make to make a judgment about what Leo said. For another 3 of them, they were required to evaluate the message quality of Leo's utterance. I will

---

9 Leo always provided the description.
address each type in turn.

For the say/mean judgment tasks, with 'red flower' as an example of an ambiguous message resulting in a communication failure, children answered the following type of judgment question on three separate trials: "Did Leo say 'the small, red flower'?" (item (i), disambiguated version of original message). For two other communication failures, they heard the following types of questions once each:

ii) "Did Leo say the red flower" (verbatim version)
iii) "Did Leo say the blue flower" (incorrect version)

If the children correctly answered 'no' to items (i) and (iii), they were then asked what the puppet did say. In order to be scored correct, children had to respond 'no' to the first question and respond with 'red flower' (or whatever was said for a given trial) to the second. Children's scores were summed for the three type (i) trials to provide a score out of 3 for the say/mean judgment task. Items (ii) and (iii) were included as they were a part of the original study, but their scores were not used in subsequent analyses (children were virtually at ceiling on these questions).

On three other trials that resulted in a communication failure, follow-up questions were asked in which children had to evaluate the quality of Leo's message. The questions were drawn from Robinson et. al (1983) and are very similar to those in Apperly and Robinson (1999). Specifically, they asked:

i) "Did Leo tell you enough about his card? Could he have helped you any more?"
if child answered correctly ('no' to the first question or 'yes' to the second), then the following question was asked:
ii) "What should Leo have said?"

To be scored correct on these evaluation trials, children had to answer 'no' to the first question and respond to the second question with an unambiguous description of the card Leo was holding. Children's scores on these three trials were summed to provide their score out of 3 for the say/mean evaluation task (see Appendix H for all trials).
6.5 Procedure and Design

Each child was interviewed individually in two sessions separated by at least one day, and not more than five. The procedure was the same for all participants; they were introduced to the researcher, given a brief introduction as to what was happening (they were informed that they were going to hear stories and by doing that, would be helping me with my 'homework'). They were all told that if they wanted to, they could go back to class, and that no one would be upset if they didn’t want to hear any 'stories'. At the end of each session, each child was given some stickers or a colourful pencil (their choice) and then returned to their rooms.

During one of the two sessions, children received the three stories (and related questions), the digit span task and the say/refer task. During the other session, children received the say/mean task, the false belief tasks, and the vocabulary measure. Each session lasted approximately 15 minutes. The two sessions, the key questions within the opacity task and the say/refer task, and the measures were ordered differently across participants.

Three children were excluded from participating in this study for not being able to correctly answer transparent context questions on their second attempt\(^{10}\). Of the 64 children who participated, 62 were correct on the first attempt across the three stories and 2 were correct on two of the three stories (mean number correct on the first attempt is 2.97/3). The results reported below are for the 64 children who, by at least their second attempt, were able to correctly answer the transparent context questions.

6.6 Results\(^{11}\)

\(^{10}\) Remember that if children answered transparent context questions incorrectly, they were re-read key points of the story and asked the transparent questions again.

\(^{11}\) As with Study 1, all significant results will be reported. Non-significant results of interest will be reported as well.
6.6.1 Methods of Analysis. The data were entered on a Macintosh computer using SPSS (Release 6.0). The primary statistical tests used were ANOVAs, hierarchical regression and partial correlations. Participants' performance was analyzed using their raw/summed scores on each task/measure. Children's ages were measured in months for correlational and regression purposes, and coded into 3 groups (5, 6 and 7 years) for use in the ANOVAs.

A repeated measures ANOVA was performed with Opacity (2 levels) as the within subject factor, and sex of the participant (2 levels), session order (2 levels), the order of the key questions for the opacity tasks (3 levels) and the order of the key questions for say/refer task (2 levels) as between subject factors. The results showed no main effects for sex, session order and forms (order of key questions). Further, no interaction effects were found for these variables. Therefore, orders, forms and sexes were combined for further analysis.

6.6.2 ANOVA Results. The first analysis for this second study was a repeated measures ANOVA with Opacity entered as the within subject factor. The Opacity factor had two levels (referential and quotational). Age (coded into 5, 6 or 7 years) and ESL status (non-ESL or ESL) were entered as the between subject factors (see Table 11 for a summary of performance on the within subject factor summed across age).

The ANOVA\(^\text{12}\) was performed using a planned Helmert contrast for the Age factor. Recall that a Helmert contrast compares performance on each level in the factor to the averaged performance on the remaining levels (the experimenter determines the orders of the levels within the factor). The first contrast was between the youngest age group (5-year-olds) and the averaged performance of the two oldest age groups (6- and

\(^{12}\) In regard to the assumption of homogeneity of variance, ANOVA\(^*\) are considered conditionally robust if the group sizes (at each level) are equal or approximately equal (Stevens, 1996), as is the case in this study.
7-year-olds), and the second contrast between the two oldest age groups (6- vs. 7-year-olds).

**Age.** The results from this ANOVA established that children improve significantly with age ($F(2, 58) = 22.50, p < .001$), and the planned contrast revealed that there is a significant difference between the youngest group and the average of the two oldest groups ($F(1, 58) = 37.36, p < .001$), and between the two oldest groups ($F(1, 58) = 7.63, p < .01$). See Table 11 for mean scores by age and Figure 8 for performance on opacity types by age.
Table 11.

Descriptive Statistics for Opacity and Metalinguistic Tasks for Study 2.

<table>
<thead>
<tr>
<th>Measure</th>
<th>All Participants (N=64)</th>
<th>5-year-olds (n=22)</th>
<th>6-year-olds (n=21)</th>
<th>7-year-olds (n=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Referential Opacity*</td>
<td>1.33</td>
<td>1.29</td>
<td>.41</td>
<td>.67</td>
</tr>
<tr>
<td>Quotational Opacity*</td>
<td>1.33</td>
<td>1.20</td>
<td>.55</td>
<td>.86</td>
</tr>
<tr>
<td>Say/Refer*</td>
<td>2.36</td>
<td>1.80</td>
<td>1.23</td>
<td>.97</td>
</tr>
<tr>
<td>Say/Mean – Judgment*</td>
<td>1.64</td>
<td>1.28</td>
<td>.95</td>
<td>1.25</td>
</tr>
<tr>
<td>Say/Mean – Evaluation*</td>
<td>1.30</td>
<td>1.29</td>
<td>.68</td>
<td>1.21</td>
</tr>
</tbody>
</table>

* out of 3.  b out of 7.
Figure 8. Mean scores (+SE) on referential and quotational opacity by age.
**ESL Status.** There was no main effect for ESL status.

**Opacity.** There was a no effect for Opacity ($F (2, 58) = .00, \text{ns}$) indicating that performance was the same for both types of opaque contexts.

**Interaction Effects.** There were no significant interaction effects.

6.6.3 *Hierarchical Regressions.* Hierarchical regressions were performed with referential opacity as the dependent variable and age (in months), vocabulary score, digit span, false belief score and metalinguistic awareness as the independent variables. In order to determine which metalinguistic awareness measures were able to predict performance on the referential opacity tasks, a hierarchical regression was performed using each of them: (i) say/refer; (ii) say/mean – judgment; and (iii) say/mean – evaluation. Age, vocabulary and digit span were entered first in each regression to ensure that the contributions made by the remaining factors were independent of age, general language ability and short-term memory. These three factors were entered in a single step (as a block) because there was no theoretical reason for entering them separately\(^\text{13}\). False belief was entered in the second step because of the prediction that false belief scores could account for variance in referential opacity performance even after the effects from age, vocabulary and digit span were removed. I will report each in turn.

**Say/Refer.** Table 12 shows the results of this first analysis (model – a). First, in predicting referential opacity performance, after step 1 with age, vocabulary and digit span in the equation, $R^2 = .36, F (3, 60) = 11.2, p < .001$. After the addition of false belief in step 2, $R^2$ change = .06, $F_{\text{inc}} (4, 59) = 6.4, p < .05$. That is, the addition of false belief scores reliably improves $R^2$. After the addition of say/refer in step 3, $R^2$ change = .12, $F_{\text{inc}} (5, 58) = 14.89, p < .001$. That is, the addition of say/refer scores reliably improves $R^2$.

\(^{13}\) Recall that these measures were included in the study in order to partial out their effects.
A second regression was performed to see whether false belief would still account for some of the variance in referential opacity performance when entered as the last variable. Table 12 also shows the results of this analysis (model – b). In the second hierarchical regression using say/refer as the metalinguistic awareness measure, say/refer scores were entered in the second step and false belief in the third. As before, in predicting referential opacity performance, after step 1 with age, vocabulary and digit span in the equation, $R^2 = .36$, $F(3, 60) = 11.2, p < .001$. After the addition of say/refer in step 2, $R^2$ change = .15, $F_{inc}(4, 59) = 17.2, p < .001$. That is, the addition of say/refer scores reliably improves $R^2$. After the addition of false belief scores in step 3, $R^2$ change = .04, $F_{inc}(5, 58) = 4.6, p < .05$. That is, the addition of false belief scores reliably improves $R^2$ even after age, vocabulary, digit span, and say/refer have been entered.

It is worth noting that even when taken together, age, vocabulary performance, digit span, false belief and metalinguistic ability (say/refer) were only able to account for about 54% of the variability of performance on referentially opaque contexts.
Table 12.

Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and the say/refer task (N = 64).

<table>
<thead>
<tr>
<th>Model/Step/Variable</th>
<th>B</th>
<th>SE_b</th>
<th>β</th>
<th>R² or R²Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.01</td>
<td>.01</td>
<td>.10</td>
<td>.36**</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.03</td>
<td>.04</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td>.03</td>
<td>.05</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>Model – a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.20</td>
<td>.09</td>
<td>.26</td>
<td>.06*</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>say/refer</td>
<td>.27</td>
<td>.07</td>
<td>.40</td>
<td>.12**</td>
</tr>
<tr>
<td>Model – b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>say/refer</td>
<td>.27</td>
<td>.07</td>
<td>.40</td>
<td>.15**</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.20</td>
<td>.09</td>
<td>.26</td>
<td>.04*</td>
</tr>
</tbody>
</table>

Note. Reported regression coefficients are from final step.

* p < .05,  ** p < .001
Say/Mean – Judgment. Table 13 shows the result from this analysis (model – c).
First, as above, in predicting referential opacity performance, after step 1 with age, vocabulary and digit span in the equation, $R^2 = .36$, $F (3, 60) = 11.2$, $p < .001$, and after addition of false belief in step 2, $R^2$ change = .06, $F_{\text{inc}} (4, 59) = 6.4$, $p < .05$. After the addition of say/mean – judgment in step 3, $R^2$ change = .03, $F_{\text{inc}} (5, 58) = 1.94$, $\text{ns}$ That is, the addition of say/mean - judgment scores did not reliably improve $R^2$.

Another regression analysis was performed to see whether false belief would still account for some of the variance in referential opacity performance when entered as the last variable. Table 13 also shows the results of this analysis (model – d). In this hierarchical regression, the say/mean – judgment score was entered in the second step and false belief in the third. As before, in predicting referential opacity performance, after step 1 with age, vocabulary and digit span in the equation, $R^2 = .36$, $F (3, 60) = 11.2$, $p < .001$. After the addition of say/mean - judgment in step 2, $R^2$ change = .07, $F_{\text{inc}} (4, 59) = 7.6$, $p < .01$. That is, the addition of the say/mean - judgment scores reliably improves $R^2$. After the addition of false belief scores in step 3, $R^2$ change = .02, $F_{\text{inc}} (5, 58) = 2.3$, $\text{ns}$ That is, the addition of false belief scores does not reliably improve $R^2$. 
<table>
<thead>
<tr>
<th>Model/Step/Variable</th>
<th>$B$</th>
<th>$SE_{B}$</th>
<th>$\beta$</th>
<th>$R^2$ or $R^2\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, Vocabulary,</td>
<td>.02</td>
<td>.02</td>
<td>.13</td>
<td>.36***</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.07</td>
<td>.05</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>Model – c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.17</td>
<td>.11</td>
<td>.22</td>
<td>.06*</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Say/Mean – Judg.</td>
<td>.24</td>
<td>.13</td>
<td>.25</td>
<td>.03</td>
</tr>
<tr>
<td>Model – d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Say/Mean – Judg.</td>
<td>.24</td>
<td>.13</td>
<td>.25</td>
<td>.07**</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.17</td>
<td>.11</td>
<td>.22</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note. Reported regression coefficients are from final step.

* $p < .05$, ** $p < .01$, *** $p < .001$
**Say/Mean – Evaluation.** Table 14 shows the result from this analysis (model – e). As before, age, vocabulary and digit span (step 1), and false belief scores (step 2) reliably improve $R^2$. After the addition of say/mean – judgment in step 3, $R^2$ change $= .02$, $F_{inc} (5, 58) = 3.4$, ns That is, addition of say/mean - evaluation scores did not reliably improve $R^2$.

Another regression was performed to see whether false belief would still account for some of the variance in referential opacity performance when entered as the last variable. Table 14 also shows the results of this analysis (model – f). In this hierarchical regression, the say/mean – evaluation score was entered in the second step and false belief in the third. As before, in predicting referential opacity performance, after step 1 with age, vocabulary and digit span in the equation, $R^2 = .36$, $F (3, 60) = 11.2$, $p < .001$. After the addition of say/mean - evaluation in step 2, $R^2$ change $= .05$, $F_{inc} (4, 59) = 4.43$, $p < .05$. That is, the addition of the say/mean - evaluation scores reliably improves $R^2$. After the addition of false belief scores in step 3, $R^2$ change $= .04$, $F_{inc} (5, 58) = 3.83$, $p = .055$. That is, the addition of false belief scores does not reliably improve $R^2$.

Table 15 provides a summary contrasting the results of the hierarchical regressions when each of the metalinguistic awareness scores was entered as the last variable in the regression equation (models a, c and e).
Table 14.

Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and say/mean – evaluation (N = 64).

<table>
<thead>
<tr>
<th>Model/Step/Variable</th>
<th>B</th>
<th>SE_β</th>
<th>β</th>
<th>R² or R²Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.01</td>
<td>.02</td>
<td>.08</td>
<td>.36**</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.04</td>
<td>.05</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td>.07</td>
<td>.06</td>
<td>.19</td>
<td></td>
</tr>
<tr>
<td>Model – e</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.21</td>
<td>.11</td>
<td>.27</td>
<td>.06*</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Say/Mean – Eval.</td>
<td>.16</td>
<td>.11</td>
<td>.16</td>
<td>.02</td>
</tr>
<tr>
<td>Model – f</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Say/Mean – Eval.</td>
<td>.16</td>
<td>.11</td>
<td>.16</td>
<td>.05*</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td>.21</td>
<td>.11</td>
<td>.27</td>
<td>.04</td>
</tr>
</tbody>
</table>

Note. Reported regression coefficients are from final step.

* p < .05, ** p < .001
Table 15.

Summary of hierarchical regression analysis predicting referential opacity from age, vocabulary, digit span, false belief and metalinguistic awareness measures (N = 64).

<table>
<thead>
<tr>
<th>Model/Step/Variable</th>
<th>B</th>
<th>SE_b</th>
<th>β</th>
<th>R² or R²Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>All models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td>.36**</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Belief</td>
<td></td>
<td></td>
<td></td>
<td>.06*</td>
</tr>
<tr>
<td>Model - a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>say/refer</td>
<td>.27</td>
<td>.07</td>
<td>.40</td>
<td>.12**</td>
</tr>
<tr>
<td>Model - c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
</tr>
<tr>
<td>Say/Mean – Judg.</td>
<td>.24</td>
<td>.13</td>
<td>.25</td>
<td>.03</td>
</tr>
<tr>
<td>Model - e</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>Say/Mean – Eval.</td>
<td>.16</td>
<td>.11</td>
<td>.16</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note. Reported regression coefficients are from final step. Coefficients are not included for steps 1 and 2 because they differ (at the final step) for the three models. See Tables 12, 13 and 14 for the coefficients for steps 1 and 2.

* p < .05,  ** p < .001
6.6.4 Correlational Analyses. When performing correlations, age was measured in months, raw scores were used for vocabulary and digit span, and summed scores were used for all other measures (i.e., for both types of opacity, the say/refer task, the say/mean judgment task, and the say/mean evaluation task). Table 16 provides a correlation matrix for all measures/scores with nothing partialled out. As can be seen, every measure/score is significantly correlated to every other measure/score (some at $p < .001$, some at $p < .005$ and two at $p < .05$; all two-tailed). Partial correlations were performed to address specific questions.
Table 16.

Correlation Matrix for all Measures (N = 64).

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Digit Span</th>
<th>FB</th>
<th>Say/Refer</th>
<th>Quot. Opacity</th>
<th>Ref. Opacity</th>
<th>Say/Mean - Eva.</th>
<th>Say/Mean - Judg.</th>
<th>Vocab.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Digit Span</td>
<td>.66***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. FB</td>
<td>.60***</td>
<td>.58***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Say/Refer</td>
<td>.36**</td>
<td>.45***</td>
<td>.42**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Quotational Opacity</td>
<td>.56***</td>
<td>.49***</td>
<td>.51***</td>
<td>.32*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Referential Opacity</td>
<td>.51***</td>
<td>.53***</td>
<td>.58***</td>
<td>.66***</td>
<td>.50***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Say/Mean Evaluation</td>
<td>.40**</td>
<td>.34*</td>
<td>.52***</td>
<td>.39**</td>
<td>.39**</td>
<td>.46***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Say/Mean Judgment</td>
<td>.39**</td>
<td>.44***</td>
<td>.61***</td>
<td>.51***</td>
<td>.41**</td>
<td>.55***</td>
<td>.67***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Vocabulary</td>
<td>.63***</td>
<td>.62***</td>
<td>.56***</td>
<td>.41**</td>
<td>.48***</td>
<td>.52***</td>
<td>.46***</td>
<td>.59***</td>
<td></td>
</tr>
</tbody>
</table>

Note: * p < .05,  ** p < .005,  *** p < .001 (two-tailed)
When age, vocabulary and digit span were partialled out, there was no significant relation between false belief and quotational opacity, as expected (r (59) = .19, ns). However, contrary to expectations, no relation was found between quotational opacity and any of the metalinguistic tasks (i.e., say/refer, say/mean - evaluation, and say/mean - judgment) after age, vocabulary and digit span were partialled out.

A final partial correlation was performed examining the relation between both types of opacity once age, digit span and vocabulary were partialled out. It was found that the correlation between referential and quotational opacity was not significant, contrary to expectations (r (59) = .23, two-tailed, p = .07).

6.6.5 False Belief Tasks. Performance was consistent with other work in this area (see Figure 9). Given that the two older groups of children were virtually at ceiling (4-year-olds mean = 3.32/6; 5-year-olds mean = 5.24, and the 6-year-olds 5.90) the scores were converted into categorical data and Chi-square tests for independence were performed to determine whether there was an age effect. Summed scores were converted into pass/fail scores (i.e., summed scores were converted into dichotomous data). A pass was considered 5 or more out of 614. The results from this analysis revealed an age effect (X² (2, N=64) = 30.8, p<.0001; see Table 17 in Appendix I) thereby supporting the view that children's performance on false belief tasks improves with age.

---

14 Analyses using a stricter pass criterion (i.e., 6/6) yielded the same results.
Figure 2. Mean scores (+SE) on false belief tasks by age.
6.7 Discussion

The results confirmed the hypotheses with a few notable exceptions. The results will be discussed within the framework provided by the Hypotheses and Results sections above.

6.7.1 Age. Not surprisingly, the hypothesis that children’s sensitivity to opacity improves as they get older was supported by the results, which showed an increase in performance for both types of opaque context between the ages of 5 and 7 years. As with Study 1, there is not too much to say about this finding other than that it is consistent with a great deal of developmental research.

6.7.2 Opacity. In answer to the question raised by Hypothesis 2, I did not find a difference in performance between referential and quotational opacity, suggesting that they are not ordered by difficulty and that they share a similar developmental trajectory. However, partial correlations revealed that surprisingly (and contrary to expectations), performance on the two tasks was not related (after the effects of age, vocabulary and digit span were removed). With the same factors partialled out, there was a correlation between referential opacity and metalinguistic awareness (all types), but not between quotational opacity and any of the metalinguistic awareness measures. As expected, performance on referential opacity, but not quotational opacity, was related to false belief performance (after the effects of age, vocabulary and digit span were removed), thus supporting the view that metarepresentational understanding is related to one type of opacity, but not the other.

Taken together, these findings suggest that referential and quotational opacity are similar in difficulty, but for different reasons: while metalinguistic and metarepresentational skills are related to success on referential opacity (which I will discuss in detail shortly), neither is related to performance on quotationally opaque contexts. Study 2, though capable of suggesting what might account for variability on
referential opacity tasks, is not capable of doing the same for quotationally opaque contexts by appealing to other related skills. This topic will have to be the focus of future research.

6.7.3 Predicting performance on referentially opaque contexts. It was no surprise to find that metarepresentational ability (as measured by false belief performance) accounted for a significant portion of the variance in referential opacity task scores, even after age, digit span and vocabulary were partialled out. This finding is identical to that in Study 1 of this thesis, and consistent with other research (Kamawar & Olson, 1999). What is new, and very interesting, is the evidence demonstrating that at least one type of metalinguistic awareness (say/refer) is able to account for a significant proportion of the variance in scores from referentially opaque questions even after the above factors have been entered into the regression equation. Further, even when metarepresentational ability (i.e., false belief scores) is entered into the regression equation as the final step, it is still able to account for a significant portion of the variability in performance on referential opacity tasks, thereby demonstrating that both metarepresentational ability and metalinguistic awareness (say/refer) make significant and independent contributions to children's performance on referentially opaque contexts.

The results support Apperly and Robinson's (1999) findings that performance on referentially opaque contexts is related to performance on the say/mean task (evaluation version). However, these results go much further. While the types of metalinguistic awareness that the say/mean tasks capture were able to predict opacity performance after age, vocabulary and digit span were entered into the regression equation, they were not able to do so after metarepresentational ability was entered into the equation (both say/mean tasks were related to false belief). Therefore, while the say/mean task scores predict referential opacity performance, they are not able to do so
above and beyond false belief tasks. This finding is very similar to that in other studies in showing a strong relation between false belief understanding and metalinguistic ability (e.g., Doherty & Perner, 1998).

The fact that the metalinguistic awareness measures used in this study differed in their ability to predict referential opacity performance warrants further discussion. While all three measures picked up some aspect of children’s metalinguistic awareness, it is clear that they picked up on different aspects of it. While the say/mean tasks required the child to be able to deal with referentially ambiguous descriptions, the say/refer tasks required the child to be able to handle referentially unambiguous, but partial, descriptions. Recall that referentially opaque contexts do not require one to deal with ambiguous descriptions of referents; they require the ability to deal with partial representations, and to realize that knowledge of a referent under one description does not entail knowledge of it under all descriptions (i.e., the ability to deal with partial knowledge).

Stated this way, then, it is not surprising that say/refer performance was a better predictor of referential opacity performance. However, I want to point out once again that the say/refer task is not simply another kind of opacity task; that is not why it shows such a strong relation to referential opacity performance. This metalinguistic task differs from opaque contexts in some significant ways. To reiterate, to succeed on this task, one does not theoretically require the ability to metarepresent because no questions are asked about what someone knows (or believes etc.), nor are any of the questions premised on what someone knows (or believes etc.). Further, everyone involved in the task, namely the participant and the story characters, have complete access to all of the possible descriptions of the referents (i.e., cards), so competence in dealing with partial knowledge is not required either. The main similarity between the two tasks is that they both require the ability to recognize, and deal with, partial (not
ambiguous) descriptions of referents.

Children's ability to deal with referentially opaque contexts has been examined in a number of studies and by a number of researchers (e.g., Apperly & Robinson, 1998; 1999; de Villiers, 1994; Kamawar & Olson, 1999), each developing their own vocabulary and way of talking about the skills that must be in place for successful performance. It has been discussed in terms of: (1) the ability to deal with the syntactic structure of complementation, with that as the link between performance on false belief tasks and referentially opaque contexts (de Villiers, 1994); (2) the ability to deal with metarepresentation, with that as the link between performance on false belief tasks and referentially opaque contexts (Kamawar & Olson, 1999); (3) the ability to deal with partial knowledge (Apperly & Robinson, 1999; Kamawar & Homer, in press; this thesis); and (4) metalinguistic awareness, though it is not always called such (Apperly & Robinson, 1999; this thesis).

I believe that this study has gone a long way toward showing that the answer to the question of what predicts performance on referential opacity tasks is not limited to just one of the above options; it may in fact prove to be the case that all of them make independent and significant contributions to sensitivity to opacity. However, this study is only able to demonstrate that metarepresentational and metalinguistic ability do so. Examining the role that the other abilities play in predicting performance on opacity tasks will have to be addressed by future research.

6.8 Summary and Conclusion

Taken as a whole, the results are supportive of the hypotheses presented in Section 6.2 above. As with the previous study, I will summarize the main findings by revisiting the questions that provided the framework for this study. I will discuss the questions in order of increasing importance.

The second question was "What does the developmental trajectory for
performance on quotationally opaque contexts look like and how is it related to referential opacity?". The answer to this question was that it is quite similar to the developmental trajectory of referential opacity. Surprisingly however, the results did not demonstrate a relation between the two tasks, leading me to conclude that the similarity in difficulty is not due to an underlying similarity in the skills required for successful performance.

The first, and far more interesting, question was "In addition to metarepresentational ability, does metalinguistic ability predict performance on referentially opaque contexts?". The data are clear in that the answer is 'yes'. Not only must one be able to represent another's knowledge (belief, etc.), one must also be able to treat the proposition expressing that belief as an object of thought. Indeed, it may be the case that in thinking of another's knowledge, one must represent it in a way much like imagining what she would say. For example, in representing what Mark (in the sample story given earlier) would say about which ball he has in his hands (while still in the dark room), not only must one represent Mark's knowledge of the referent as 'the big ball', if asked what Mark would say about the object in his hands, one must be able to say that Mark would describe it as 'the big ball'. In other words, one must be aware of the linguistic expression Mark would use to characterize the referent in question.

The way in which these results should be considered most significant is that they demonstrate that success on referentially opaque contexts is not explained by any single factor. Taken together, even age, vocabulary performance, digit span, false belief and metalinguistic ability (say/refer) were not able to account for all of the variability in performance on opacity tasks. It looks like there's a need for more work to be done.
Chapter 7. General Discussion and Conclusions

The main objective of this thesis was to investigate children's developing proficiency with a variety of opaque contexts with a focus on: (a) exploring their ability to deal with a variety of opacities (i.e., referential, linguistic, intentional, and quotational opacity); and (b) explaining some of the variability in performance on referential opacity tasks by appeal to metarepresentational ability and metalinguistic awareness. In Chapters 5 and 6, children's performance was discussed in terms of the relations among the tasks and other measures (e.g., vocabulary and digit span). However, it is now time to take a step back and discuss children's performance on the opacity, metarepresentational, and metalinguistic tasks more conceptually.

7.1 Study 1

As the title of this thesis indicates, I was interested in examining the development of children's understanding of the opaque uses of language. While other studies have confined their research to contexts that are opaque because they contain mental attitudes directed toward propositions (e.g., Apperly & Robinson, 1998; de Villiers, 1994; Kamawar & Olson, in press), the first study of this thesis set out to explore children's performance across a range of opaque contexts. Though different, the three contexts examined in this first study met the definition of an opaque context: A context is opaque if replacing a term within it with a coreferential term can affect the truth value of the whole sentence (Hookway, 1988).
To review, Study 1 examined 4- to 8-year-olds' performance on referentially opaque contexts (opaque as a result of a mental attitudes directed toward a proposition), linguistically opaque contexts (opaque because of indirect quotation containing a proposition), intentionally opaque contexts (opaque because of the explicit mention of an intention), a digit span task, false belief task, and a vocabulary measure. To recap, the main findings from this study are that children's performance on the three opacity tasks and the false belief tasks differs significantly, with false belief easiest and intentional opacity hardest. Interestingly, the results also reveal that the three opacity tasks were still related to one another even after age, digit span, and vocabulary were partialled out. The results of this first study also support the view that children found opaque contexts with action referents significantly easier than those with object referents.

This first study can be best seen as providing new insight into children's developing proficiency with opacity across a variety of contexts. The results are significant in that they support both the view that there are some fundamental similarities across opaque contexts, and the view that different types of opacity vary in complexity. I contend that this commonality stems from the fact that in order to succeed on these three types of opacity, children need to realize that access to an object (or action) under one description does not ensure access to it under all, and that therefore, one must pay attention to the particular description used. I also contend that this commonality stems from the fact that children, when faced with an opaque context, need to be able to reflect on and evaluate the referring expression as an object of thought (what I refer to as metalinguistic awareness). However, since I did not examine the role of metalinguistic awareness in this first study, this remains an open empirical

---

1 At least for the three types measured.
question, and as such, suggests a useful follow-up study.

So while there exists a commonality in performance across opacity types, there are also exist differences. However, the exact nature of what it is that makes one type of opaque context different from another is still not completely clear. It may prove to be the case that the three types of opacity examined vary along a single dimension (i.e., there is an additive relation from the easiest to the hardest), or, more likely, that they have specific requirements that differ (e.g., memory for exact words versus competence with sophisticated verbs of intention). Therefore, the topic of future research may be to try to specify which of the two previous options is the appropriate one. For example, in Chapter 5, I hypothesized that the differences between referential and intentional opacity are due to the fact that children acquire more sophisticated ways of talking about intentions later than they acquire ways of talking about knowledge. If this view is correct, then only children who are sensitive to intentional opacity should be able to use and understand a wide range of terms that express intention (e.g., 'plan to' and 'intended to'), or vice versa. Conversely, all children who understand intentional opacity should succeed on referential opacity.

This study can be seen as examining some of the more subtle and sophisticated ways in which language is connected to theory of mind development. Astington (in press) writes of theory of mind that it “underlies [children’s] ability to attribute beliefs, desires, intentions, and emotions to the self and others in order to explain and predict behavior” (p. 9). This study can be seen as examining theory of mind and language development because: (1) it examines two verbs that refer to mental states that are important to theory of mind understanding (i.e., knowing and intending); and (2) it examines children’s sensitivity to how those verbs constrain the changes that can be made to the propositions that follow them when a propositional attitude is attributed to someone. So, while children as young as 4 years of age have some understanding of
these mental verbs (as evidenced by false belief tasks using the terms 'know' and 'think'), and can perhaps use them correctly in particular situations, they still lack a complete understanding of those terms, an understanding that would reveal the verbs as operating in very specific, and constrained ways. In other words, this study demonstrates that aspects of theory of mind development continue well beyond the age of 4 years. While this is not new information to the field of theory of mind research, it is a new way in which to demonstrate later development of a theory of mind.

In summary, this first study demonstrates that sensitivity to opaque contexts varies for the specific contexts examined. The way in which these results should be considered most significant is that they demonstrate that children's understanding of opacity is not all-or-nothing; it develops over time and varies across situations.

7.2 Study 2

I believe Study 2 to be the more informative, and important, of the two studies conducted for this thesis. While Study 1 was rather exploratory in nature, Study 2 had a more specific goal: to test the hypothesis that performance on referentially opaque questions is related to performance on metarepresentational and metalinguistic measures. Though not related to the primary goal of the second study, quotationally opaque contexts were included in order to determine the developmental trajectory of another type of opaque context often discussed by philosophers of language (quotational opacity). I have discussed each of these issues in Chapter 6, but I will now discuss the above hypothesis, the findings, and my explanation of the findings in more detail.

I argued in Chapter 3 that the ability to deal successfully with referentially opaque contexts is related to three other abilities: (1) to represent another's representations; (2) to understand that someone can have knowledge of an object (referent) under one description without necessarily having knowledge of it under all
allowable descriptions (i.e., understand that partial knowledge is possible); and (3) to treat a referring expression as an object of thought, that is, to recognize that when attributing a belief or intention toward a referent to another person, it is not sufficient that the description makes reference to the right thing; what is important is that the description makes reference in the right way (i.e., by using a description that is known to the person to whom the belief, intention, etc., is attributed). The above argument was first laid out in Chapter 3 (Cognitive Requirements), but will be reviewed again in light of the results.

The main finding of Study 2, I believe, is very clear in supporting the view that being able to think about another's (partial) representations and being able to make language an object of thought are both related to being able to deal with referential opacity².

Before getting into why I think that both thinking about thoughts and thinking about the words used to refer are instrumental to sensitivity to referential opacity, I want to make certain distinctions among some related terms. The first term that needs to be clarified is 'metacognition'. Though I have not used the term in this thesis, I think it is important to spell it out so that it can be contrasted with other important terms (i.e., metarepresentation and metalinguistic). Metacognition has been defined as "any knowledge or cognitive process that refers to, monitors, or controls any aspect of cognition" (Moses & Baird, 1999, p. 533). Moses and Baird go on to say that

Metacognition is now seen as a central contributor to many aspects of cognition, including memory, attention, communication, problem solving, and intelligence, … In this sense at least, metacognition is a domain-general facet of cognition (p. 533).

This definition highlights the fact that the term 'metacognition' is domain-general and

² Though not tested directly, given the nature of opacity tasks, children who succeeded must have been able to maintain multiple, distinct representations of the referents.
can therefore be applied to a very wide range of processes or types of knowledge. To say that some process requires metacognitive ability is to say that it requires the ability to reflect or treat some aspect (any aspect) of cognition as an object of thought; for example, metamemory would be an example of metacognition. Certainly, under this broad description, the key tasks used in this study (opacity, say/refer, say/mean and false belief tasks) can be considered metacognitive tasks.

However, I think that there are more specific ways in which to talk about the metacognitive skills that are related to successful completion of referential opacity tasks. I believe that by using more precise terms, I can be clearer about what I think is going on in the mind of the young child who is/is not successful on referentially opaque contexts.

The first term I want to discuss (really, to re-examine), is 'metarepresentation'. Sperber (1999) provides this useful description/definition of metarepresentation:

> Cognitive systems are characterized by their ability to construct and process representations of objects and states of affairs. Mental representations and public representations such as linguistic utterances are themselves objects in the world, and therefore potential objects of second-order representations, or "metarepresentations". (p. 541)

Under this description, it seems that 'metarepresentation' is a specific type of 'metacognition': metarepresentation meets the definition of metacognition, but goes further by pertaining to internal (mental) and linguistic representations (but not to other aspects of cognition, like memory; see Moses & Baird, 1999 quote above). Thus, I want to say that the tasks, while requiring metacognitive skills, can be more precisely said to involve metarepresentational ones. However, I can be even more precise if I turn to the term 'metalinguistic'.

Sperber (1999), while defining 'metarepresentation', alluded to what it means to be 'metalinguistic':

> It has long been observed that human languages have the semantic and
syntactic resources to serve as metalanguages. In direct and indirect quotations, utterances and meanings are metarepresented. The study of such *metalinguistic* devices has been developed in semiotics, ... in philosophy of language, and in pragmatics. (p. 541, my italics)

Extrapolating only slightly from Sperber's (1999) use of the term 'metalinguistic', I believe that it is possible to consider metalinguistic knowledge (sometimes referred to as awareness) to be a sub-type of metarepresentational knowledge; that is, the term 'metalinguistic' specifically pertains to the metarepresentation of language. There does not, however, seem to be a corresponding term for metarepresentation that specifically pertains to *mental representations*. For many, the term 'metarepresentation' itself is used to refer to representations of mental representations, such as beliefs about beliefs (e.g., Leslie, 1988 as cited in Astington, 1993), or to refer to "the mind’s activity in forming beliefs and other mental states" (Astington, 1993, p. 58; describing Perner, 1991) without any mention of the term applying to public (linguistic) representations. Therefore, for the rest of this thesis, I will use the terms 'metacognitive', 'metarepresentation' and 'metalinguistic' as outlined above, and will use the term 'mental–metarepresentation' (metarepresentation for short) to refer to metarepresentation that is specific to metarepresentations about *mental representations* (as opposed to public ones like language). The term metarepresentation is necessary because the term 'metarepresentation' is often used to mean only 'mental–metarepresentation'. Even Sperber (1999) blurs the distinction somewhat when he notes that

much attention has been paid to different degrees of metarepresentational competence that may be involved in attributing mental states to others. In particular, the ability to attribute *false* beliefs is seen as a sufficient, if not necessary, proof of basic metarepresentational competence (p. 541)

It seems that we need to come to some consensus on how to use the term 'metarepresentation'.

There also is no clear consensus as to what constitutes metalinguistic awareness; in fact, there is quite a range as to what is taken to count as metalinguistic ability. In the
strictest philosophical sense, metalinguistic ability is the ability to use a language to talk about another language, which is also known as the object language (Bach, 1995). However, within developmental psychology, the definition of metalinguistic awareness is sometimes seen as paying "attention to language forms and the manipulation of linguistic units" (Bialystok, 1993, p. 213). At other times, metalinguistic awareness is seen as the ability to manipulate language structures, such as being able to delete phonemes, re-order words, delete words or recognize rhyming (e.g., Bialystok & Ryan, 1985a; 1985b; Bryant & Bradley, 1985; all cited in Garton & Pratt, 1989). Garton and Pratt (1989) point out that in talking about metalinguistic awareness, it is possible to talk about it at different levels, such as

phonological awareness, word awareness, syntactic awareness, and pragmatic awareness [and that] in considering these aspects of language, it will become clear that although the knowledge involved is different in each case, each involves increasing control over reflecting upon language. (p. 151)

Therefore, when using the term 'metalinguistic', I will have to be very precise as to what I mean when I use the term. I believe my use of the term to be consistent with the definitions that I presented in Chapter 3, namely, that offered by Cazden (1974, cited in Bialystok, 1993) who defined it as "[the] ability to make language forms opaque, and attend to them in and for themselves ..." and that offered by Tunmer and Herriman (1984) who define metalinguistic awareness as "the ability to reflect upon and manipulate the structural features of spoken language, treating language itself as an object of thought" (p. 12, my italics).

So, at the highest level of abstraction, this study can be viewed as examining the interrelations among metacognitive tasks. At a more fine-grained level, this study can be viewed as examining the interrelations among metarepresentational tasks. At a yet finer-grained level of description, this study can be viewed as examining the interrelations among metarepresentational and metalinguistic tasks.
It seems clear that the key tasks make up a single domain of understanding; that is, they all require metarepresentational understanding, and more broadly, metacognitive understanding. Stated this way then, it is no surprise that performance among the tasks is generally correlated.

With the terms now laid out (at least within the context of this thesis), the hypothesis that was the driving force behind Study 2 can be rephrased as: Performance on referentially opaque questions is related to performance on metarepresentational and metalinguistic measures.

Recall that the main finding of this second study is that even after the variability from children's age (measured in months), their vocabulary scores and their digit span tasks are taken into account, metarepresentational ability (false belief performance) still accounts for a significant amount of the remaining variability on referential opacity tasks. Further, even after all of the above variables are entered into the regression equation, metalinguistic awareness (specifically, the say/refer task) still significantly explains some of the remaining variability. Further, when performance on the say/refer task is entered in the second step (after age, digit span and vocabulary) and false belief scores are entered in the third, both account for significant amounts of the variance on the referential opacity tasks. Thus, metarepresentational ability and metalinguistic ability make independent contributions to variance in referential opacity task scores.

I will now turn to the task of explaining what I think is going on in the mind of the child trying to deal with a referentially opaque context. My argument as to why I think that both thinking about thoughts and thinking about words (i.e., descriptions and names) are instrumental to being sensitive to referential opacity is as follows. A young child, when faced with multiple descriptions for a given referent and a person with limited knowledge of the referent, has a complex task if she or he is to answer a question about the person's mental attitude toward a proposition correctly (referential
opacity). In order to do so, the child must be able to represent the other’s *partial* representation (e.g., 'big ball') of the referent in order to distinguish it from her own, *more complete*, representation (e.g., 'big, yellow ball'). In order to maintain a representation of another's representation, the child, by definition, needs to *metarepresent*. It is not that being able to *metarepresent* causes the child to maintain a representation of another's representation; rather, what the child is doing is characterized by saying that it is *metarepresentational* in nature.

So far, the argument is no different than that presented elsewhere (Kamawar & Olson, in press). Where this argument differs from earlier ones is in its insistence that being able to think about thoughts does not tell the whole story. If it did, then performance on false belief tasks and performance on referential opacity tasks would be indistinguishable, but they are not (Apperly & Robinson, 1998; Kamawar & Olson, in press). So, in order to explain at least some of the difference between the two tasks, I also appeal to the notion of metalinguistic awareness.

As mentioned earlier, there is a wide range of skills that are considered to be metalinguistic in nature, so I need to be more exact in my use of the term. While some research has shown performance on referential opacity to be related to phoneme deletion and word deletion after age is partialled out (Artuso & Olson, 1999), and other research has shown it to be related to the say/mean-evaluation task after age is partialled out (e.g., "Did puppet tell us enough about his card?"; Apperly & Robinson, 1999), I am concerned with a different, but very specific, type of metalinguistic ability. My choice of this particular metalinguistic ability in explaining differences in performance on false belief and referential opacity tasks is based both on the theoretical reasons that were the basis of the construction of the say/refer task (see Chapters 3 and

---

3 As mentioned earlier, false belief tasks are used because they are a widely accepted measure of metarepresentational ability (*metarepresentational* in my terms).
6), and on the results of Study 2. This study demonstrates that performance on the say/refer task is able to account for a significant amount of variance on the referential opacity tasks even after age, digit span, vocabulary and false belief are entered into a hierarchical regression.

The specific kind of metalinguistic skill that I believe to be related to dealing with referentially opaque contexts (and measured by the say/refer task), is the ability to make s about coreferential names or descriptions; that is, the ability to distinguish between which utterances said the same thing from those that said something about the same thing. The say/refer task measured children's ability to attend to the exact description used because the key items could only be answered correctly by attending to the descriptors contained in the utterances (the referents gave no indication of who said what). In other words, the say/refer task measures the ability to treat a name/description as an entity which can be judged separately from that to which it refers. So then, the child trying to deal with a referentially opaque context, in addition to representing another's partial representation (mmetarepresent), also has to reflect on the particular description used to pick out a referent and make a as to whether it picks out the referent in the right way. Stated this way, it is not surprising that performance on this task is related to referential opacity performance: both require the child to treat the language used to refer as being more important than that to which the language refers. Said another way, both require the child to override considering the referent when evaluating an utterance, and instead, to pay attention to exact words used.

In summary, when trying to attribute a propositional attitude to another person, children do so by both thinking about thoughts, and by thinking about the particular language used to express those thoughts.

7.3 Studies 1 and 2

The design of the studies in this thesis only allows me to answer whether the
performance on two tasks is related; the studies say nothing about which abilities drive or allow for, which. However, this is a very interesting question, and one that I would like to address (at least theoretically). The question of whether children understand opacity by virtue of understanding what another could think (metarepresentation) or whether they understand what another could think by virtue of being able to use the appropriate linguistic form (propositional attitudes) is, I think, asking the wrong question. I don’t want to argue that metarepresentational development drives opacity development, nor the other way around. Rather, I propose that performance on all three tasks is driven by the development of an underlying competence with representation. In other words, I consider the tasks to be measuring the same thing: the ability to think about beliefs (minds) and linguistic expressions (words) metarepresentationally.

It is worth acknowledging that I have not shown the identity of false belief, say/refer and opacity tasks; these tasks differ in terms of difficulty. I want to point out that the measures are not merely independent tasks ordered by complexity. If that were the case the relation among them would disappear when covariates such as age, digit span and vocabulary were removed. Rather, some common factors must be present in addition to their slightly different complexities. That factor, I have suggested, is their common reliance on understanding representations, whether in minds or in language.

It is only with development that children begin to think of language as representational rather than referential. The results are consistent with the view that both passing false belief tasks and being able to deal appropriately with opaque contexts are manifestations of a developing understanding of representation more generally, and that children’s theories of mind and language follow similar developmental paths.

7.4 Future Research

A number of ideas have been mentioned in this chapter regarding future
research, but only one will be presented in detail. The research idea to follow is a summary of my proposed post-doctoral research project, the major objective of which is to map out children's understanding of the opaque uses of language across a variety of contexts and to develop an explanatory categorization system for such contexts.

The proposed line of research, then, is to map out children's developing ability to handle different types of opacity with an aim at understanding what makes some contexts 'more opaque' than others. The contexts to be examined include those already addressed by this thesis, but would extend to include desire verbs (e.g. 'want') and the class of verbs considered speech acts (e.g. 'promise', 'permit', and 'apologize') (Austin, 1962). The study will consist of three parts. The goal of Part One will be to map out the developmental trajectory of the specified opaque contexts. This will be done by telling children stories about characters with access to (partial) information about an object referent. While the story framework will remain the same, characters will have intentions, desires, beliefs etc. toward the object/action only under a specific representation (as appropriate to the specific opaque context being examined). Participants will answer questions about the story characters' intentions, desires, beliefs etc. toward the object under the description the characters are familiar with (e.g. 'the big ball'), as well as under the description that they are not (e.g. 'the yellow ball') to determine their proficiency with the contexts. The results of this part of the research project will lead directly into the second phase.

The goal of the second part of this study is to develop a system for categorizing the various contexts so as to shed light on their order of acquisition; in other words, to explain why some contexts are recognized as opaque more readily than others. The categorization system will be one that groups like contexts together by determining an objective set of criteria that explain their order of acquisition. This will be accomplished by examining literature from the philosophy of language, philosophy of mind, speech
act theory, and psychology of mind (e.g. theory of mind) in more detail.

The goal of the third part of this study is to test the categorization worked out in the second part. This will be done by performing small scale longitudinal studies in which participating children (at age ranges to be determined by part one), will be tested at fairly short intervals. For example, say it was found that two opaque contexts (e.g. those using ‘think’ and ‘know’), are theorized to have a common factor of ‘belief’ and develop between the ages of 4 and 5. What I would then do is frequently test a group of 4-year-olds to determine whether it is the case that the two types of opaque contexts (i.e., those using ‘think’ and ‘know’) do in fact develop simultaneously (examples will be counterbalanced across testing times). Further, in order to test the categorization system worked out in part two, the children will be given at least one other type of opaque context (e.g., that using the verb ‘want’) that the theory in Part Two predicts should also load on to the factor ‘belief’. If it does, I will have supporting evidence for my categorization system. At the test times, children will also be given a number of other tasks that are expected to be related to (perhaps even predictive of) opacity performance, such as metarepresentational, metalinguistic, general language and specific language (i.e., complementation) measures. By running this as a longitudinal study, it will become possible to determine what skills, if any, are necessary for successful completion of tasks with opaque contexts.

This research will prove interesting to both psychologists concerned with children’s cognitive development and philosophers of language concerned with semantics because the results will get at why sensitivity to opaque contexts is not all-or-nothing (it varies according to the context at issue) but can be explained by appeal to the categorization system I will develop, and will shed light on which skills are predictive of opacity performance.
7.5 Conclusion

Given the high degree of inter-relation among the tasks examined in each of the two studies, it seems that they make up a single domain of understanding. In other words, the tasks used seem to tap into a common pool of cognitive resources (e.g., metarepresentational ability, language proficiency). The lack of a significant correlation between every pair of key tasks (after factors such as age, digit span and vocabulary were partialled out), and the fact that every task is related to at least one other task, leaves me with the opinion that the tasks are related by family resemblance. All of the key tasks (i.e., everything but digit span and vocabulary) are united in that they require the child to represent an event/referent as a thought or as an expression cut off from the actual situation and to process that representation. In other words, all of these tasks require the child to either metarepresent or to pay attention to the language forms: they are all metarepresentational tasks.
References


Robinson, E. J., Goelman, H., & Olson, D. R. (1983). Children’s understanding of the relation between expressions (what is said) and intentions (what was meant). British Journal of Developmental Psychology, 1, 75 - 86.


Appendix A: Information Letter and Consent Form

Dear Parent/Guardian:

We are conducting a study of children's cognitive development and would like to include your child in our study. We are interested in children's understanding that sometimes you do or say things that you didn't mean to do or say, even those things that are done intentionally. For example, if someone turned on the outside light and it scared away a raccoon that she didn't know was there, we can't say that she meant to scare away the raccoon, even though she did mean to turn on the light and it scared the animal away.

The Associate Director of the Scarborough Board of Education, at the recommendation of the Research and Evaluation Advisory Committee, has given permission for this study to be carried out in your child's school. The study will be conducted in accordance with the ethical standards for research of both the Ontario Institute for Studies in Education and the Scarborough Board of Education.

The participating children will hear stories, some acted out with dolls and puppets, and some accompanied by colourful illustrations. They will be asked some questions about the stories. The stories will be read to each child individually by a female graduate student. The questions concern the story characters' thoughts, beliefs and intentions. Children who take part in these studies usually enjoy them.

The children will also complete short tasks that look at general language ability. Please note that we are not assessing individual children but observing them in order to increase our knowledge of children's development in general. Your child's responses will not be identified by name and we will not use any private information from school records. Their answers will give us important information about their understanding of people's thoughts and what can be said about their actions. There will be two sessions each lasting about fifteen minutes.

Would you please complete and sign the attached consent form and return it to your child's classroom teacher as soon as possible? If you allow your child to be included in the study, he or she will be asked if they want to take part; no children will be forced to participate.

We sincerely appreciate your cooperation. If you would like to receive more information about the study, please contact Deepthi Kamawar at 923-4909 or Prof. Olson at 923-6641, ext. 2572.

Thank you very much,

Prof. D. R. Olson, Ph.D.          Ms. Deepthi Kamawar, M.A.
Professor                        Graduate Student
Consent Form

I have read the attached letter that describes the study being conducted by Ms. Deepthi Kamawar and Prof. David R. Olson. I understand that the information she collects will be kept confidential and not be referred to except as a part of the total body of data that she collects. I also understand that my child’s name will not be used. Furthermore, I understand that participation is completely voluntary and the results have no bearing on any school or classroom evaluation of my child.

Child’s name ____________________________

Child’s date of birth ____________________________

All languages that your child speaks ____________________________

Child’s teacher ____________________________

I give permission for my child, named above, to participate in the University of Toronto study conducted by Prof. D. R. Olson and Deepthi Kamawar. I understand that my child will be asked if (s)he wants to take part and will not be required to do so if (s)he is shy or unwilling.

PLEASE CHECK HERE:  YES ___  NO ___

Signature of parent/guardian ____________________________

Date _______________
Appendix B: Opacity Stories and Questions for Study 1

- Note: the illustrations that were used with the participants were in colour.
- Stories 1-3 contain multiple descriptions of an object and stories 4-6 contain multiple descriptions of an action.

General Introduction to the stories: “I’m going to tell you some stories about two friends, Mark and Cathy. As we go along, I’m going to ask you some questions about them, and what they did. So listen carefully, okay?”

The Big Yellow Ball

One day, after Cathy finished talking to her friend Jenny on the phone, she got ready to have her friend Mark come over to play soccer.

Cathy got out all of her toy balls and put them on her bed. She had a new yellow ball, it was the biggest one, and she also had a red one and a green one.

The light bulb burnt out and she was in the dark. She didn’t have time to ask her mom to change the bulb because just then, the doorbell rang. She went downstairs to let Mark in.

Cathy asked Mark to get a ball from her room. She told him that all of the balls were on the bed because she knew that they would be playing soccer. She said, “Get any ball you
want except for the yellow ball because I don’t want to get it dirty”. As Mark goes up to Cathy’s room he says, “I don’t care what colour it is. I’m going to take the big ball.” He wanted the big ball because it’s best for playing soccer.

Upstairs in Cathy’s room, Mark can’t see anything because the light is out. He finds his way to the bed, he feels around until he finds the biggest one [dramatize]. He can’t tell that it’s the yellow ball.

He picks it up and starts to leave Cathy’s room. [note: questions are asked before Mark leaves the room]

Questions:

Comprehension:
Did Mark take the big ball? [correct answer: yes] – transparent context
Did he take the yellow ball? [yes] – transparent context
Did he know that Cathy was talking on the phone before he came over? [no]

Referential Opacity:
Did he know that he took the big ball? [yes]
Did he know that he took the yellow ball? [no]

Intentional Opacity:
Did Mark mean to take the big ball? [yes]
Did he mean to take the yellow ball? [no]

1 Note: The questions did not appear in this order.
2 Either ‘Mark’ or ‘he’ was used, depending on the order in which the questions appeared.
Linguistic Opacity:
Remember what Mark said to Cathy. He said, "I'm going to take ..."
Did he say that he was going to take the yellow ball? [no]
Did he say that he was going to take the big ball? [yes]
The Doctor’s Office

One day after school Mark and Cathy were playing on the swings. They were having contests to see who could swing the highest.

Mark was winning, but then he fell off the swing and hurt his knee. It wasn’t too hurt badly, but it did need a bandage so they went to the school health room.

Cathy went with Mark to keep him company. Here they are in the Health Room waiting for an adult to help. Mark said to Cathy, “Please go and get someone to put a bandage on my knee.” Cathy said, “Maybe I should get a doctor.” and Mark said, “Please don’t get a doctor. I don’t need one.” Cathy said, “Okay, I won’t get a doctor”. Cathy went to the office to get an adult to help.

While she was gone, Mark ate a small chocolate bar he had in his pocket. He finished it before Cathy came back.
When Cathy got to the office, she saw Mark’s mom. She had come to the school to pick Mark up. Cathy said to herself, “There’s Mark’s mom. I should ask her to help” so she went up to Mark’s mom and said, “Mark hurt himself on the swings. Could you come look at Mark’s knee and put a bandage on it?” Cathy couldn’t tell that Mark’s mom was a doctor.

She took her to the Health Room. She put a bandage on Mark’s knee and he felt all better.

Comprehension Questions:
Did Cathy take Mark’s mom to the Health room? [correct answer: yes]
Did she take a doctor to the Health room? [yes]
Does she know that Mark ate a chocolate bar? [no]

Referential Opacity:
Does Cathy know that she took Mark’s mom to the health room? [yes]
Does she know that she took a doctor to the health room? [no]

Linguistic Opacity:
Remember what Cathy said to mark. She said, "I won’t get …"
Did she say that she wouldn’t get a doctor? [no]
Did she say that she wouldn’t get Mark’s mom? [yes]

Intentional Opacity question:
Did Cathy mean to get Mark’s mom to help? [yes]
Did she mean to get a doctor to help? [no]

3 Either ‘Cathy’ or ‘she’ was used, depending on the order in which the questions appeared.
Police Officer

On the street there was a Police Car. By the car, the Police Officer was tying his shoes.

Then he got out up and started walking to a red house. Mark was walking home from school and he walked down that street. He saw the Police Officer drop a set of keys on the sidewalk. The Police Officer did not notice that he dropped the keys.

When Mark got to where the keys were, he picked them up and said, “I should give these keys to the Police Officer”.

The Police Officer walked back to where he had come from. Mark walked over to the Police Officer. Mark said, “Hi, you dropped you keys”. Mark couldn't tell that the Police Officer was Cathy's dad. Mark gave the Police Officer the keys, and the Police Officer said, “Thank you”.

Comprehension Questions:
Did Mark give the keys to the Police Officer? [correct answer: yes]
Did Mark give the keys to Cathy’s dad? [yes]
Does Mark know that the Police Officer tied his shoelaces? [no]
Referential Opacity:
Does Mark know that he gave the keys to the Police Officer? [yes]
Does Mark know that he gave the keys to Cathy’s dad? [no]

Linguistic Opacity:
Remember what Mark said. He said, "I should give these keys to ..."
Did he say that he should give the keys to Cathy’s dad? [no]
Did he say that he should give the keys to the Police Officer? [yes]

Intentional Opacity:
Did Mark mean to give the Police Officer the keys? [yes]
Did Mark mean to give Cathy’s dad the keys? [no]
One night, Mark was sitting in the kitchen having a snack. The back door of the house was right by the kitchen. Outside, right by the back door, there was a raccoon in the dark. It was very, very quiet so Mark didn’t hear it.

Mark’s dad was upstairs wrapping a surprise gift for Mark. He called down, “Mark, please turn on the light by the back door.”

Mark got up and flicked the light switch and the light just above the back door turned on.

Mark said to his dad, “Okay Dad, I turned on the light”. The light scared away the raccoon and it ran away so quietly Mark couldn’t tell that a raccoon was ever there.

Questions:

Comprehension:
Did Mark turn on the light? [correct answer: yes] – transparent context
Did he scare away the raccoon? [yes] – transparent context
Did he know that his dad was wrapping a surprise gift? [no]
Referential Opacity:
Did he know that he turned on the light? [yes]
Did he know that he scared away the raccoon? [no]

Intentional Opacity:
Did Mark mean to turn on the light? [yes]
Did he mean to scare away the raccoon? [no]

Linguistic Opacity:
Remember what Mark said to his dad. He said, "Okay Dad, I ..."
Did he say that he turned on the light? [yes]
Did he say that he scared away the raccoon? [no]
Pouring water/watering a plant

On a Saturday afternoon, Cathy was sitting by the window working at her desk. She had a glass of water next to her. Cathy was busy drawing a picture for her dad. Downstairs, Dad was hiding a surprise birthday cake for Cathy.

Cathy was about to drink her water when she saw a bug floating in it. She didn’t want to drink it anymore, so

she poured out all of the water out of the window and it landed on a plant growing outside.

Cathy took her glass downstairs to get more water. She went into the kitchen where Dad was. Dad said, “You must be very thirsty Cathy - you’re drinking so much water” and Cathy said; “I didn’t drink all of that water. There was a bug in it so I poured the water out of the window”. Cathy couldn’t tell that when she poured the water out of her window, it fell on to a plant outside that was very thirsty and watered it.
Comprehension Questions:
Did Cathy pour the water out of the window? [yes]
Did Cathy water the plant? [yes]
Does Cathy know that her dad hid a surprise cake for her? [no]

Referential Opacity:
Did Cathy know that she poured the water out of the window? [yes]
Did Cathy know that she watered the plant? [no]

Linguistical Opacity:
Remember what Cathy said to her Dad. She said, "I ..."
Did she say that she watered the plant? [no]
Did she say that she poured the water out of the window? [yes]

Intentional Opacity:
Did Cathy mean to pour the water out of the window? [yes]
Did Cathy mean to water the plant? [no]
Bread Crumbs/Ants

One sunny day, Cathy and Mark decided to have a picnic lunch. Cathy went to her house to get some fruit and juice and

Mark went home to make sandwiches. When Mark was making the sandwiches, he was listening to the radio and he heard his favourite song.

After they got the food, they met at the park. They spread out a blanket and sat down to have lunch. It was warm outside and Cathy and Mark were happy to be sitting in the sun. They both had the sandwiches, apples and orange juice. As they ate their sandwiches, little bread crumbs fell all over the picnic blanket.

After they were done, they packed up all of their things. Cathy picked up the picnic blanket and shook off the crumbs. Mark asked, "Cathy, what are you doing?". Cathy said, "I am shaking the blanket clean." and she shook all of the crumbs off of it. Cathy couldn't tell that there were ants in the grass and
and that they were fed by the bread crumbs.

Comprehension Questions:
Did Cathy shake the blanket? [correct answer: yes]
Did Cathy drop food for the ants to eat? [yes]
Does Cathy know that Mark listened to his favourite song when he made the sandwiches? [no]

Referential Opacity:
Did Cathy know that she shook the bread crumbs off of the blanket? [yes]
Did Cathy know that she dropped food for the ants to eat? [no]

Linguistical Opacity:
Remember what Cathy said to Mark. She said, "I ..."
Did she say that she fed the ants? [no]
Did she say that she shook the blanket clean? [yes]

Intentional Opacity:
Did Cathy mean to shake the blanket clean? [yes]
Did Cathy mean to drop food for the ants? [no]
Appendix C: False Belief Tasks

Task 1: Unexpected contents

Look at this... What's in here? [usual answer: crayons/colours]

Let's open it and have a look. What is it? [correct answer: stickers]

some brief conversation about stickers, e.g. yes! stickers!
I just put them in this box to keep them safe..... Well... let's put them back into the box...

(When box is closed) What's in the box? [correct: stickers]

*False Belief for self:
What did you think was inside the box BEFORE we opened it? [correct: what child said for first question]

*False Belief for other:
X (friend's name) hasn't seen inside this box.
What will s/he think is inside it before s/he opens it? [correct: crayons/colours]

Task 2: Appearance/Reality Distinction

Look at this... What is it? [usual answer: rock/stone]

Look, I can squeeze it, I can wipe things up with it.
Let child hold it, squeeze it, then take it back. What is it? [usual answer: sponge/foam]

*False belief for self:
What did you think it was when you first saw it, before you played with it?
[correct: rock/stone]

*False belief for other:
X (friend's name) hasn't seen this before.
What will s/he think it is when s/he FIRST sees it, BEFORE s/he plays with it? [correct: rock/stone]

**Appearance question:
What does this LOOK LIKE? A rock [or whatever child called it, e.g. stone] or a sponge? [correct: rock/stone/what child initially called it]

**Reality question:
What is this REALLY? A rock or a sponge? [correct: sponge/foam]
Task 3: Change of location

I'm going to tell you a story, listen carefully, then I'll ask you some questions. OK?

The boy has a toy turtle. He puts it away in his red box.
He goes upstairs. (move doll under one side of table)
While he's gone his sister takes the turtle out of the red box.
She plays with it, then she puts the turtle away.
She puts it away in the yellow box.
Then the girl goes outside. (move doll under OTHER side of table)
The boy comes back. He wants to play with his turtle.

*False belief for other:
Where will the boy look for the turtle? [correct: red box]

Memory check:
Where did he put the turtle before he went upstairs? [correct: red box]

Reality check:
Where is the turtle really? [correct: yellow box]

Task 4: True belief

Look at this... What's in here? [answer: bandaids]

Let's open it and have a look. What are they? [correct: bandaids]

some brief conversation about bandaids... Then... Let's play a trick on X (friend). Let's take the bandaids out and put pennies in here. [do so and close box]

(When box is closed) What's in the box? [correct: pennies]

True belief memory:
What did you think was inside the box BEFORE we opened it? [correct: bandaids]
Appendix D: Tables for Study 1 (Chapter 5)

Table 2.

Performance on Intentional Opacity by Performance on Referential Opacity Across Referent Types

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Intentional Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Fail</td>
<td>23</td>
</tr>
<tr>
<td>Pass</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

χ² (1, N=73) = 10.56, p<.01

Note. A pass is 5/6 or better.

Table 3.

Performance on Intentional Opacity by Performance on Referential Opacity for Object Referents

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Linguistic Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>Fail</td>
<td>22</td>
</tr>
<tr>
<td>Pass</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>73</td>
</tr>
</tbody>
</table>

χ² (1, N=73) = 16.40, p<.001

Note. A pass is 2/3 or better.
Table 4.
Performance on Intentional Opacity by Performance on Referential Opacity for Action Referents

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Linguistic Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>15</td>
</tr>
<tr>
<td>Pass</td>
<td>58</td>
</tr>
<tr>
<td>Fail</td>
<td>15</td>
</tr>
<tr>
<td>Pass</td>
<td>58</td>
</tr>
<tr>
<td>31</td>
<td>42</td>
</tr>
</tbody>
</table>

$\chi^2(1, N=73) = 25.58, p < .001$

Note. A pass is 2/3 or better.

Table 5.
Performance on Linguistic Opacity by Performance on Referential Opacity Across Referent Types

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Linguistic Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>23</td>
</tr>
<tr>
<td>Pass</td>
<td>49</td>
</tr>
<tr>
<td>Fail</td>
<td>1</td>
</tr>
<tr>
<td>Pass</td>
<td>30</td>
</tr>
<tr>
<td>42</td>
<td>31</td>
</tr>
</tbody>
</table>

$\chi^2(1, N=73) = 21.5, p < .001$

Note. A pass is 5/6 or better.
Table 6.

Performance on Linguistic Opacity by Performance on Referential Opacity for Object Referents

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Linguistic Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>Fail: 19, Pass: 3</td>
</tr>
<tr>
<td>Pass</td>
<td>Fail: 16, Pass: 35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>19</td>
</tr>
<tr>
<td>Pass</td>
<td>51</td>
</tr>
<tr>
<td>35</td>
<td>38</td>
</tr>
</tbody>
</table>

\[ \chi^2 (1, N=73) = 18.62, p < .001 \]

Note. A pass is 2/3 or better.

Table 7.

Performance on Linguistic Opacity by Performance on Referential Opacity for Action Referents

<table>
<thead>
<tr>
<th>Referential Opacity</th>
<th>Linguistic Opacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>Fail: 11, Pass: 4</td>
</tr>
<tr>
<td>Pass</td>
<td>Fail: 10, Pass: 48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>21</td>
</tr>
<tr>
<td>Pass</td>
<td>58</td>
</tr>
<tr>
<td>21</td>
<td>52</td>
</tr>
</tbody>
</table>

\[ \chi^2 (1, N=73) = 18.3, p<.001 \]

Note. A pass is 2/3 or better.
Table 8.

Performance on the False Belief Tasks by Age.

<table>
<thead>
<tr>
<th>False Belief</th>
<th>Age (in years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Fail</td>
<td>15</td>
</tr>
<tr>
<td>Pass</td>
<td>8</td>
</tr>
</tbody>
</table>

\[ \chi^2 (2, N=73) = 29.84, p<.0001 \]

Note. A pass is 5/6 or better.
Appendix E: Correlation Matrix for All Measures

Table 9.

Correlation Matrix for all Measures.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Ref. Opacity</td>
<td>.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Lin. Opacity</td>
<td>.76</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Int. Opacity</td>
<td>.69</td>
<td>.68</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Object Ref.</td>
<td>.80</td>
<td>.81</td>
<td>.88</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Action Ref.</td>
<td>.74</td>
<td>.91</td>
<td>.86</td>
<td>.82</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Obj. – Ref.</td>
<td>.73</td>
<td>.94</td>
<td>.75</td>
<td>.69</td>
<td>.84</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Obj. – Lin.</td>
<td>.70</td>
<td>.60</td>
<td>.86</td>
<td>.67</td>
<td>.86</td>
<td>.69</td>
<td>.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Obj. – Int.</td>
<td>.60</td>
<td>.50</td>
<td>.61</td>
<td>.89</td>
<td>.84</td>
<td>.59</td>
<td>.55</td>
<td>.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Act. – Ref.</td>
<td>.66</td>
<td>.94</td>
<td>.68</td>
<td>.59</td>
<td>.67</td>
<td>.87</td>
<td>.76</td>
<td>.54</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Act. – Lin.</td>
<td>.61</td>
<td>.71</td>
<td>.87</td>
<td>.57</td>
<td>.65</td>
<td>.84</td>
<td>.71</td>
<td>.48</td>
<td>.46</td>
<td>.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Act. – Int.</td>
<td>.65</td>
<td>.72</td>
<td>.69</td>
<td>.93</td>
<td>.76</td>
<td>.87</td>
<td>.68</td>
<td>.62</td>
<td>.64</td>
<td>.66</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. FB</td>
<td>.54</td>
<td>.71</td>
<td>.54</td>
<td>.52</td>
<td>.58</td>
<td>.67</td>
<td>.67</td>
<td>.41</td>
<td>.38</td>
<td>.66</td>
<td>.52</td>
<td>.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Vocabulary</td>
<td>.65</td>
<td>.69</td>
<td>.70</td>
<td>.61</td>
<td>.68</td>
<td>.73</td>
<td>.66</td>
<td>.56</td>
<td>.49</td>
<td>.64</td>
<td>.64</td>
<td>.61</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Digit Span</td>
<td>.79</td>
<td>.72</td>
<td>.73</td>
<td>.66</td>
<td>.70</td>
<td>.77</td>
<td>.68</td>
<td>.59</td>
<td>.52</td>
<td>.66</td>
<td>.67</td>
<td>.66</td>
<td>.57</td>
<td>.65</td>
<td></td>
</tr>
</tbody>
</table>

Note. All correlations have N = 73 and are significant at \( p < .001 \) (exception is underlined, significant at \( p < .01 \).
Appendix F: Opacity Stories and Questions for Study 2

Note: the illustrations that were used with the participants were in colour.

---

**General Introduction to the stories:** “I’m going to tell you some stories about two friends, Mark and Cathy. As we go along, I’m going to ask you some questions about them, and what they did. So listen carefully, okay?”

---

**The Big Yellow Ball**

One day, Cathy got out all of her toy balls and put them on her bed. She had a new yellow ball, it was the biggest one, and she also had a red one and a green one.

The light bulb burnt out and she was in the dark. She didn’t have time to ask her mom to change the bulb because just then, the doorbell rang. She went downstairs to let Mark in.

Cathy asked Mark to get a ball from her room. She told him that all of the balls were on the bed because she knew that they would be playing soccer. She said, “Get any ball you want except for the yellow ball because I don’t want to get it dirty”. As Mark goes up to Cathy’s room he says, “I don’t care what colour it is. I’m going to take the big ball.” He wanted the big ball because it’s best for playing soccer.
Upstairs in Cathy's room, Mark can't see anything because the light is out. He finds his way to the bed, he feels around until he finds the biggest one [dramatize]. He can’t tell that it’s the yellow ball.

He picks it up and starts to leave Cathy's room.

Questions:

Comprehension:
Did Mark take the big ball? [correct answer: yes] – transparent context
Did he take the yellow ball? [yes] – transparent context
Did Cathy know that her room was dark? [yes]

Referential Opacity:
Did he know that he took the big ball? [yes]
Did he know that he took the yellow ball? [no]

Quotational Opacity:
Remember what Mark said to Cathy: He said, “I’m going to take ...”
Did he say “I’m going to take the big ball”? [yes]
Did he say “I’m going to take the yellow ball”? [no]

*Note: The questions did not appear in this order.*
The Doctor’s Office

One day after school Mark and Cathy were playing on the swings. They were having a contest to see who could swing the highest. Mark was winning, but then

he fell off the swing and hurt his knee. It wasn’t hurt badly, but it did need a bandage so they went to the school health room.

Cathy went with Mark. Here they are in the Health Room waiting for help. Mark said to Cathy, “Please get someone to put a bandage on my knee.” Cathy said, “I will get a doctor.” and Mark said, “Please don’t get a doctor. I don’t need one.”. Cathy said, “Okay, I won’t get a doctor”. Cathy went to the office to get help.

When Cathy got to the office, she saw Mark’s mom. And guess what - Mark’s mom is really a doctor. She had come to the school to pick Mark up. Cathy said to herself, “There’s Mark’s mom. I’ll ask her to help” so she went up to Mark’s mom and said, “Mark hurt himself on the swings. Could you come look at Mark’s knee and put a bandage on it?”. Cathy couldn’t tell that Mark’s mom was a doctor.
She took her to the Health Room. She put a bandage on Mark's knee and he felt all better.

Questions

Comprehension:
Did Cathy take a doctor to the Health room? [correct answer: yes] – transparent context
Did Cathy take Mark's mom to the Health room? [yes] – transparent context
Did she know that Mark hurt his knee? [yes]

Referential Opacity:
Did she know that she took Mark's mom to the health room? [yes]
Did she know that she took a doctor to the health room? [no]

Quotational Opacity:
Remember what Cathy said to Mark: She said, "Okay, I won't ..."
Did she say "Okay, I won't get a doctor"? [yes]
Did she say "Okay, I won't get Mark's mom"? [no]
Police Officer

Mark was walking home from school one day when he saw a Police Officer drop a set of keys on the sidewalk. The Police Officer did not see that he dropped the keys.

When Mark got to where the keys were, he picked them up and said, "I should give these keys to the Police Officer".

The Police Officer got to the house and no one was home so he walked back to his car. Mark walked over to the Police Officer. Mark said, "Hi, you dropped you keys". And guess what - the Police Officer was really Cathy’s Dad. Mark couldn’t tell that the Police Officer was Cathy’s dad. Mark gave the Police Officer the keys, and the Police Officer said, “Thank you”.

Questions:

Comprehension:
Did Mark give the keys to the Police Officer? [correct answer: yes] – transparent context
Did he give the keys to Cathy’s dad? [yes] – transparent context
Did he know that the Police Officer dropped his keys? [yes]
Referential Opacity:
Did he know that he gave the keys to the Police Officer? [yes]
Did he know that he gave the keys to Cathy's dad? [no]

Quotational Opacity:
Remember what Mark said: He said, “I should give these keys ...”
Did he say “I should give these keys to the Police Officer”? [yes]
Did he say “I should give these keys to the Cathy’s dad”? [no]
Appendix G: The Say/Refer Task

Introduction:
• “These two children (Playmobil dolls), Anna and Mark, are going to play a game. In this game they each have the same pictures and they sit on different sides of a little wall – and they can’t see through the wall. They are going to take turns picking cards and see if they pick the same card as each other.”

• “Let’s look at one turn. Mark is picking a card and putting in front of him. Anna is picking a card and putting it in front of her. Now they are going to tell each other which one they have.”

Practice Trials:

Trial I
Mark: “I have a tree.”
Anna: “I have a tree.”
Q: “Do they have the same one?”
** give corrective feedback if necessary

Practice Trial II
• “Let’s look at another turn. Anna picks a card, then Mark picks one.”
Anna: “I have a tree.”
Mark: “I have a house.”
Q: “Do they have the same one?”
** give corrective feedback if necessary

Concept Check:
• “Let’s look at all the cards Anna and Mark will be playing with. [go through cards in random order asking what it is, and what colour it is, at end, ask child to show the big/small version of the star/balloon/flower].

Warm-up Trials:

• “Let’s watch them take another turn. Listen carefully to what the kids say because I am going to ask you questions about what Anna and Mark say. Sometimes the right answer to my question will be Mark said it, sometimes it will be Anna said it, sometimes it will Anna and Mark said it, and sometimes it will be that nobody said it. Let’s start okay?”

Picture Array contains: - heart - house - car - tree

Warm-up Trial A
"Anna and Mark have each picked a picture and put in front of them. Now they are going to tell each other which one they have."
A: "I have a tree"
M: "I have a heart"
Qa: Do they have the same one?
Qb: Who said “I have a tree”? - was it just Anna, just Mark, Anna and Mark or nobody?
If incorrect, follow up with:
"Now remember, Anna said, "I have a tree" and Mark said, "I have a heart"
Qb': So, who said "I have a tree?" - was it just Anna, just Mark, Anna and Mark or nobody?

Warm-up Trial B
"Anna and Mark have each picked another picture and put in front of them. Now they are going to tell each other which one they have."
M: "I have a house"
A: "I have a house"
Qa: Do they have the same one?
Qb: Who said "I have a house"? - was it just Anna, just Mark, Anna and Mark or nobody?

If incorrect, follow up with:
"Now remember, Anna said, "I have a house" and Mark said, "I have a house".
Qb': So, who said "I have a house"? - was it just Anna, just Mark, Anna and Mark or nobody?

Warm-up Trial C
"They have each put a picture in front of them and they're going to tell each other which one they have."
A: "I have a car"
M: "I have a tree"
Qa: Do they have the same one? [no]
Qb: Who said "I have a heart"? - was it just Anna, just Mark, Anna and Mark or nobody?

If incorrect, follow up with:
"Now remember, Anna said, "I have a car" and Mark said, "I have a tree".
Qb': So, who said "I have a heart"? - was it just Anna, just Mark, Anna and Mark or nobody?

Second try at warm-up: (only given if children were incorrect at second try for any of the above warm-up trials)

Warm-up Trial D
"Anna and Mark have each picked a picture and put in front of them. Now they are going to tell each other which one they have."
M: "I have a heart"
A: "I have a car"
Qa: Do they have the same one? [no]
Qb: Who said "I have a heart"? Was it just Anna, just Mark, Anna & Mark, or nobody?

Warm-up Trial E
"Anna and Mark have each picked another picture and put in front of them. Now they are going to tell each other which one they have."
A: "I have a tree"
M: "I have a tree"
Qa: Do they have the same one? [yes]
Qb: Who said "I have a tree"? - was it just Anna, just Mark, Anna and Mark or nobody?
Warm-up Trial F
"They have each put a picture in front of them and they're going to tell each other which one they have."
M: "I have a tree"
A: "I have a house"
Qa: Do they have the same one? [no]
Qb Who said "I have a heart"? - was it just Anna, just Mark, Anna and Mark or nobody?

Test Items:
Scores from Qb on items 1, 3, 8, 9, 10, 11, and 12 were summed to provide a composite score on this task.

#1
M: "I have an orange star"
A: "I have a big star"
Qa: "Do they have the same one?"
Qb: Who said "I have an orange star"? - was it just Anna, just Mark, Anna and Mark or nobody?

#3
M: "I have a yellow balloon"
A: "I have a big balloon"
Qa: "Do they have the same one?"
Qb: Who said, "I have the big balloon" - was it just Anna, just Mark, Anna and Mark or nobody?

#8
A: "I have a small star"
M: "I have a small flower"
Qa: "Do they have the same one?"
Qb: Who said, "I have a blue flower" - was it just Anna, just Mark, Anna and Mark or nobody?

#9
M: "I have a small flower"
A: "I have a blue flower"
Qa: "Do they have the same one?"
Qb: Who said, "I have a small, blue flower" - was it just Anna, just Mark, Anna and Mark or nobody?

#10
A: "I have a big star"
M: "I have an orange star"
Qa: "Do they have the same one?"
Qb: Who said, "I have a big, orange star" - was it just Anna, just Mark, Anna and Mark or nobody?

---

The items were not presented in this order.
#11
M: "I have a big, yellow balloon"
A: "I have a big, yellow balloon"
Qa: "Do they have the same one?"
Qb: Who said, "I have a yellow balloon" - was it just Anna, just Mark, Anna and Mark or nobody?

#12
A: "I have a big, red flower"
M: "I have a big, red flower"
Q: "Do they have the same one?"
Qb: Who said, "I have a big flower" - Just Anna, just Mark, Anna and Mark, or nobody?

#2
A: "I have a small flower"
M: "I have a yellow balloon"
Qa: "Do they have the same one?"
Qb: Who said, "I have a yellow balloon" - was it just Anna, just Mark, Anna and Mark or nobody?

#4
A: "I have an orange star"
M: "I have a small balloon"
Qa: "Do they have the same one?"
Qb: Who said, "I have an orange star" - was it just Anna, just Mark, Anna and Mark or nobody?

#5
M: "I have a big star"
A: "I have a big star"
Qa: "Do they have the same one?"
Qb: Who said, "I have a big star" - was it just Anna, just Mark, Anna and Mark or nobody?

#6
A: "I have a green star"
M: "I have a green star"
Qa: "Do they have the same one?"
Qb: Who said, "I have a green star" - was it just Anna, just Mark, Anna and Mark or nobody?

#7
M: "I have a small balloon"
A: "I have a small balloon"
Qa: "Do they have the same one?"
Qb: Who said "I have a small balloon" - was it just Anna, just Mark, Anna and Mark or nobody?
Appendix H: Say/Mean Task

Directions:
"Now you're going to play a game with puppet. This is Leo and he is going to sit over here, on this side of this wall, and you're going to sit here, on the other side. You and puppet with both have the same set of cards. Let's look at them."

[show child the cards in 'random' order] "What's this? what colour is it? etc. until all cards are down on table in front of child. Then ask "Show me a big flower [pause], show me a small flower [pause], show me a big star [pause], show me a small star" (to make sure they know the sizes).

"Okay. In this game, puppet is going to tell you to pick up a card and you have to pick up the card that she/he tells you to get. Then we'll see if you both have the same card, okay? Good. Let's try one."

Warm-up with feedback:

Trial A:
puppet: "the big, red flower"
to child: "pick the card that puppet said to pick".
* if incorrect, say, let's listen to puppet again. "The big, red flower" "You should pick the card that has the big, red flower on it. Which card is that?"

Trial B:
puppet: "the small, yellow star"
to child: "pick the card that puppet said to pick".
* if incorrect, say, let's listen to puppet again. "The small, yellow star" "You should pick the card that has the small, yellow star on it. Which card is that?"

Continue with trials C/D if necessary (i.e., if incorrect on A or B):

Trial C:
puppet: "The big, black star"
to child: "pick the card that puppet said to pick".
correct: [yes] [no]

Trial D:
puppet: "The small, blue flower"
to child: "pick the card that puppet said to pick."
correct: [yes] [no]

Test Items:
The scores from items 7, 11 and 13 were summed to give the say/mean – evaluation score.
The scores from items 2, 6 and 15 were summed to give the say/mean – judgment score.

* The items did not always appear in this order.
puppet: "a small, blue flower"
same card chosen:

puppet: "a yellow star"
E: "'a big/small yellow star', is that what puppet said?"
    if no then: what did puppet say?

puppet: "a small, black star"
same card chosen:

puppet: "a red flower"
E: 'a red flower', is that what puppet said?

puppet: "a big, black star"
same card chosen:

puppet: "a big flower"
E: "'a big, red/blue flower, is that what puppet said?"
    if no then: what did puppet say?

puppet: "a big star"
E: Did puppet tell you enough about his card?
    if 'no' then: What should puppet have said?

puppet: "a small, yellow star"
same card chosen:

puppet: "a black star"
E: 'a yellow star', is that what puppet said?

puppet: "a small, red flower"
same card chosen:

puppet: "a small star"
E: Did puppet tell you enough about his card?
    if 'no' then: What should puppet have said?
#12
puppet: "a big, black star"
same card chosen:

#13
puppet: "a small flower"
E: Did puppet tell you enough about his card?
    if 'no' then: What should puppet have said?

#14
puppet: "a big, blue flower"
same card chosen:

#15
puppet: "a big star"
E: "'a big, black/yellow star, is that what puppet said?"
    if no then: what did puppet say?
Appendix I: Table for Study 2 (Chapter 6)

Table 17.  
Performance on the False Belief Tasks by Age.

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>16</td>
<td>3</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>Pass</td>
<td>6</td>
<td>18</td>
<td>21</td>
<td>45</td>
</tr>
</tbody>
</table>

\( \chi^2 (2, N=64) = 30.8, p<.0001 \)

*Note.* A pass is 5/6 or better.