The Cognitive Effects of Bilingualism: Does Knowing Two Languages Impact Children's Ability to Reason about Mental States?

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Katherine Ruth Gordon Millett

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Maria D. Sera

April 2010

Acknowledgements

I would like to thank my advisor, Maria Sera, for her guidance throughout this project and throughout my graduate school career. I would also like to thank my committee members: Stephanie Carlson, Melissa Koenig, and Kathryn Kohnert for all of their suggestions and feedback throughout the entire dissertation process. Each of their unique perspectives was invaluable. Many thanks to Taryn Halverson's for her help with data collection and the parents and children who were willing to participate in the study. Funding from the Institute of Child Development made this project possible and is much appreciated as well as the support from ICD faculty, students, and staff. Finally, I would like to thank my parents and family for inspiring me to set high educational goals and for their encouragement throughout the process.

Dedication

This dissertation is dedicated to my husband, Ben, for believing in me and my abilities. It is also decided to our son for motivating me to finish before he joined our family.

Abstract

In a number of studies, bilingual children have been shown to outperform monolingual children in false belief tasks, thus providing evidence that bilingualism affects children's ability to reason about the mental states of others. However, there are two limitations to this past work. The first limitation is that false belief tasks only measure a limited aspect of children's mental state reasoning abilities. Thus, performance in false belief tasks cannot be assumed to reflect a general ability to reason about the mental states of others. Secondly, the language skills of the bilingual groups included in this past work were only reliably measured in one language. Thus, we do not have a good understanding of how language proficiency across both languages impacts mental state reasoning abilities. In order to address these limitations, 3- to 5-year-old Spanish-English bilingual children and English monolingual children were tested using Wellman and Liu's (2004) scale which assesses a variety of aspects of mental state reasoning. The scale includes the following tasks: a Diverse Desires task, a Diverse Beliefs task, a Knowledge Access task, a Contents False Belief task, an Explicit False Belief task, a Belief-Emotion task, and a Real-Apparent Emotion task. Additionally, the language proficiency of the bilingual group was measured in both English and Spanish using standardized measures of vocabulary comprehension (the PPVT and the TVIP). Results indicate that when English vocabulary level was controlled, the bilingual children outperformed the monolingual children in the Diverse Desires task. Furthermore, effect sizes suggest that the bilingual children also outperformed the monolingual children in the Knowledge Access, Belief-Emotion, and Real-Apparent Emotion tasks when English vocabulary level was controlled. Overall, these findings provide evidence that

bilingualism contributes to a broader effect on mental state reasoning than has been previously found.

Table of Contents

List of Tables

Chapter 1: Introduction

Parents and educators have long asked questions about how bilingualism impacts the language and cognitive development of children. Although these groups are interested in these questions for different reasons, they are both extremely invested in gaining a better understanding of the long-term effects of exposing a child to two languages early in life. For example, parents who immigrate to the United States often worry about the impact of exposing their children to both English as well as the language of their country of origin. Some of these parents fear that teaching their child two languages will compromise the child's ability to learn English and subsequently their ability to succeed in the United States. Conversely, many native English-speaking parents within the United States enroll their children in bilingual programs with the hope that early exposure to a second language will not only lead to the ability to speak and understand two languages, but will also lead to cognitive advantages as well (King & Fogle, 2006). Likewise, because of the growing number of non-native English speakers who are entering the school system (U.S. Census Bureau, 2003), educators have a lot of questions about how to help children who are learning more than one language succeed in the classroom and in general society.

Like parents and educators, psychologists are dedicated to helping children achieve positive language and cognitive outcomes. However, they are also interested in understanding the relationship between bilingualism and cognition for theoretical reasons. Part of that theoretical exploration involves pinpointing the specific aspects of cognition that are impacted by bilingualism. Although some parents believe that exposing children to two languages will increase their overall intelligence, evidence for

bilingual effects on cognition has only been shown in specific areas (King & Fogle, 2006). For example, there is a substantial body of evidence that bilingual children outperform monolingual children in Executive Function tasks (Bialystok, 1999; Bialystok, Craik, & Luk, 2008; Bialystok, Majumder, & Martin, 2003; Bialystok & Senman, 2004; Bialystok & Shapero, 2005; Carlson & Meltzoff, 2008; Costa, Hernández, & Sebastián-Gallés, 2008; Martin-Rhee & Bialystok, 2008). Recent work also suggests that bilingualism has a positive effect on children's ability to reason about the mental states of others (Bialystok & Senman, 2004; Goetz, 2003; Kovács, 2009), which is the focus of this dissertation.

While researchers have found a link between bilingualism and the ability to reason about mental states, there are several key unanswered questions about this relationship. First, because the majority of this work has been conducted using false belief tasks, we do not know whether bilingualism also has an effect on other areas of mental state reasoning. In other words, we lack an understanding of the *nature* and *limitations* of the bilingual effect on the development of mental state reasoning. Secondly, although there are some existing hypotheses about how bilingualism impacts mental state reasoning, researchers have yet to explore these hypotheses in depth. Thus, we lack an understanding of the *causes* of the bilingual effect on mental state reasoning. The purpose of this dissertation is to address the first question: What is the nature of the bilingual effect on mental state reasoning? In order to address this question, the current study compares bilingual and monolingual children's performance in a wider variety of mental state tasks than have been used in the past. The second question (i.e. the causes of the bilingual effect on mental state reasoning) is not directly explored in the current

study. However, the results of this study can direct the focus of subsequent research as researchers more carefully explore the causes of this effect.

This dissertation is organized into eight chapters. In Chapter 1, I will give an overview of the theoretical perspectives and empirical evidence for the Language and Cognition Debate (i.e. to what extent does language impact other areas of cognition). In Chapter 2, I will discuss the challenges and considerations when defining "bilingualism" in psychological research. Chapter 3 will include an overview of past research on the effects of bilingualism on language and cognitive outcomes. In Chapter 4, I will give a review of the literature on the development of mental state reasoning as well as a specific look at the research on mental state reasoning in bilingual groups. Chapter 5 will provide an explanation of the methodological challenges of comparing bilingual and monolingual groups as well as a justification for specific methodological choices for the current study. Chapters 6 and 7 will include the methods and results of the current study respectively, and Chapter 8 will include a discussion of the implications of those results as well as suggestions for future research.

The Relationship Between Language and Cognition

For many decades, researchers have explored the relationship between language and cognition. Through this exploration, two distinct perspectives have developed; 1- Language influences and is influenced by other areas of cognition and 2- Language develops and functions separately from other areas of cognition. While there are a number of researchers who hold strongly to one or the other of these two viewpoints, there are many whose views fall somewhere in between the two. Thus, in the current

chapter I will review some of the main questions and pieces of empirical evidence about the nature of the relationship between language and cognition.

Historical Background

Interest in questions about the relationship between language and cognition was sparked in the 1950s with the republication of Sapir and Whorf's work on the Linguistic Relativity Hypothesis (Sapir, 1929; Whorf, 1956). The claims made by Sapir and Whorf about the relationship between language and cognition are summarized by Brown (1976) in the following way:

> I Structural differences between language systems will, in general, be paralleled by nonlinguistic cognitive differences, of an unspecified sort, in the native speakers of the two languages.
>
> II The structure of anyone's native language strongly influences or fully determines the world-view he will acquire as he learns the language (p. 128).

Thus, according to the Linguistic Relativity Hypothesis, differences in linguistic structure either determine or strongly influence differences in non-linguistic cognition. Most of Whorf's evidence for the Linguistic Relativity Hypothesis came from his comparisons between speakers of English and speakers of Hopi, a Native-American language (1956). However, a major criticism of his work was that because his tasks were linguistic in nature, his data did not provide any direct evidence of an effect of language on non-linguistic cognition (Brown, 1976).

Despite interest in the Linguistic Relativity Hypothesis in the 1950s, support for this hypothesis waned in the 1960s with the rise of modularity theories of language development, which focused on the development of grammar. According to these theories, linguistic processes are universal and develop independently of other areas of

cognition. Major proponents of this view were Noam Chomsky (1967), Steven Pinker (1994) and Jerry Fodor (1983). For example, Chomsky (1967) emphasized the innate and universal aspects of language learning in order to explain how the vast majority of children learn grammar quickly and easily despite impoverished stimuli and lack of explicit feedback. Because of this emphasis on universal principles of language development, these theories left little room for different outcomes due to cross-linguistic differences. However, despite the skepticism for the Linguistic Relativity Hypothesis during the 1960s, recent years have seen a revival in questions about the nature of the relationship between language and cognition. Given the broad nature of the constructs *language* and *cognition*, it is not surprising that this exploration has broken into many sub-areas of inquiry. Thus, some of the primary ways in which this relationship has been explored will be reviewed below. Various ways in which researchers have explored language differences will be reviewed first, followed by a review of various cognitive outcomes due to those language differences.

Language Differences

Research with monolingual speakers that addresses questions about the impact of language on cognition has focused on two primary areas: 1-The influence on cognition of knowing one language versus knowing another language and 2-The influence of language development on cognitive development throughout childhood. Additionally, researchers have looked at the influence of knowing a language versus not knowing a language through research with non-human animals. Other questions have also been raised when considering bilingual populations such as whether there are cognitive differences between

individuals who know one language and individuals who know more than one language. Each of these areas will be discussed in turn.

Knowing One Language Versus Knowing Another

As mentioned earlier, Whorf's (1956) early work on the Linguistic Relativity Hypothesis focused on the cognitive differences between monolingual speakers of different languages. Similarly, as researchers have continued to explore the relationship between language and cognition, they have focused on the specific aspects (e.g. syntax, semantics etc.) of a language that affect the non-linguistic cognition of speakers of that language. Some of the primary questions explored in this area include how members of different language communities conceptualize shape categories (Roberson, Davidoff, & Shapiro, 2002), color categories (Davidoff, Davies, & Roberson, 1999; Özgen & Davies, 2002; Roberson, Davidoff, Davies, & Shapiro, 2005) and concepts of number (Gelman & Gallistel, 2004; Gordon, 2004; Gumperz & Levinson, 1997; Pica, Lemer, Izard, & Dehaene, 2004).

For example, members of various language communities divide the color spectrum in different ways based on the color terms used in that language. In research that examines these cross-linguistic differences, speakers of a language are asked to label and categorize different colors. Colors that are most consistently given the same label by native speakers are identified as the focal colors of that language. Through this work, researchers have shown that languages vary both in the number of color terms that are used as well as how the color spectrum is divided under those color terms. For example, English has eleven basic color terms; Himba (a language found in Namibia) has five basic terms (Roberson, et al., 2005 & Shapiro, 2005); and Dani (a language found in New

Guinea) has two basic terms, one term for light colors and one term for dark colors (Rosch, 1973). These color terms label varying subsets of colors within that language. For example, some languages include separate terms for green and blue while some languages label them with the same term (Roberson, et al., 2005).

One of the main questions in this line of research is whether variations in how colors are labeled and categorized affect the ease at which members of different language communities learn the color terms specific to their language, discriminate between different colors, and remember different colors. In an early study, Rosch (1973) provided evidence against the Linguistic Relativity Hypothesis by finding support for universal principles of color perception and categorization. In her study, she tested native speakers of Dani (who only use two basic color terms) on their ability to learn and remember various groups of colors. She found that speakers of Dani more easily learned and remembered color groupings when they were based on English focal colors than when they were grouped in other ways.

However, subsequent studies have provided evidence against universal principles of color perception and categorization. For example, Roberson et al. (2005) found that speakers of Himba were more likely to remember that they had seen a specific color previously when that color was a focal color in their language, and did not show greater recognition for English focal colors. Similarly, a variety of other researchers have found evidence of differences in color perception and categorization in various language communities (Kay & Kempton, 1984; Özgen & Davies, 2002; Roberson, et al., 2005). While these subsequent studies provide evidence for an influence of language on cognition, the nature of that influence is still hotly debated. In general, researchers have

rejected the strong Whorfian hypothesis (i.e. language determines thought), but many still adhere to a weak Whorfian hypothesis (i.e. language influences thought). Although many researchers hold to this idea, the particular nature and extent of the influence of cross-linguistic differences on cognition is still under considerable discussion.

Language Development and Cognitive Development

In addition to research that compares individuals from different language communities, there is also a substantial body of work that looks at how language and cognition interact as they develop within the same child. In this branch of work, researchers focus on how the acquisition of various components of one's native language affects cognitive development throughout childhood. One of the key questions in this area of inquiry is whether language is simply mapped on to and used to express different aspects of cognition or whether learning language enhances or alters cognitive development.

For example, Gentner (2003) argues that learning relational terms significantly drives or enhances children's ability to reason about abstract relationships between objects. In a series of studies, Gentner and Ratterman (Gentner & Rattermann, 1991; Rattermann & Gentner, 1998) explored how 3- to 5-year-old children's use of relational terms affects their ability to understand relational concepts. In these studies the child and the experimenter both had their own array of three objects that varied in size. The experimenter hid a sticker under one of her objects and instructed the child to find the sticker "in the same place" in the child's array of objects. In one condition the two sets of objects where identical. However, in another condition, the two sets of objects were different and the child was only able to find the sticker if he looked under the object that

was the same relative size (e.g. the biggest object) as the object in the experimenter's set. While 3- and 4-year old children could solve this task when the objects were identical, they performed poorly in the condition that required them to match based on relative size. However, children's performance drastically improved after they were trained on terms that denote relative size (e.g. big, little, and tiny). In fact, children who received training with these terms where able to solve this task when new object sets were presented and when the terms that denote relative size were no longer used. Thus, Gentner (2003) argues that learning language offers insight into how to think abstractly about the relationships between objects.

In addition to this evidence for the impact of language on relational thinking, some researchers have found evidence that learning mental state language enhances children's ability to reason about the mental states of others (Astington & Baird, 2005; P. L. Harris, De Rosnay, & Pons, 2005). For example, Astington and Jenkins (1999) found that language abilities were predictive of later performance in false belief tasks, but performance in false belief tasks was not predictive of later language abilities in preschool children. Likewise, there is evidence that learning certain syntactic constructions such as tensed complements (e.g. He thought he saw a unicorn) is important for developing an understanding of false belief (de Villiers & Pyers, 2002) and that training children in mental state terms and the syntactic constructions used to talk about mental states improves children's performance in false belief tasks (Lohmann & Tomasello, 2003). Furthermore, deaf children who are not exposed to sign language from birth are delayed in their performance in both standard (Peterson & Siegal, 2000)

and low-verbal (Figueras-Costa & Harris, 2001; Woolfe, Want, & Siegal, 2002) false belief tasks.

Although the studies mentioned above provide evidence that learning language impacts children's cognitive development, the direction of this relationship is not always clear. There is much debate about whether the development of specific language abilities are necessary for the development of cognitive abilities or visa versa (de Villiers, 2007). While it is possible that one is a necessary antecedent of the other, it is also possible that the relationship between language development and cognitive development is bidirectional. In other words, language skills and cognitive abilities may aid each other as they develop simultaneously within the same child. Thus, once again the key questions are how much and in what ways do language and cognitive development influence each other over time.

Knowing a Language Versus Not Knowing a Language

In order to compare the impact of the presence of language input on development, one would have to conduct the "forbidden experiment" by depriving a child of language in order to see how a lack of language leads to different cognitive outcomes. Although there are a few cases of children raised with very little language input, such as Genie and Victor (Curtiss, Fromkin, Krashen, Rigler, & Rigler, 1974), these cases offer little insight into this question because of the large number of unknown and confounding factors in each. While conducting the forbidden experiment with more careful controls in place may provide some answers to this question, this experiment will never be willfully conducted because of the extreme ethical concerns that it raises.

Some researchers have attempted to address the question of the impact of language on cognition through research with non-human animals. Typically in these experiments, researchers train animals to use aspects of linguistic and symbolic systems in order to see how this training changes the animals' ability to perform in a variety of cognitive tasks. Through this work, there is some evidence that language training does enhance the cognitive abilities of non-human animals. For example, Boysen and her colleagues conducted a series of experiments that explored the effect of language training on chimpanzees' performance in the less-is-more task (Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996). In this task, the chimpanzee is shown two arrays of treats that vary in quantity. The chimpanzee is then encouraged to point to one of the arrays and is rewarded the array that he or she did not point to. Chimpanzees consistently fail in this task because they persist in pointing to the array with the greater quantity of treats, thus receiving the array with the lesser quantity. However, chimpanzees that had been trained to use numeric symbols are able to point to the symbol representing the smaller quantity in order to receive the larger quantity of treats. This finding suggests that the symbolic nature of language aids the chimps in their ability to succeed in this task. Similar findings have been found with children in that children succeed more often in this task when symbols rather than actual quantities are used (Carlson, Davis, & Leach, 2005).

While some researchers suggest that teaching non-human animals rudimentary language systems leads to qualitative changes in animal cognition (Premack, 1983), others suggest that this language training only leads to quantitative differences in cognitive processing (Langer, 2000). One limitation of this work is that it is difficult to

demonstrate whether language training enhances the cognitive capacities of non-human animals, or whether language training merely enhances their ability to express the cognitive capacities that they already possessed (Tyack, 1993). Additionally, there are limitations in extending this work to humans because language learning may affect human cognition in different ways than it affects non-human cognition (Kuczaj & Hendry, 2003).

Knowing One Language Versus Knowing Two Languages

In addition to questions about the relationship between language and cognition in monolingual speakers, there are also questions about how learning more than one language leads to different cognitive outcomes. As mentioned previously, there is evidence that there are cognitive differences between individuals who learn one language and individuals who learn more than one language. For example, in the area of color perception, Caskey-Sirmons and Hickerson (1977) asked Korean-English bilingual and Korean monolingual speakers to label colors based on eleven basic Korean color terms. They found more variability among the bilingual than the monolingual speakers in which colors were labeled with each term and found differences between the focal colors of the bilingual and the monolingual speakers. One example was that monolingual speakers most consistently used the term *paran sekj* (blue) to describe a greenish-blue, while bilingual speakers used the same term to describe a purplish-blue. Thus, learning a second language affected the way in which Korean speakers categorized colors when using their native language. While this is just one example, there are numerous other studies on the cognitive differences between monolingual and bilingual speakers (see Bialystok, 2001, 2007; Bialystok, Martin, & Viswanathan, 2005 for reviews). Because

the cognitive differences between bilingual and monolingual speakers are the focus of this dissertation, this research will be covered in more detail in Chapter 3.

Cognitive Outcomes Due to Language Differences

In addition to the various ways in which language differences can be explored, researchers have also examined different cognitive outcomes based on those diverse language experiences. While some researchers have looked at how language influences the formation of concepts and categories, others have looked at how language influences cognitive processes. For example, evidence that language differences impact the formation of categories includes the color work previously discussed. The implication of this work is that the language that one speaks may impact whether, for example, one labels variations of green and blue with different color terms, and subsequently whether one thinks of these colors as being in different categories or not. However, in addition to the work on how language affects the formation of categories, researchers have also asked questions about whether differences between languages lead to differences in cognitive processes. For example, researchers have looked at how cross-linguistic differences in number terms affect an individual's ability to reason about exact quantities of objects. Along these lines, Gordon (2004) conducted research with speakers of Pirahã, a language that has only three terms for amounts: *one*, *two*, and *many*. In a series of tasks, he found that speakers of Pirahã struggled with remembering specific amounts greater than 8 to 10. Thus, a lack of numerical terms appears to affect these individuals' ability to reason about specific amounts. The way in which language differences are thought to affect cognitive processes will be covered in more detail in Chapter 3 as I more closely examine the cognitive differences between bilingual and monolingual

individuals. However, before I review these differences, it is important to discuss how researchers have defined the term “bilingual” in psychological research. Thus, the challenges and considerations of defining “bilingualism” will be the focus of the next chapter.

Chapter 2: Bilingualism

The Prevalence of Bilingualism

Within the United States, it is common to view monolingualism as the norm and bilingualism is the rare exception to that norm. Not only is this view prevalent in general U.S. society, but it has also been perpetuated by researchers in that a large body of language research has been conducted with monolingual samples. However, this research is misrepresentative of the language experiences of many individuals both within the U.S. and worldwide as it is common to gain some skills in a second language during one's lifetime (Bialystok, 2001). Language skills in a second language can be gained in a wide variety of ways. For example, some children learn both a local and a national dialect, or are exposed to a second language once they enter school. Likewise, both children and adults can be exposed to another language when they move permanently to a different language community (e.g. immigration, adoption) or live in a different language community for a short amount of time (e.g. study abroad, travel, international service). Similarly as trade, communication, and media become more globalized; individuals are more and more likely to come in contact with different language communities and thus have exposure to different languages.

Although there are a wide variety of individuals who are exposed to and gain some knowledge of a second language during their lifetime, it is difficult to gather statistics on the rate of bilingualism worldwide. This is largely the result of the difficulty of defining who is "bilingual," which is discussed in the remainder of this chapter. Although statistics for the rate of bilingualism worldwide do not exist, there are some estimates of the rate of bilingualism in various countries and regions. For example,

according to the 2000 U.S. Census, 18% of the total U.S. population speaks a language other than English in their home and 55% of these people reported that they spoke English "very well" (U.S. Census Bureau, 2003). However, this statistic probably underestimates the rate of bilingualism in the U.S. because it does not include individuals who have knowledge of a second language, but do not use that language in their homes. In comparison to the U.S., rates of bilingualism are estimated to be much higher in Europe and other countries. For example, Tabouret-Keller (2004) has estimated that the rate of bilingualism in Europe is as high as 50% of the population.

While the exact rates of bilingualism worldwide are unknown, many people do learn more than one language throughout their lifetime. Thus, theories about language development and the interaction between language and cognition should include these diverse language experiences. Researchers are certainly moving in that direction by studying bilingual samples more often, but a lot more work needs to be done in order to understand how bilingualism affects both language and cognitive processes. Additionally, while researchers have made some strides toward a better understanding of the effects of bilingualism on language and cognitive abilities, it is important to remember that there are individuals who learn three or more languages during their lifespan. Because most of the research on multilingualism has been done with bilingual samples, the primary focus of this dissertation will be the effects of learning two languages on cognition. However, I believe that by gaining a better grasp of the cognitive and language effects of bilingualism we will also move in the direction of gaining a better understanding of the various effects of a wider variety of language experiences.

The Nature of Bilingualism

In order to look at the relationship between bilingualism and cognition, one of the key questions we must ask is, *What does it mean to be bilingual?* Typically when we think of a bilingual individual, we think of someone who can understand and speak two languages. However, there are many individuals who have some knowledge of a second language that do not fit within that definition. For example, consider an individual who can read a second language quite well but lacks the ability to speak that language. Would we consider that person to be bilingual? Additionally, although we tend to think of bilingualism as a categorical variable, bilingualism is inherently a continuous variable because of the wide variety of language skills among different individuals (Bialystok, 2001). Thus, when we think about bilingualism we need to consider both the type of language proficiency (e.g. abilities in language production versus language comprehension, abilities in spoken versus written language, etc.) as well of the degree of language proficiency in each language of a given individual.

In addition to variations in type and degree of proficiency, bilingual individuals also vary a great deal in the sociolinguistic experiences that they have had. For example, bilinguals vary in which languages they are exposed to, the timing of exposure to those languages, the amount of exposure to those languages, how often they use each language, and the contexts in which they use each language. It is important that we pay attention to these differences in sociolinguistic experiences as they may contribute to language and cognitive outcomes in unique ways beyond language proficiency. Type of proficiency, degree of proficiency, and differences in sociolinguistic experiences will each be discussed in more detail below.

Type of Language Proficiency

Because there are so many different aspects of language knowledge and use, it is not surprising that an individual's level of proficiency will vary between these different sub-areas of language. For example, variations in exposure to and use of a language can impact a given individual's abilities in language comprehension and production; or in their ability to speak, read, or write in a language. Thus, when we measure an individual's level of proficiency in a language, we must remember that their level of proficiency will vary based on the type of language ability that we assess.

Another important consideration is that even within a specific sub-area of language ability, there are many different aspects of that area that could be assessed. For example, if we wanted to measure an individual's ability to understand spoken French, we could measure the number of words he understands, his understanding of different syntactic constructions, or his ability to understand pragmatic aspects of French. Although researchers interested in bilingualism typically focus on individuals who can both speak and understand two languages, there are vast differences in how they measure those language abilities. Thus, when we consider research that includes bilingual groups, we need to pay careful attention to the type of language skills that are being measured as well as how those language skills are being measured across various studies.

Degree of Language Proficiency

In addition to type of language proficiency, we need to consider how degree of language proficiency is used to define bilingualism. The continuous nature of language skills among bilingual individuals leads to challenges in conducting bilingual research because researchers must make decisions about who to include and who to exclude from

the bilingual category. According to Bialystok (2001) setting an arbitrary cut-off point of proficiency is problematic because it inevitably excludes individuals from the bilingual category who have some abilities in a second language. In order to avoid treating bilingualism as a categorical variable, researchers should: (1) include a variety of levels of language proficiency within and across studies, and (2) include measures of language proficiency that provide continuous rather than categorical data. In this way, researchers can gain a better understanding of how degree of proficiency is related to the dependent variables that they are studying.

Additionally, when measuring the degree of language proficiency in a given individual, we are often interested in how that individual's language skills compare to other speakers of that language. For example, if we measured an individual's vocabulary level in English, we would be interested in how her vocabulary skills compare to the vocabulary skills of other English speakers in her environment. While this seems fairly straightforward, it becomes more complicated when we look at bilingual individuals because they have language skills in more than one language. Thus, we can compare how a bilingual individual's language skills compare to other speakers of each language, and we can compare the relative level of language skills within that individual. For example, in many cases individuals are more proficient in one of their languages than the other (Ben-Zeev, 1977; Bialystok, 1988; Nicoladis & Genesee, 1997; Umbel, Pearson, Fernández, & Ollie, 1992). Although language abilities vary both within and across individuals, these factors are often confused or talked about simultaneously in bilingual research. However, a failure to separate these two factors could lead to misrepresentations of a bilingual's language abilities. For example, two different

individuals could have the same level of balance between their two languages, but one could have comparable language skills to monolinguals in both languages while the other's level of proficiency could be significantly below his monolingual counterparts. Thus, focusing on only one of these factors gives us an incomplete picture of the language abilities of the bilinguals whom we are studying.

An added complication to determining an individual's level of proficiency is that level of proficiency does not remain stable across the lifespan. For example, an individual can lose proficiency in a native language after he has moved to a new language environment. Conversely, he can regain unpracticed language abilities if he re-enters an environment in which that language is frequently used. It is particularly challenging to assign a level of language proficiency to bilingual children because children are in the *process* of acquiring language skills and thus their level of language proficiency is constantly changing (Bialystok, 2001). Although researchers often measure a bilingual child's level of language proficiency at the time of testing, we must keep in mind that that data point represents a small part of a long-term trajectory of language change and growth.

Sociolinguistic Factors

In addition to different types and levels of language proficiency, we must also consider variations in the sociolinguistic experiences of bilingual individuals. Of course, these factors are intrinsically tied to levels of language proficiency. For example, there is evidence that variations between bilingual individuals in the languages that they are exposed to, the timing of first being exposed to a language, the amount of language exposure, the duration of language exposure, and the context of language exposure are all

related to language proficiency (Holowka, Brosseau-Lapre, & Petitto, 2002; Petitto, et al., 2001). However, differences in sociolinguistic factors may also contribute to differences in language and cognitive abilities in unique ways beyond language proficiency. For example, bilingual children vary in how much they are exposed to various cultures, how much they participate in environments in which only one language is spoken or environments in which both languages are spoken, and how much they interact with monolingual or other bilingual speakers. Therefore, when studying the differences between bilingual and monolingual groups in their language and cognitive abilities we must consider how these sociolinguistic factors may be related to the dependent variables we are studying. Now that I have explored the challenges and considerations with defining "bilingualism," I can turn to the research that looks at the differences between bilingual and monolingual groups in their language and cognitive abilities.

Chapter 3: The Effect of Bilingualism on Language and Cognitive Abilities

Historical Background

During the early part of the twentieth century, the predominant opinion was that learning two languages would confuse children and lead to language deficits. In support of this viewpoint, researchers found evidence of language delays in bilingual children (see Arsenian, 1937; Darcy, 1953, 1963; Macnamara, 1966 for reviews) such as a lower vocabulary level (Barke & Perry-Williams, 1938; Grabo, 1931; Saer, 1923) and poorer grammatical abilities (C. W. Harris, 1948; Saer, 1923). In addition to these language deficits, there was also evidence of cognitive deficits in bilingual children. In one of the early studies on the topic, Saer (1923) compared bilingual and monolingual children's performance on the Stanford-Binet Intelligence Test and found that bilingual children performed worse than monolingual children overall. Due to this research, bilingualism was viewed as deleterious to both children's language and general cognitive development (see Arsenian, 1937; Diaz, 1983; Jensen, 1962 for reviews).

However, there were two major flaws in this early research (Darcy, 1953; Hakuta & Diaz, 1985). First, most studies did not control for the socio-economic differences between the bilingual and monolingual groups. Because the bilingual children were typically at a lower socio-economic level than the monolingual children, this put the bilinguals at a marked disadvantage. Secondly, the level of language proficiency of the bilingual groups was often not measured or controlled. Children were deemed to be bilingual based on where their parents were born, the foreignness of the parent's last name, or where the family currently lived (Darcy, 1953). For these reasons, early studies on the effects of bilingualism often included children who had limited exposure to or

competence in a second language (Diaz, 1983), which put the bilinguals at a disadvantage because most of the tests were administered in that second language (Hakuta, 1986). A failure to measure and control for the socioeconomic status and language proficiency of the bilingual samples compromised both the validity and the generalizability of these early studies. Therefore, most of these results are unusable today. However, since that time, researchers have made great strides in understanding the differences in the language and cognitive abilities between bilingual and monolingual groups by including careful controls for these factors.

The Effect of Bilingualism on Language Abilities

In marked contrast to early opinions about bilingualism, later work has provided evidence that bilingualism does not lead to overall language delays or deficits (see Petitto & Holowka, 2002 for a review). One of the biggest innovations in this line of research came through looking at the bilingual child's language skills across both languages instead of looking at each language separately. While, researchers have found that bilinguals tend to have smaller vocabularies than their monolingual peers in each language (Bialystok & Fang, 2009; Oller & Eilers, 2002; Perani, et al., 2003; Portocarrero, Burright, & Donovick, 2007), bilingual children's vocabulary level is comparable to monolinguals if you include their skills in both languages. For example, if you combine the total number of lexical items that bilingual children produce, bilinguals reach language milestones (e.g. first word, first fifty words, first two-word combinations, word spurt) at about the same age that monolingual children do (Holowka, et al., 2002; Pearson, Fernández, & Oller, 1993; Petitto, et al., 2001). It is important to note that in these studies, the bilingual children were not simply producing translational equivalents

in both languages. Instead, the total number of unique words that bilingual children produced followed the same developmental timetable as monolingual children.

In other work, researchers have found that bilingual individuals are slower in lexical access and retrieval than monolinguals (Bialystok, et al., 2008; Roberts, Garcia, Desrochers, & Hernández, 2002). This deficit appears to be related to the two lexicons that bilinguals must sort through in order to access the right word (Bialystok, 2009). However, there is evidence that negotiating two language systems contributes to a greater metalinguistic awareness in bilingual groups. For example, bilingual individuals show an advanced understanding of the arbitrary nature of language (Ben-Zeev, 1977; Cummins, 1978; Feldman & Shen, 1971; Ianco-Warrall, 1972), greater skills at making grammatically judgments (Bialystok, 1986, 1988; Bialystok & Majumder, 1998; Cromdal, 1999) and greater phonemic awareness (Bialystok, et al., 2003) than monolingual individuals. Because bilinguals show advantages in some areas and disadvantages in others, we cannot state that bilingualism leads to overall language deficits. Instead, differences between the language skills of bilingual and monolingual individuals are better conceptualized as a trade-off of language abilities.

The Effect of Bilingualism on Cognitive Abilities

Similar to language abilities, public opinion about the impact of bilingualism on cognition has significantly changed in recent decades. This change was sparked by a notable study by Peal and Lambert (1962) in which they compared bilingual and monolingual children's performance in a variety of intelligence tests. In marked contrast to earlier work, they controlled for the language proficiency of the bilingual group by using tests to identify children who were competent in both English and French.

Additionally, they controlled for the socioeconomic status of the bilingual and monolingual groups. To their surprise, the bilinguals outperformed the monolinguals, not just in the verbal tasks as they predicted, but in the majority of verbal and non-verbal tasks. Overall, they found that the bilingual children did particularly well in tasks that required mental or symbolic flexibility, which lead them to the conclusion that bilingualism fosters greater cognitive flexibility in children.

Many researchers (including Peal and Lambert themselves) have pointed out that the selection of the bilingual and monolingual samples for this study may have favored the bilingual group (Diaz, 1983). For example, the bilinguals were from a higher grade in school than the monolinguals and the bilingual group was limited to children who passed a certain level on the Peabody Picture Vocabulary Test (possibly excluding bilinguals with a lower intelligence level). However, despite these criticisms, Peal and Lambert's (1962) study provided an important first step toward more careful controls in bilingual research as well as changing public opinion about the effects of exposure to two languages on children's cognitive development.

Since that time, there continues to be a growing body of evidence that bilingual individuals demonstrate greater cognitive flexibility than monolingual individuals. Defining cognitive flexibility is difficult, given the wide variety of ways in which this term is used across different research studies. In general, cognitive flexibility has been used to describe bilingual individuals' ability to switch between different tasks and to perform well in tasks that include conflicting or misleading cues. For example, evidence that bilingualism fosters skills in task switching and in negotiating conflicting cues includes bilingual children's superior performance in the dimensional change card sort

task (Bialystok, 1999; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008). In this task, children are shown a series of cards one at a time that vary on two dimensions: shape and color (Zelazo & Frye, 1998). Children are first taught one version of the task in which they are asked to sort the cards into boxes based on one of the dimensions (e.g. color). Later, children are taught another version of the task in which they are asked to sort the cards by the other dimension (e.g. shape). Researchers have found that bilingual children between 4- and 5-years-old outperform monolingual children of the same age in this task (Bialystok, 1999; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008). This finding is thought to demonstrate bilingual children's greater cognitive flexibility in that they are better able to negotiate the conflicting aspects of the task and switch back and forth between sorting by shape and sorting by color.

Additional evidence for advanced skills in cognitive flexibility among bilingual individuals comes from their performance in the flanker task. In a typical flanker task, an individual sees a row of arrows and must respond by pressing the appropriate button to indicate which way the central arrow is pointing. Flanker tasks include both congruent (all the arrows point the same way) and incongruent (the center arrow points a different way from all the other arrows) trials. Costa, Hernández, Costa-Faidella, and Sebastian-Galles (2009) discuss two differences between the performance of bilingual and monolingual adults in this task. Bilinguals have been shown to respond faster than monolinguals in the incongruent trials, which is thought to demonstrate greater skills in negotiating conflicting cues. However, bilinguals have also been shown to respond faster than monolinguals when the task switches often between congruent and incongruent trials, which is thought to demonstrate greater flexibility in task switching.

In addition to the tasks described above, bilingual children have shown an advantage in a wide variety of other tasks that require skills in task switching and negotiating conflicting and misleading cues such as the Simon task (Martin-Rhee & Bialystok, 2008), the attentional network task (ANT) (Costa, et al., 2008), the stroop task (Bialystok, et al., 2008; Bialystok & Senman, 2004), and the reversing ambiguous figures task (Bialystok & Senman, 2004; Bialystok & Shapero, 2005). Also, bilingual individuals have shown greater abilities in cognitive flexibility at a wide variety of ages such as infancy (Kovács & Mehler, 2009), childhood (Bialystok, et al., 2008; Bialystok & Senman, 2004; Bialystok & Shapero, 2005; Martin-Rhee & Bialystok, 2008), adulthood (Bialystok, 2006; Colzato, et al., 2008; Costa, et al., 2008), and later adulthood (Bialystok, et al., 2008; Bialystok, Craik, & Ryan, 2006; Bialystok, Craik, Klein, & Viswanathan, 2004; Kavé, Eyal, Shorek, & Cohen-Mansfield, 2008).

For example, Kovács and Mehler (2009) trained 7-month-old infants to look toward one side of a screen in order to see a puppet when they heard a verbal cue. The researchers then switched the reward (seeing the puppet) to the other side of the screen. They found that in these post-switch trials, infants who had been exposed to two languages from birth looked more often at the second location while infants who had only be exposed to one language where more likely to continue to look toward the first location. The responses of the children exposed to more than one language from birth is thought to demonstrate greater cognitive flexibility because they were better able to inhibit their desire to continue to look toward the first location. Also, Bialystok (2008) found that for adults between the ages of 60 and 70, bilinguals and monolinguals perform similarly on working memory tasks while bilinguals outperform monolinguals in

executive control tasks. Furthermore, bilingualism tends to slow the deterioration of various cognitive skills (Bialystok, et al., 2008; Bialystok, et al., 2004). Thus, for bilingual individuals, superior abilities in cognitive flexibility may begin in infancy and continue across the entire lifespan.

Despite this evidence for greater cognitive flexibility in bilingual individuals, it is important to note that these findings do not go undisputed. For example Yang and Lust (2004) failed to replicate the bilingual advantage in the dimensional change card sort task and Morton and Harper (2007) reported that bilingual children do not show an advantage in the Simon task when socioeconomic status and ethnicity are more carefully controlled. Thus, one of the biggest unanswered questions with regards to these cognitive differences is why bilinguals outperform monolinguals in some cases and not others. As discussed in Chapter 2, there are a wide variety of differences among bilingual individuals. Thus, these diverse findings could be related to the specific bilingual and monolingual groups that are included in these studies. Also, different factors may influence the *magnitude* of these bilingual effects, which would lead us to find significant differences between bilingual and monolingual groups in some cases and not others. This issue will be discussed in more detail later in the current chapter, but in order to discuss factors that impact the magnitude of these effects it is useful to review hypotheses about the causes of these bilingual effects.

Causes of Bilingual Effects on Cognitive Abilities

One of the key questions in research with bilingual individuals is how they are able to select and use the appropriate language in different sociolinguistic situations. In addressing this question, researchers initially claimed that each language could be

"switched on" or "switched off" based on the language needed at the time (MacNamara & Kushnir, 1971; Penfield & Roberts, 1959). However, more recently, researchers have provided evidence that when doing a task in one language, both languages are activated (Grainger & Dijkstra, 1992; Guttentag, Haith, Goodman, & Hauch, 1984). Because of this evidence for the simultaneous activation of both languages, some researchers have suggested that bilingual individuals must develop skills to select the appropriate language and inhibit the inappropriate language to use in different situations (Green, 1998; Kroll & de Groot, 1997). Additionally, researchers have hypothesized that practice in selecting the appropriate language and inhibiting the inappropriate language impacts children's general inhibition and control processes which is why bilinguals show superior abilities in task switching and negotiating conflicting and misleading cues (Bialystok & Viswanathan, 2009).

For example, Costa et al. (2009) claim that bilingual individuals' superior performance in the Flanker task is driven by the practice that bilingual individuals have in managing their two languages. They posit that bilingual individuals' ability to respond faster in incongruent trials is related to the practice they receive in inhibiting the inappropriate language in various sociolinguistic situations. Additionally, they claim that bilingual individuals' ability to switch more easily between congruent and incongruent trials is related to the practice that bilinguals have in monitoring and selecting the appropriate language to use in different situations. Thus, they suggest that this language monitoring leads to a greater ability at monitoring when inhibition skills (i.e. skills needed to negotiate conflicting and misleading cues) are needed which is distinct from an advantage in inhibitory control.

While this is the most discussed hypothesis to explain the relationship between bilingualism and cognition, other researchers have offered other explanations. For example, Carlson and Choi (2009) have posited that because bilinguals develop an advanced metalinguistic awareness of the arbitrary nature of language, they have an advanced awareness of the symbolic nature of linguistic forms. This awareness is hypothesized to lead to greater skills in symbolic reasoning, which may contribute to their ability to perform well in tasks in which symbolic distancing aids with inhibition (see Carlson, et al., 2005). Others have suggested that because bilinguals constantly learn more than one label for the same object or concept, they may have more flexibility in how they conceptualize the world (Baker, 2006; Goetz, 2003). For example, Baker (2006) notes that in Welsh the word *ysgol* means *school*, but it also means *ladder*. Thus, because synonyms across languages have slightly different meanings, bilinguals might develop both an advanced metalinguistic awareness and an advanced ability to conceptualize things in different ways. While further research is needed to investigate these various hypotheses, an important point to acknowledge is that it is unlikely that there is a single cause for the cognitive differences between bilingual and monolingual groups. For example, the practice that bilingual individuals have in managing their two languages may lead to some cognitive outcomes while an advanced metalinguistic awareness may lead to others. Thus, it is possible that these different hypotheses explain different cognitive outcomes in bilingual individuals.

Magnitude of Bilingual Effects

One way in which we can better understand the causes of these bilingual effects is to look at factors that influence the magnitude of these effects. When considering the

magnitude of bilingual effects on cognition, two factors that have been considered are the degree of bilingualism (i.e. the level of language proficiency across both languages), and the amount of exposure that children have had to a second language. For example, Cummins (1976) proposed a model to explain why bilingual children outperform monolingual children in some studies while the reverse is true in others. According to Cummins' Threshold Model, bilingual children who are highly proficient in both languages will show cognitive advantages over monolingual children, bilingual children who are proficient in one language and not the other will show no differences from monolingual children in cognitive tasks, and bilinguals with a low level of proficiency in both languages will show cognitive deficits in comparison to monolingual children.

There is evidence that language proficiency across both languages does impact the magnitude of bilingual effects on cognition. For example, Ricciardelli (1992) divided French-English bilingual children into various subgroups based on their level of proficiency in both languages. She found that only the bilingual children who had a high level of proficiency in both French and English outperformed English monolingual children in a variety of cognitive tasks (including a task that measured divergent thinking and another that measured grammatical awareness). Although this work provides support for the Threshold model, there are several limitations to using this model to explain differences in the magnitude of the bilingual effects. First of all, the word "threshold" implies that there is a specific level of proficiency that must be reached in order for bilingual children to show advantages over monolingual children. However, with the wide variation in language proficiency among different bilingual individuals it is likely

that these effects are a matter of degree rather than some specific cut-off point of proficiency.

Another consideration is that because language proficiency is related to a variety of other variables, level of proficiency may not be the factor that mediates the magnitude of bilingual effects. For example, Costa and his colleagues (2009) have suggested that the amount of practice that bilinguals have in switching back and forth between different languages could be related to the magnitude of the bilingual effect on an individual's ability to switch back and forth between congruent and incongruent trials in the flanker task. They hypothesize that bilinguals who interact primarily with other bilinguals will be switching back and forth between the two languages more often, while bilinguals who use primarily one language in various settings (e.g. one language at home and a different language at work), will have less practice in language switching. Thus, in this case, the amount of practice that the bilingual individual has in switching back and forth between the two languages could be impacting both language proficiency and cognitive flexibility.

In addition to language proficiency, amount of experience with a second language has been shown to be related to the magnitude of these bilingual effects. For example, Carlson and Meltzoff (2008) found that children who had been exposed to Spanish and English from birth outperformed monolingual children in a battery of Executive Function tasks. However, native English-speaking children who had been enrolled in a half-day immersion program (in either Spanish or Japanese) for six months performed similarly to their monolingual peers. Additionally, Bialystok (1988) found that fully bilingual children (Canadian children who had been educated entirely in French and had some exposure to French at home) scored higher than partially bilingual children (children who

had attended a French immersion school for 2 years) in tasks that tested metalinguistic awareness.

Again, amount of exposure to a second language could be mediating the magnitude of these bilingual effects, but it could also point to a wide variety of other causal factors. This is because the amount of exposure to each language, degree of language proficiency across languages, and a variety of sociolinguistic factors are all related. Thus, while a variety of studies have looked at how "degree" or "level" of bilingualism are related to these various cognitive outcomes, researchers have yet to pinpoint the specific factors that are driving these bilingual effects as well as how variations in these factors impact the magnitude of these effects. Now that I have reviewed the literature on the differences between the language and cognitive abilities of bilingual and monolingual groups, I will turn to the literature that specifically explores bilingual children's ability to reason about mental states. In order to examine this literature, it is useful to gain an understanding of the development of mental state reasoning throughout childhood, and the tasks most commonly used to assess mental state reasoning. Thus, research on mental state reasoning will be the focus of the following chapter.

Chapter 4: Reasoning about the Mental States of Others

Children's ability to reason about the thoughts, beliefs, and desires of themselves and others, improves drastically between the ages of 3 and 5 (P. L. Harris, 2006). While often called a "Theory of Mind," this ability is currently called by a variety of terms including social-cognition, and mental state reasoning. In the past, children's ability to reason about mental states has been assessed primarily through false belief tasks. However, the role of false belief tasks as the definitive measure of mental state reasoning has been soundly criticized for several reasons (Bloom & German, 2000). First of all, in order to pass false belief tasks, children need more than an ability to think about the mental states of others as variations in memory abilities, inhibition, attention, and language abilities are all related to performance in these tasks. Secondly, there are many other aspects of mental state reasoning (such as understanding other's emotions) that are not captured in false belief tasks. Because of these and other criticisms, researchers have designed a wider variety of tasks in order to gain a broader view of how mental state reasoning develops in children during the preschool years. In order to give a background of research on mental state reasoning, I will begin this chapter by describing the false belief tasks most commonly used to assess mental state reasoning. Then, I will discuss the comparative performance of bilingual and monolingual groups in these false belief tasks. Finally, I will explain how researchers have developed additional tasks in order to gain a more comprehensive picture of children's ability to reason about mental states, and how the use of these tasks can help us gain a richer understanding of the differences between bilingual and monolingual children in mental state reasoning.

False Belief Tasks

There are three tasks that have been commonly used to assess children's ability to reason about mental states. All three tasks ask the child to distinguish between the actual state of the world and what someone believes or thinks about the state of the world (i.e. reasoning about a false belief). The first is known as the "Sally-Ann Task," but has also been called the "false belief", "Maxi", or "unexpected transfer task" (Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983). This task involves a character named Sally, who hides an object in one of two locations (e.g. a basket or a box). While Sally is out of the room, another character, named Ann, moves the object from one location to the other. Then Sally returns to the room, and the child is asked where Sally will look for the object. Three-year-old children typically state that Sally will look in the location where the object is currently located rather then the location where Sally saw it last. In contrast, 4- and 5-year-old children typically indicate that Sally will look in the location where she saw it last, demonstrating that they have some understanding of the difference between where Sally believes the object is located and where the object is actually located. In other words, they are able to recognize that Sally has a belief that is false (P. L. Harris, 2006).

A second task commonly used to assess the ability to reason about mental states is the unexpected contents task (Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987). This task involves a clearly marked box that contains something other than what the external packaging indicates should be in the box. In one version of this task, the child is shown a Smarties box and asked what she thinks is in the box. After the child indicates that she thinks that there are Smarties in the box, the box is opened to reveal

that it has pencils inside. The box is then closed and the child is introduced to a character and told that that character has never seen inside the box before. The child is asked what the character will think is inside the box. Three-year-old children typically respond that the character will think that there are pencils inside of the box. However, 4- and 5-year-old children typically respond that the character will think that there are Smarties inside of the box despite the fact that they know that there are pencils inside (P. L. Harris, 2006).

A third common false belief task is the appearance-reality task (Flavell, Flavell, & Green, 1983). In this task the child is shown an object that appears to be something different from what it actually is. In one version of the task, the child is shown a sponge that is painted to look like a rock, and asked what he thinks that the object is (Flavell, 1986). At that point, the child is allowed to touch the sponge and again asked what he thinks that it is. Following this, the child is asked a number of questions about the object's appearance known as appearance questions (e.g. What does the object look like?), and a number of questions about the true nature of the object known as reality questions (e.g. What is the object, really?). The child is also asked questions about what a character who had never touched the sponge before would think that it is. Once again, 3-year-olds are fairly unsuccessful at this task in that they indicate that a character who had never touched the sponge before would think that it is a sponge, while 4- and 5-year-olds typically respond that this character would think that it is a rock because it looks like a rock (P. L. Harris, 2006).

Bilingual Children's Performance in False Belief Tasks

As discussed previously, researchers have found evidence that bilingual children outperform monolingual children in false belief tasks. For example, Kovács (2009) tested 2- and 3-year-old Romanian-Hungarian bilingual and Romanian monolingual children in a standard unexpected transfer task (i.e. Sally-Ann task). She found that twice as many bilingual children passed the task as monolingual children. Similarly, Goetz (2003) tested 3- and 4-year-old Mandarin-English bilingual, Mandarin monolingual, and English monolingual children in an unexpected transfer task and an unexpected contents task. She found that the bilingual children outperformed both groups of monolingual children in the unexpected contents task. However, in contrast to Kovács findings, Goetz found no reliable differences between the bilingual and two monolingual groups in the unexpected transfer task.

Researchers have also found evidence of a bilingual advantage in appearance-reality tasks. Goetz (2003) found that English-Mandarin bilinguals outperformed Mandarin monolinguals in an appearance-reality task (the bilingual group also outperformed the English monolingual group, but the results only approached significance). Additionally, Bialystok and Senman (2004) compared 4- and 5-year-old bilingual children (from a wide variety of language backgrounds who all spoke English at school) with English monolingual children in an appearance-reality task. When language proficiency was controlled, they found that the two language groups performed similarly on the appearance questions (e.g. "What did you think this was when you first saw it?" and "Tigger didn't see or hear what we were doing. What will Tigger think this is?"), but

bilingual children outperformed monolingual children in the reality questions (e.g. "What is it really?").

While these findings could be interpreted to indicate that bilingualism fosters greater mental state reasoning abilities in children during the preschool years, there are alternate explanations to why bilinguals are outperforming monolinguals in these false belief tasks. For example, researchers have found a correlation between children's inhibitory control skills and their ability to pass false belief tasks (Carlson, Moses, & Breton, 2002). Thus, bilingual children's superior performance in these tasks may be due to their advanced skills in negotiating conflicting and misleading cues instead of an advanced ability to reason about mental states. In fact, one of the major criticisms of false belief tasks is that because they require a variety of different abilities to pass, it is difficult to pinpoint what the task is actually measuring. For this and other reasons, researchers have developed a wide variety of other tasks that assess mental state reasoning, which will be discussed in the following section. Additionally, comparing monolingual and bilingual children in a wider variety of mental state tasks would be useful because we could further investigate whether bilingualism contributes to an advantage in mental state reasoning abilities or whether these findings point to some other advantage in bilingual groups.

Moving Beyond False Belief Tasks

Since the development of the three most commonly used false belief tasks, there have been a multitude of studies using these tasks. However, as discussed at the beginning of this chapter, there has been a lot of criticism for using false belief tasks as the definitive measure of mental state reasoning abilities. One of the major criticisms is

that false belief tasks only assess a narrow aspect of mental state reasoning which is in fact a multifaceted and complex ability. Thus, through subsequent work, researchers have divided the development of mental state reasoning into several different understandings that children develop between infancy and the age of 5. These understandings include an understanding of goals and desires, an understanding of knowledge and knowledge access, an understanding of beliefs (including false beliefs), and an understanding of emotion (see P. L. Harris, 2006).

A Developmental Timeline of Mental State Reasoning Abilities

Although researchers have collected a substantial amount of data on the development of different aspects of mental state reasoning, a cohesive developmental story of how children come to develop an advanced ability to reason mental states was lacking. In order to address this problem, Wellman and Liu (2004) conducted a meta-analysis of research on children's ability to reason about mental states through which they identified seven unique mental state understandings, the age at which children typically demonstrate each understanding, and the tasks most commonly used to measure those understandings. These seven understandings were scaled in the following order (the task passed at the youngest age is listed first): an understanding of diverse desires, an understanding of diverse beliefs, an understanding of knowledge access, an understanding of contents false beliefs, an understanding of explicit false beliefs, an understanding of beliefs and emotion, and an understanding of the difference between real and apparent emotions. Children typically pass the easiest of these tasks around the age of 3 and pass all of the tasks around the age of 5. In addition to the meta-analysis,

Wellman and Liu (2004) conducted an experimental study with these seven tasks, which confirmed the scale that they had developed.

As mentioned previously, there is value in comparing bilingual and monolingual children's abilities in mental state reasoning through using a wider variety of tasks. Wellman and Liu's (2004) seven scaled mental state tasks may provide a useful tool through which to do this for several reasons. First of all, use of these tasks allows researchers to measure a variety of different aspects of mental state reasoning (e.g. reasoning about desires, beliefs, knowledge, and emotions), not just a child's ability to reason about false beliefs. Secondly, these tasks are scaled in a specific order, and the age at which children typically pass the various tasks is identified. Thus, by comparing monolingual and bilingual children in these tasks we can get an idea of whether bilingual children follow a similar developmental pattern to monolingual children in the development of their ability to reason about mental states. For these reasons, the current study includes a comparison of bilingual and monolingual children's performance in these tasks. However, in order to compare these two language groups in these tasks, there are several methodological issues that must be considered. Thus, the following chapter includes a discussion of the methodological challenges with comparing monolingual and bilingual groups and a justification of methodological choices that were made in order to compare a bilingual and a monolingual group in Wellman and Liu's (2004) scale.

Chapter 5: Methodological Challenges in Bilingual Research and a Justification of the Methodological Choices for the Current Study

In order to compare bilingual and monolingual children's performance in Wellman and Liu's (2004) scale, there are several methodological issues that must be addressed. One issue is the type of bilingual and monolingual children that will be included in the study. A major challenge in selecting bilingual and monolingual samples is that it is nearly impossible to find two groups that vary only in the number of languages that they speak and no other factors. This is because many bilinguals vary in systematic ways based on the conditions that lead to bilingualism in the first place such as moving to a different country, or being exposed to various cultures (Bialystok, 2001). Therefore, it is difficult to match bilingual and monolingual children on age, SES, language proficiency, and a variety of other factors simultaneously. Even if we were able to find a sample of bilingual children who matched monolingual children on all of the factors mentioned above, this sample would not be very representative of typical bilingual children thus limiting the generalizability of our study (see Carlson & Meltzoff, 2008).

Other major methodological challenges include how to measure language proficiency in each language within the bilingual group and how to select the language that will be used when testing both bilingual and monolingual children. Measuring language proficiency in both languages is often a challenge because of the limited number of standardized measures that have been developed to measure proficiency in non-English languages. Thus, the language proficiency of bilingual children is often only measured in one language, which has lead to some limitations in past work.

Additionally, because bilingual children are typically less proficient than monolingual children in each of the languages that they speak, selecting the testing language that will be used to test the bilingual children is a challenge because choices in testing language can lead to confounds in these studies. Each of these issues will be discussed in more detail below.

Measuring Language Proficiency in Both Languages

As discussed in the previous chapter, there is evidence that degree of bilingualism can impact the magnitude of the bilingual effect on cognitive abilities. However, the impact of language proficiency across both languages on various cognitive outcomes is not fully understood. This is partially because, in many studies, the language proficiency of the bilingual group has only been measured in one language or has not been measured at all. However, a failure to measure language proficiency across both languages is a major limitation to past work because differences in degree of bilingualism could explain why we find cognitive differences between some monolingual and bilingual groups and not others.

For example, one limitation of the three studies on bilingual children's performance in false belief tasks is the ways in which the language proficiency of the bilingual samples was measured. Kovács (2009) did not include any independent measure of language proficiency, but instead her bilingual sample was composed of children who were regularly exposed to two languages and had parents who spoke different languages. Goetz (2003) and Bialystok and Senman (2004) tested the bilingual children's vocabulary comprehension in English with the Peabody Picture Vocabulary Test (PPVT) , and Goetz (2003) translated the English PPVT into Mandarin in order to

test the bilingual group's vocabulary comprehension abilities in Mandarin. However, Goetz (2003) admits that because the Mandarin translation was not standardized, it is not a reliable measure of language proficiency. Because they lacked reliable measures for the bilingual child's language skills in both languages, Kovács (2009), Goetz (2003) and Bialystok and Senman (2004) could not look at how language proficiency across both languages was related to performance in false belief tasks.

Selecting the Testing Language

In addition to measuring language proficiency, selecting the testing language to use when comparing bilingual and monolingual groups is also a challenge. Different confounds are introduced based on whether you decide to test the monolingual and bilingual children in the same language or whether you decide to translate the tasks. Researchers have addressed this issue in a wide variety of ways namely: (1) Test all the children in the same language and control for language proficiency statistically (2) Translate the tasks and test each child in his/her most proficient language (3) Test the bilingual group twice, once in each language or (4) Minimize the verbal aspects of the tasks. None of these approaches eliminates all possible confounds when comparing bilingual and monolingual groups as each includes inherent strengths and weaknesses.

For example, there are different problems that arise when researchers decide either to translate or to not translate various tasks. Testing all of the children in the same language could be a problem because some of the bilingual children will be less competent in that language than their monolingual peers, which could underestimate their abilities. However, translating the tasks can introduce other problems. One translation issue is that you must be careful about word choice, as there is evidence that the use of

specific terms can impact children's performance in some tasks. For example, in Mandarin, there are three separate verbs that mean to think or to believe: *xiang*, *yiwei*, and *dang*. X*iang* has a neutral connotation, *yiwei* connotes a belief that may be false, and *dang* has a strong connotation of a mistaken belief. Lee, Olson, and Torrance (1999) tested a group of 3- and 4-year-old monolingual Mandarin speaking children in an unexpected contents, an unexpected transfer, and an appearance-reality task. They found that when the researcher used the verbs *yiwei* or *dang* children performed better than when the researcher used the neutral verb *xiang*. Similarly, some languages make distinctions between permanent and temporary characteristics of people and objects, which can improve children's performance in appearance-reality tasks. For example, Spanish contains two distinct verbs for the verb *to be*: *ser* which connotes properties and characteristics of objects and people that are more permanent and *estar* which connotes properties and characteristics that are more temporary. Sera, Bales and Pintado (1997) found that monolingual Spanish-speaking children outperformed monolingual English-speaking children in an appearance-reality task when the Spanish verb *ser* was used with the reality questions and *estar* was used with the appearance questions.

Challenges in selecting the testing language are illustrated by the following example. If a researcher wanted to compare a group of Mandarin-English bilinguals and English monolinguals in a false belief task, she may decide to test each child in his or her most competent language by developing both English and Mandarin versions of the tasks. However, when she translates the tasks into Mandarin, she would have to make decisions about which term to use when she talks about thoughts and beliefs. If she chose to use the Mandarin term *dang*, she would be providing the Mandarin speakers with a cue that

she is talking about a belief that is false, possibly giving the Mandarin speakers an advantage in the task. However, if she decided to use the neutral term *xiang*, the Mandarin speakers may assume that she is not talking about a belief that is false because she is not using the term *dang*. Lee, Olson, and Torrance's (1999) study provides direct evidence that these translational choices can affect how children perform in these tasks.

Another option would be to test all the children in English. However, this would introduce other issues because the English version of the task would include verb forms of *think* and *believe* that do not make distinctions between true and false beliefs as the Mandarin verbs do. Because the bilingual children know explicit markers of true and false beliefs in Mandarin, they may struggle when faced with a language that does not make these distinctions and may code and process these English terms differently than the monolingual group. Thus, one of the biggest problems in deciding whether to translate a task or not is that it is possible that the monolingual and bilingual children are coding and processing the verbal aspects of the task in different ways.

Some researchers have addressed the problem of translation by testing the bilingual children twice (once in each language) and testing the monolingual children twice in the same language. This was the strategy employed by Goetz (2003) in that she tested Mandarin-English bilingual children in two different forms of the same false belief tasks. She counterbalanced the children so that half of the bilingual children were tested in English first, while half of the bilingual children were tested in Mandarin first. She also tested a Mandarin monolingual group and an English monolingual group in the two forms of the tasks in the same language. However, the results of her study were mixed in that the bilingual group outperformed the monolingual groups in some tasks at one time

point but not the other. These mixed results appear to be due in part to practice effects for the monolingual, but not the bilingual group. Thus, a challenge with testing the children twice (once in each language) is that if bilingual children outperform monolingual children at one time point, but not the other, it is difficult to interpret these mixed results.

A final solution is to reduce the verbal components of the tasks. These low-verbal or non-verbal versions in many ways eliminate the complications of translation and comprehension. However, there are some limitations to using low-verbal or non-verbal versions of mental state tasks. First of all, it is difficult to compare children's performance in these tasks to past research because different versions of the tasks must be designed. Although there is some evidence that children pass non-verbal versions of mental state tasks at the same rate as verbal versions (Call & Tomasello, 1999), the researcher would need to compare non-verbal and verbal versions of any new task he designs in order to compare results to past research. Secondly, non-verbal versions of mental state tasks may contain different processing demands than verbal versions of the tasks (Call & Tomasello, 1999). Thus, when a researcher compares bilingual and monolingual children in mental state tasks, he may get different results for verbal and non-verbal versions of the tasks. Now that I have reviewed some of the methodological challenges in research with bilingual populations, I will provide a justification of the sample selected and the testing language used in the current study.

Justification for the Sample Selected

In order to compare bilingual and monolingual children in Wellman and Liu's (2004) scale, I decided to compare Spanish-English bilingual children to English

monolingual children in the United States. I selected this particular population for several reasons. First of all, a standardized version of the PPVT exists in Spanish, which is called the Test de Vocabulario en Imágenes Peabody or TVIP (L. W. Dunn, Lugo, Padilla, & Dunn, 1986). Because standardized versions of this task exist in both English and Spanish, I will be able to test the language proficiency of the bilingual group in both languages, which would allow me to investigate how the relative balance of skills across the two languages affects mental state reasoning abilities. Another advantage for selecting a Spanish-English bilingual sample is that this population is fairly representative of bilingual children within the United States as it has been estimated that 75% of the children who speak a language other than English in their homes speak Spanish (August & Hakuta, 1997; U.S. Census Bureau, 2003).

In order to control for differences between the bilingual and monolingual groups, I elected to match the two groups on age, gender, and the education level of the parents. However, I elected not to match the two groups on their level of vocabulary comprehension in English. Because monolingual samples typically outperform bilingual samples on measures of vocabulary comprehension in each individual language, I would have to include monolingual children who score atypically low, or bilingual children who scored atypically high in order to get matched samples. However, matching the two groups in this way would provide me with samples that are not representative of normative populations. Thus, I decided to control for level of vocabulary comprehension statistically in order to address these differences in language proficiency.

Justification for the Testing Language

As discussed in this chapter, there are both strengths and weaknesses with deciding to translate or to not translate the mental state tasks or to developing low-verbal versions of these tasks. In order to be able to compare my findings to Wellman and Liu's (2004) findings, I elected to use the same verbal versions of the tasks that they used instead of low-verbal versions of the tasks. However, this left the question of whether the tasks should be translated into Spanish or not. Because of the added complications of translation (linguistic effects of particular word choice) I elected to administer the tasks to all of the children in English. While this does eliminate possible effects of word choice for the Spanish versions of the tasks, it could put the bilingual children who are less proficient in English at a disadvantage. For this reason, I elected to use level of English vocabulary comprehension as a covariate in the data analysis in order to control for the potential effect of the testing language.

Chapter 6: Research Methods

Participants

Twenty-six Spanish-English bilingual children and twenty-six English monolingual children participated in the study. The bilingual sample included twelve 3-year-olds, three 4-year-olds, ten 5-year-olds, and one 6-year-old (mean age = 4;6, range = 3;0 to 6;2; female = 14, male = 12). The monolingual sample included twelve 3-year-olds, three 4-year-olds, ten 5-year-olds, and one 6-year-old (mean age = 4;5, range = 3;0 to 6;4; female = 14, male = 12). The bilingual children were recruited from local schools, the University of Minnesota's Como Student Housing, personal contact by the researcher, and through referrals of other research participants. Similarly, the monolingual children were recruited from local schools, the University of Minnesota's Como Student Housing, the Institute of Child Development Participant Pool, and through personal contact by the researcher.

In order to find bilingual children who had some skills in speaking both languages, the researcher asked the director of a local bilingual preschool as well as the instructor of a Kindergarten class at a local monolingual school to identify children who spoke both English and Spanish. Likewise, when asking for referrals, the researcher asked the parents of bilingual children to identify other children who could speak both English and Spanish. Also, an advertisement was put in the University of Minnesota's Como Student Housing asking for the participation of bilingual children who spoke both English and Spanish.

After identifying children who had some skills in both English and Spanish, children's vocabulary levels were tested with the Peabody Picture Vocabulary Test

(Third Edition) and the Test de Vocabulario en Imágenes Peabody. The PPVT is a standardized test of English vocabulary comprehension abilities (L. W. Dunn, Dunn, & Williams, 1997) and the TVIP is an equivalent standardized test of Spanish vocabulary comprehension abilities (L. W. Dunn, et al., 1986). Children were excluded from the study if they tested below the lowest possible level on the PPVT (raw score = 22, age equivalent score = 1;9) or the TVIP (raw score = 3, age equivalent score = 2;6).

Bilingual children who scored above the lowest level on the PPVT and the TVIP tended to be children of parents who had attended college. Of the 26 bilingual participants, 22 of them had at least one parent who completed some college. Five additional bilingual children were excluded from the study for the following reasons. Two of the bilingual children were excluded because their age equivalent score on the PPVT was below 1;9, one bilingual child was excluded because his/her age equivalent score on the TVIP was below 2;6, and two bilingual children were excluded because they did not complete the second session. The monolingual sample was matched to the bilingual sample in gender, age, and educational level of the parents, which was assessed through a parental questionnaire. Of the monolingual participants, all of them had at least one parent who completed some college. Three additional monolingual children were excluded from the study for the following reasons. Two of the monolingual children were excluded because of experimenter error and one child was excluded because he/she was regularly exposed to a language other than English.

Procedures

Each bilingual child completed the PPVT, the TVIP, and a battery of seven mental state tasks. All tasks were administered to the bilingual children by the same

researcher who is fluent in both English and Spanish. The bilingual children completed the tasks during two sessions and were tested primarily in their homes although some of them were tested on the campus of the University of Minnesota. During the first session, the bilingual children were tested on the PPVT and the TVIP and were allowed to select whether they wanted to complete the "Spanish book" or the "English book" first. During the second session, the bilingual children completed the seven mental state tasks, which were given in a randomized order. The first and second sessions were given about a week apart.

Each monolingual child completed the PPVT and the seven mental state tasks. The monolingual children were tested by the same researcher who tested the bilingual children although some were tested by a second researcher. They were either tested in their homes or on campus and were given all tasks during the same session. They completed the mental state tasks first followed by the PPVT, and were given the mental state tasks in a randomized order. The parents of both the bilingual and monolingual children were given a brief questionnaire about the child's exposure to English as well as other languages (Appendix A). The questionnaire was translated into Spanish (Appendix B) and the parents of the bilingual children could select whether they wanted to complete the English or the Spanish version. The questionnaire and its translation were developed by Carlson and Meltzoff (2008).

The mental state tasks included the seven tasks used by Wellman and Liu (2004); which included a Diverse Desires task, a Diverse Beliefs task, a Knowledge Access task, a Contents False Belief task, an Explicit False Belief task, a Belief-Emotion task, and a Real-Apparent Emotion task. Each task is described in more detail below. The tasks

were given following the procedures described in Wellman & Liu (2004) except that in some cases the character names were changed or the toy animal used was a different animal. Additionally, in order to reduce the likelihood that children would pass the mental state tasks by chance, a second version of each of Wellman and Liu's tasks was designed (described below). The tasks were given to the child in one of 7 different randomized orders using a 7 x 7 Latin Square design with the two versions of each task always being presented one after the other. Each child was given a score between 0-2 for each aspect of mental state reasoning (1 point for passing Wellman and Liu's version of the task and 1 point for passing the second version of the task) and a composite score between 0-14 for their overall performance in all seven tasks. Both the bilingual and monolingual children were given the mental state tasks in English in order to reduce effects due to translation. Each task is described in detail below. The descriptions of the Wellman and Liu versions of the tasks are in the original words of Wellman and Liu (2004).

Diverse Desires

Wellman & Liu Version

Children saw a toy figure of an adult and a sheet of paper with a carrot and a cookie drawn on it. "Here's Mr. Jones. It's snack time, so, Mr. Jones wants a snack to eat. Here are two different snacks: a carrot and a cookie. Which snack would you like best?" This is the *own-desire* question.

If the child chose the carrot: "Well, that's a good choice, but Mr. Jones really likes cookies. He doesn't like carrots. What he likes best are cookies." (Or, if the child chose the cookie, he or she is told Mr. Jones likes carrots.) Then the child was asked the

target question: "So, now it's time to eat. Mr. Jones can only choose one snack, just one. Which snack will Mr. Jones choose? A carrot or a cookie?"

To be scored as correct, or to pass this task, the child must answer the *target* question opposite from his or her answer to the *own-desire* question.

Second Version

The same procedure was followed except that children are shown a paper with a block and a kite drawn on it and the following changes were made to the script. "Here's Mr. Jones. It's playtime, so, Mr. Jones wants to play with a toy. Here are two different toys: a block and kite. Which toy would you like best?" This is the *own-desire* question.

If the child chose the block: "Well, that's a good choice, but Mr. Jones really likes kites. He doesn't like blocks. What he likes best are kites." (Or, if the child chose the kite, he or she is told Mr. Jones likes blocks.) Then the child was asked the *target question*: "So, now it's time to play. Mr. Jones can only choose one toy, just one. Which toy will Mr. Jones choose? A block or a kite?"

Diverse Beliefs

Wellman and Liu Version

Children saw a toy figure of a girl and a sheet of paper with bushes and a garage drawn on it. "Here's Linda. Linda wants to find her cat. Her cat might be hiding in the bushes or it might be hiding in the garage. Where do you think the cat is? In the bushes or in the garage?" This is the *own-belief* question.

If the child chose the bushes: "Well, that's a good idea, but Linda thinks her cat is in the garage. She thinks her cat is in the garage." (Or, if the child chose the garage, he

or she is told Linda thinks her cat is in the bushes.) Then the child is asked the *target* question: "So where will Linda look for her cat? In the bushes or in the garage?"

To be correct the child must answer the *target* question opposite from his or her answer to the *own-belief* question.

Second Version

The same procedure was followed except that children are shown a paper with a bedroom scene on it and the following changes were made to the script. "Here's Linda. Linda wants to find her socks. Her socks might be in this drawer or they might be under the bed. Where do you think her socks are? In the drawer or under the bed?" This is the *own-belief* question.

If the child chose the drawer: "Well, that's a good idea, but Linda thinks her socks are under the bed. She thinks her socks are under the bed." (Or, if the child chose the bed, he or she is told Linda thinks her socks are in the drawer). Then the child was asked the *target* question: "So where will Linda look for her socks? In the drawer or under the bed?"

Knowledge Access

Wellman and Liu Version

Children saw a nondescript plastic box containing a small plastic toy bear inside. "Here's a box. What do you think is inside this box?" (The child could give any answer he or she liked or indicate that he or she did not know). Next, the box was opened and the child was shown the contents of the box: "Let's see…it's really a bear inside!" Close the box: "Okay, what is in the box?"

Then a toy figure of a girl named "Julia" was produced: "Julia has never ever seen inside this box. Now here comes Julia. So, does Julia know what is in the box? (the *target* question) "Did Julia see inside this box?" (the *memory* question).

To be correct the child must answer the *target* question "no" and answer the *memory* control question "no."

In Wellman and Liu's original version, they used a small plastic drawer and a toy dog. In this version a small box was used and a toy bear.

Second Version

The second version was administered the same way as the first version except that a new box was used that had a toy zebra inside.

Contents False Belief

Wellman & Liu Version

The child saw a clearly identifiable Band-Aid box with a plastic toy elephant inside the closed Band-Aid box. "Here's a Band-Aid box. What do you think is inside the Band-Aid box?" Next, the Band-Aid box was opened: "Let's see…it's really a elephant inside!" The Band-Aid box was closed: "Okay, what is in the Band-Aid box?"

Then a toy figure of a boy named Peter was produced: "Peter has never ever seen inside this Band-Aid box. Now here comes Peter. So what does Peter think is in the box? Band-Aids or an elephant? (the *target* question) "Did Peter see inside this box?" (the *memory* question).

To be correct the child must answer the *target* question "Band-Aids" and answer the *memory* question "no."

In Wellman and Liu's original version, a toy pig was used. In this version a toy elephant was used.

Second Version

The same procedure is followed except that children were shown a crayon box with a toy giraffe inside and the following changes were made to the script. "Here's a Crayon box. What do you think is inside the Crayon box?" Next, the Crayon box is opened: "Let's see…it's really a giraffe inside!" The Crayon box is closed: "Okay, what is in the Crayon box?"

Then a toy figure of a boy named Peter was produced: "Peter has never ever seen inside this Crayon box. Now here comes Peter. So what does Peter think is in the box? Crayons or a giraffe? (the *target* question) "Did Peter see inside this box?" (the *memory* question).

Explicit False Belief

Wellman & Liu Version

Children saw a toy figure of a boy and a sheet of paper with a backpack and a closet drawn on it. "Here's Daniel. Daniel wants to find his gloves. His gloves might be in his backpack or they might be in the closet. *Really,* Daniel's gloves are in his backpack. But Daniel *thinks* his gloves are in the closet."

"So, where will Daniel look for his gloves? In his backpack or in the closet?" (the *target* question) "Where are Daniel's gloves really? In his backpack or in the closet?" (the *reality* question).

To be correct the child must answer the *target* question "closet" and answer the *reality* question "backpack."

In Wellman and Liu's original version, the experimenter said "mittens". However, because mittens are less common in Latin American countries the experimenter said "gloves".

Second Version

The same procedure was followed except that children were shown a sheet of paper with a kitchen scene and the following changes are made to the script. "Here's Daniel. Daniel wants to find his favorite cup. His cup might be in his dishwasher or it might be in the drawer. *Really,* Daniel's cup is in the dishwasher. But Daniel *thinks* his cup is in the drawer."

"So, where will Daniel look for his cup? In his dishwasher or in the drawer?" (the *target* question) "Where is Daniel's cup really? In his dishwasher or in the drawer?" (the *reality* question).

To be correct the child must answer the *target* question "drawer" and answer the *reality* question "dishwasher."

Belief-Emotion

Wellman & Liu Version

Children saw a toy figure of a boy and a clearly identifiable individual-size Cheerios box with rocks inside the closed box. "Here is a Cheerios box and here is John. What do you think is inside the Cheerios box?" (Cheerios) Then the adult made John speak: "John says, 'Oh good, because I love Cheerios. Cheerios are my favorite snack. Now I'll go play.'" John was then put away out of sight.

Next, the Cheerios box was opened and the contents were shown to the child: "Let's see…there are really rocks inside and no Cheerios! There's nothing but rocks." The Cheerios box was closed: "Okay, what is John's favorite snack?" (Cheerios).

Then John came back: "John has never ever seen inside this box. Now here comes John. John's back and it's snack time. Let's give John this box. So, how does John feel when he gets this box? Happy or sad? (the *target* question) The adult opened the Cheerios box and let the toy figure look inside: "How does John feel after he looks inside the box? Happy or sad? (the *emotion-control* question).

To be correct, the child must answer the *target* question "happy" and answer the *emotion-control* question "sad."

Second Version

The same procedure was followed except that children were shown a clearly identifiable popcorn box with scraps of paper inside and the following changes were made to the script. "Here is a popcorn box and here is John. What do you think is inside the popcorn box?" (popcorn) Then the adult made John speak: "John says, 'Oh good, because I love popcorn. I love popcorn even more that Cheerios. Popcorn is my favorite snack. Now I'll go play.'" John was then put away out of sight.

Next, the popcorn box was opened and the contents were shown to the child: "Let's see…there is really paper inside and no popcorn! There's nothing but paper." The popcorn box was closed: "Okay, what is John's favorite snack?" (popcorn).

Then John came back: "John has never ever seen inside this box. Now here comes John. John's back and it's snack time. Let's give John this box. So, how does John feel when he gets this box? Happy or sad? (the *target* question) The adult opened

the popcorn box and let the toy figure look inside: "How does John feel after he looks inside the box? Happy or sad? (the *emotion-control* question).

Real-Apparent Emotion

Wellman & Liu Version

Initially, children saw a sheet of paper with three faces drawn on it: a happy, a neutral, and a sad face. The experimenter than pointed to each face in turn and asked the child "What does this face look like to you?" If children did not produce the labels "happy", "o.k.", and "sad"; the experimenter labeled the faces for the children. Then that paper was put aside, and the task began with the child being shown a cardboard cutout figure of a boy drawn from the back so that the boy's facial expression could not be seen. "This story is about a boy. I'm going to ask you about how the boy really feels on the inside and how he looks on his face. He might really feel one way inside but look a different way on his face. Or, he might really feel the same way inside as he looks on his face. I want you to tell me how he really feels inside and how he looks on his face."

"This story is about David. David's friends were playing together and telling jokes. One of the older children, Ellen, told a mean joke about David and everyone laughed. Everyone thought it was very funny, but *not* David. But, David didn't want the other children to see how he felt about the joke, because they would call him a baby. So, David tried to *hide how he felt*." Then the child gets two memory checks: "What did the other children do when Ellen told a mean joke about David?" (Laughed or thought it was funny.) "In the story, what would the other children do if they knew how David felt? (Call David a baby or tease him).

Pointing to the three emotion pictures: "So, how did David really feel, when everyone laughed? Did he feel happy, sad or okay?" (the *target-feel* question) "How did David try to look on his face, when everyone laughed? Did he look happy, sad or okay? (the *target-look* question).

To be correct the child's answers to the *target-feel* question must be more negative than his or her answer to the *target-look* question (e.g. sad for target-feel and happy or okay for target-look, or okay for target-feel and happy for target-look).

Second Version

The same procedure was followed except that the following story was told. "This is another story about David. David had a birthday party and invited all of his friends. David was really excited for his friend Sara to come to his birthday party. He started opening all the presents from his friends. He opened the present from Sara and it was some socks. David did not want socks for his birthday party, but he didn't want Sara to know that he didn't like the present because she would cry. So, David tired to *hide how he felt*." Then the child gets the following memory check: "In the story, what would Sara do if she knew how David felt?" (Cry or be sad).

Pointing to the three emotion pictures: "So, how did David really feel, when he opened the present from Sara? Did he feel happy, sad or okay?" (the *target-feel* question) "How did David try to look on his face, when he opened the present from Sara? Did he look happy, sad or okay? (the *target-look* question).

Chapter 7: Results

In order to compare the bilingual and monolingual groups' performance in the vocabulary tasks as well as the mental state tasks, several analyses were conducted, which will be described in this chapter in the following order. First of all, I compared the English and Spanish vocabulary scores within the bilingual group as well as the English vocabulary scores across the bilingual and monolingual groups. Secondly, I compared the children's performance in the Wellman and Liu versions of the tasks to the second version of each task. Additionally, in order to compare the order of difficulty in the current study to the order of difficulty found by Wellman and Liu (2004), I examined the task order of difficulty for the seven mental state tasks. Third, I calculated correlations between the raw vocabulary scores and performance in the mental state tasks for the monolingual and bilingual groups and looked at the predictive nature of age and vocabulary scores on mental state task performance through regression analyses. Fourth, I compared the bilingual and monolingual groups' performance in the seven mental state tasks. Fifth, I separated the bilingual children into groups based on their level of vocabulary comprehension in both languages and the monolingual children into groups based on their level of vocabulary comprehension in English and examined how the mental state task scores differed between the groups. Finally, I examined how the questionnaire data was related to performance in the mental state tasks within the bilingual group.

English and Spanish Vocabulary Scores

The results for the bilingual children's performance in the English vocabulary task (PPVT) and Spanish vocabulary task (TVIP) are reported in Table 1. The TVIP

standardized scores were calculated using the tables for the Hispanic Composite Scores, because the parents of the children in the bilingual sample where born in a variety of different Spanish-speaking countries. Because the raw scores for the PPVT and the TVIP are on different scales, they were not directly compared. However, the bilingual group's English and Spanish standardized scores were compared through a paired sample t-test which revealed no significant differences between the scores $t(25) = -1.09$, $p = .286$, $d = .28$. The English and Spanish standardized scores for the bilingual group were not significantly correlated with each other $r = .165$, $p = .422$. This result is likely due to the fact that some of the bilingual children scored high in one language and low in the other.

The results for the monolingual children's performance in the PPVT are also reported in Table 1. An independent samples t-test revealed that the monolingual group's raw English vocabulary scores were significantly higher than the bilingual group's $t(50) = 2.64$, $p = .011$, $d = .75$. Likewise, the monolingual group's standardized English vocabulary scores were significantly higher than the bilingual group's $t(50) = 4.556$, $p < .000$, $d = 1.29$. Because the bilingual group had a significantly lower level of English vocabulary comprehension than the monolingual group and because the mental state tasks were administered in English, children's vocabulary comprehension score (PPVT) was included as a covariate in subsequent analyses.

Order of Difficulty for the Mental State Tasks

Before I examined the order of difficulty of the mental state tasks, I wanted to verify that the two versions of each task were comparable. Thus, I conducted a chi-square analysis for each individual task in order to assess whether children who passed one version of the task, also tended to pass the other version of the task. This analysis

revealed a significant relationship between the two versions of each tasks: Diverse Desires, Pearson χ^2 (1, N = 52) = 10.84, p = .001, Cramer's V = .456; Diverse Beliefs, Pearson χ^2 (1, N = 52) = 9.25, p = .002, Cramer's V = .422; Knowledge Access, Pearson χ^2 (1, N = 52) = 28.84, p = .000, Cramer's V = .745; Contents False Belief, Pearson χ^2 (1, N = 52) = 40.98, p = .000, Cramer's V = .888; Explicit False Belief, Pearson χ^2 (1, N = 52) = 34.67, p = .000, Cramer's V = .817; Belief-Emotion, Pearson χ^2 (1, N = 52) = 33.68, p = .000, Cramer's V = .805; Real-Apparent Emotion, Pearson χ^2 (1, N = 52) = 9.57, p = .002, Cramer's V = .429. In other words, children who passed the Wellman and Liu version of each task were also highly likely to pass the second version of each task.

Table 2 lists the means for the bilingual and monolingual groups' performance in each mental state task and the ranking from easiest to most difficult based on these means. The task order of difficulty for the monolingual group exactly replicated the task order found in Wellman and Liu (2004). For the bilingual group, the task order was replicated for the three easiest tasks. However, the bilingual group differed from the monolingual group in that they scored higher on the Belief-Emotion task and the Real-Apparent Emotion task than they did on the Contents False Belief task and the Explicit False Belief task.

Additional analyses were conducted to compare the children's performance in each task within each language group. Thus, a one-way ANOVA was conducted for the monolingual group comparing their performance on each mental state task. This revealed a significant difference between the tasks F (6, 175) = 3.935, p = .001, partial η^2 = .119. Post-hoc tests (Tukey) showed that children's performance on the Diverse Desires task (the easiest task) differed significantly from their performance on the Explicit False

Belief, Belief-Emotion, and Real-Apparent Emotion tasks (the three hardest tasks) as shown in Table 3. Additionally, the children's performance in the Diverse Beliefs task (the second easiest task) differed significantly from the Real-Apparent Emotion task (the hardest task). Overall, this pattern of results provides evidence that, for the monolingual group, the tasks became more difficult in the order specified in Wellman and Liu (2004).

A one-way ANOVA was conducted for the bilingual group in order to compare their performances on each individual mental state task, which also revealed a significant difference between the tasks $F(6, 175) = 8.418$, $p < .000$, partial $\eta^2 = .224$. Post-hoc tests (Tukey) revealed that for the bilingual group, children's performance in the Diverse Desires task differed from their performance in all other tasks (Table 4). This finding was distinct from the monolingual group because the monolingual group's performance in the Diverse Desires task only differed reliably from the three hardest tasks. Additionally, in contrast to the monolingual group, the bilingual group's performance in the Explicit False Belief task was significantly worse than their performance in the Diverse Beliefs and Knowledge Access tasks. These differences confirm the finding that the order of difficulty for the tasks was different in the bilingual and monolingual groups.

Relationship Between Performance in the Mental State Tasks and Vocabulary Scores

In order to assess the relationship between the performance in the different mental state tasks and the vocabulary tasks, correlations between these variables were calculated. Table 5 and Table 6 include matrices of the correlations between age (in months), raw PPVT score, raw TVIP score, and the scores for the seven mental state tasks for the monolingual and bilingual groups respectively. Table 7 includes the differences in these

correlations between the bilingual and monolingual groups. There are several patterns to note in these differences. First of all, PPVT score and age are highly correlated in the monolingual group $r = .891$, $p < .000$. However, PPVT score and age $r = .684$, $p < .000$ as well as TVIP score and age $r = .567$, $p = .003$ were not as highly correlated in the bilingual group. Secondly, the Contents False Belief and the Belief-Emotion tasks were more highly correlated with the other mental state tasks for the monolingual than the bilingual group. These differences in the correlations between the mental state tasks in the bilingual and monolingual groups are likely due to differences between the factors that impact the development of different aspects of mental state reasoning in the two language groups. In other words, scores on the various mental state tasks were likely more correlated in the monolingual group because the different aspects of mental state reasoning are driven by the same factors in this group. However, scores on the false belief and emotions tasks were less correlated with other aspects of mental state reasoning in the bilingual group because the ability to reason about false beliefs and emotions are likely influenced by additional factors beyond what is driving general mental state reasoning in the bilingual group.

In order to assess whether age (in months), language group (bilingual or monolingual), and English vocabulary score significantly predicted the variance in the mental state task scores, a regression analysis was conducted (Table 8). This analysis revealed that these three factors (age, language group, and PPVT score) significantly predicted performance in all of the mental state tasks except the Diverse Beliefs task. Furthermore, language group significantly predicted the variance in mental state task scores for the Diverse Desires, the Belief-Emotion, and the Real-Apparent Emotion tasks

(negative β-values in Table 8 for language group indicate that membership in the bilingual group significantly predicted a higher score in the mental state tasks). Similar regression analyses were also conducted for the monolingual and the bilingual groups separately (Tables 9 and 10). For the monolingual group, age and raw PPVT score significantly predicted the composite mental state task score $R^2 = .441$, $p = .001$. However, neither age nor PPVT score significantly predicted unique variance in the composite mental state task score. This is not surprising given that PPVT and age were highly correlated in the monolingual group $r = .891$, $p < .000$.

For the bilingual group, age, raw English and raw Spanish vocabulary score did not significantly predict the composite mental state task score $R^2 = .201$, $p = .168$. Overall, these findings suggest that while age and English vocabulary score are predictive of performance in the mental state tasks within the monolingual group, other factors beyond age and vocabulary score are contributing to the differences in performance in the mental state tasks within the bilingual group. One other finding of note is that within the bilingual group; age, PPVT score, and TVIP score all contributed to a unique proportion of the variance in the Explicit False Belief task. I will return to this finding later when I discuss how the vocabulary scores in both English and Spanish are related to performance in the mental state tasks within the bilingual group.

Comparing the Two Language Groups in Mental State Task Performance (without considering language proficiency)

In order to analyze the children's performance in the mental state tasks a series of ANOVAs were conducted. First of all, a series of one-way ANOVAs revealed that for the children's performance in the mental state tasks, there were no significant effects for

gender or for task order so these variables were excluded from further analyses. Secondly, to compare the bilingual and monolingual groups on their mental state task performance, an ANOVA was conducted with Language Group (bilingual, monolingual) as the independent variable and the composite mental state task score as the dependent variable (Table 11). This ANOVA revealed that the two language groups' did not differ significantly in their composite mental state task scores. To compare the two language groups' performance in each mental state task individually, a MANOVA was conducted with Language Group (bilingual, monolingual) as the independent variable and the mental state task scores as the dependent variables (Table 12). This analysis revealed a significant difference between the two language groups $F(7, 44) = 2.822$, $p = .016$, partial $\eta^2 = .31$. Pairwise comparisons using the Bonforroni correction for multiple pairwise comparisons (because there are seven tasks, we would need a p-value of .05/7 = .007 to be significant) revealed no significant differences between the two language groups on any of the tasks. However, the difference between the two language groups' performance on the Diverse Desires task with the bilingual outperforming the monolingual group revealed a medium effect size $F(1, 40) = 4.551$, $p = .038$, $d = .60$. Likewise, the difference between the two language groups' performance on the Explicit False Belief task with the monolingual outperforming the bilingual group also revealed a medium effect size $F(1, 40) = 2.769$, $p = .055$, $d = .56$.

The Role of Language Proficiency

There was a strong correlation between raw English vocabulary score and the composite mental state score across both language groups $r = .542$, $p < .000$. For the bilingual group, the raw English vocabulary score was significantly correlated with the

composite mental state score r = .435, p = .026 while the raw Spanish vocabulary score was not r = .192, p = .348. Because the bilingual group's English vocabulary scores were significantly lower than the monolingual group's and because the mental state tasks were administered in English, it is important to control for English language proficiency as it may be masking group differences in mental state task performance. Thus, an ANCOVA was conducted with language group as the independent variable, the composite mental state task score as the dependent variable, and English vocabulary score as the covariate (Table 11). This analysis revealed no significant difference between the bilingual and monolingual groups.

In order to compare the two language groups' performance in each of the individual mental state tasks, a MANCOVA was conducted with language group as the independent variable, the mental state task scores for each task as the dependent variables, and English vocabulary score as the covariate (Table 12). This analysis revealed a significant difference between the two groups' performance in the mental state tasks $F(7, 43) = 5.06$, $p < .000$, partial $\eta^2 = .45$. Pairwise comparisons (using the Bonforroni correction for multiple pairwise comparisons) revealed that the bilingual group outperformed the monolingual group in the Diverse Desires task $F(1, 49) = 9.75$, $p = .003$, $d = .91$. Although the following tasks failed to show a significant difference with the Bonforroni correction, they did reveal medium to large effect sizes with the bilingual group outperforming the monolingual group: Knowledge Access $F(1, 49) = 7.23$, $p = .010$, $d = .78$; Belief-Emotion $F(1, 49) = 7.23$, $p = .010$, $d = .78$; and Real-Apparent Emotion $F(1, 49) = 4.87$, $p = .032$, $d = .64$.

In order to look at how language proficiency was related to mental state task performance within the bilingual group, the bilingual children were broken into subgroups based on their PPVT and TVIP standardized scores. Children were considered to be proficient in English if their standardized PPVT score was at or above 100 (the normed mean for this test). Similarly, children were considered to be proficient in Spanish if their standardized TVIP score was at or above 100. Based on their level of proficiency in each language, the bilingual children were broken into four subgroups: high proficiency in both languages, low proficiency in both languages, low proficiency in Spanish and high proficiency in English, and high proficiency in Spanish and low proficiency in English (Table 13). In a similar way, the monolingual children were broken into a high proficiency and low proficiency group based on their standardized score in English (Table 14).

A MANCOVA was conducted for the bilingual children with proficiency group (high-high, low-low, high Spanish-low English, low Spanish-high English) as the independent variable, the mental state task scores as the dependent variables, and age (in months) as the covariate. This analysis revealed a significant difference between the groups $F(21, 44) = 2.018$, $p = .025$, partial $\eta^2 = .478$. Pairwise comparisons (using the Bonforroni correction) revealed a significant difference between the four groups on the Explicit False Belief Task $F(3, 21) = 11.12$, $p < .000$, partial $\eta^2 = .614$. Additionally, although the results were not significant with the Bonforroni correction, the differences between the four groups showed a medium effect size for the Real-Apparent Emotion Task $F(3, 21) = 4.03$, $p = .021$, partial $\eta^2 = .365$. Thus, this analysis revealed that language proficiency across both languages as measured by vocabulary score was related

to children's performance in the Explicit False Belief and the Real-Apparent Emotion tasks.

Post-hoc analyses (Tukey) for the Explicit False Belief task revealed that the High-High group outperformed every other group in this task. There were no other significant differences between any of the other groups. Visual inspection of the data revealed that while children in the High-High group scored from 0-2 on this task (mean = 1.29), all other children in the other 3 groups scored a 0 on this task. Given this distribution of scores, it is not surprising that age, PPVT score, and TVIP score each predicted a significant proportion of the variance in performance in the Explicit False Belief task in the regression analysis for the bilingual group. Post-hoc analyses also revealed that for the Real-Apparent Emotion task the High-High group (mean = 1.57) outperformed the High Spanish-Low English group (mean = .25). There were no other significant differences between the other proficiency groups.

In a similar manner, a MANCOVA was conducted for the monolingual children with proficiency group (high English, low English) as the independent variable, the mental state task scores as the dependent variables, and age (in months) as the covariate. This analysis revealed no significant difference between the groups $F(7, 17) = .681$, $p = .686$, partial $\eta^2 = .219$. However, this result may have been due to the uneven sample sizes for the two proficiency groups within the monolingual sample.

Questionnaire Data

The questionnaires provided demographic information on both the bilingual and monolingual groups. However, because the questionnaire data was very similar within the monolingual group, analyses were only done with the questionnaire data from the

bilingual group. Of the twenty-six bilingual children, twenty-four of them where born in the United States, while one child was born in Peru and one child was born in Columbia. Additionally, twelve children had parents who both spoke Spanish in the home, three had parents who both spoke English in the home, and eleven had parents who spoke a combination of English and Spanish in the home.

Exposure to and Use of Each Language

The language that each bilingual child was exposed to or used (either English, Spanish, or both) was coded for each of the following questions: the language the mother speaks in the home (question 6), the language the father speaks in the home (question 6), the language the child speaks with his/her mother (question 7), the language the child speaks with his/her father (question 7), the language the child speaks with his/her siblings (question 9), the language the child speaks with his/her friends (question 10), the language of the television and radio programs in the home (question 11), the language of the print media in the home (question 12), the language the child is read to (question 14), and the language the child reads (question 16). Based on these questions, each bilingual child was given a score for his/her exposure to or use of Spanish and a score for his/her exposure to or use of English. For example, each time the parent indicated that the child was exposed to or used Spanish, 1 point was added to that child's total Spanish Exposure score. If on a given question, the parent indicated both languages (e.g. that the child uses both Spanish and English with her friends) 1 point was added to both the child's total English and total Spanish Exposure score. The English Exposure Score ranged from 0 to 11, and the Spanish Exposure Score ranged form 1 to 12.

There was a significant negative correlation between the bilingual children's English Exposure Scores and Spanish Exposure Scores when age (in months) was controlled r = - .601, p = .001. Additionally, the correlation between the English Exposure score and the raw English vocabulary score approached significance when age was controlled r = .387, p = .056, as did the correlation between the Spanish Exposure score and the raw Spanish vocabulary score r = .388, p = .055 when age was controlled. The English and Spanish Exposure Scores were not significantly correlated with mental state task performance scores on any of the tasks when age was controlled. These results appear to indicate that there is a trade-off between exposure to English and exposure to Spanish, and exposure to each language appears to be related to vocabulary score in that language. However, it is not surprising that the English and Spanish exposure scores were not correlated with mental state task performance, as each score only indicates the child's exposure to (and proficiency in) each languages separately.

Use of One Language or Both Languages with Various Groups of People

As suggested by Costa et al. (2009) the amount of experience that a bilingual individual has in interacting with either bilingual or monolingual speakers might impact the magnitude of the effect of bilingualism on cognition. Although the number of monolingual and bilingual speakers that the child interacts with or amount of time that the children spend talking to various speakers was not directly assessed, parents were asked to indicate the languages that the child used with different people or groups of people. Thus, based on the questionnaire data, we can get a rough idea of which children are using only one language as opposed to using both languages with different people. In order to roughly assess how much the bilingual children were interacting with

monolingual speakers, each child was given a score for the number of times the parents indicated that he/she used only one language on the following questions: the language that the child uses with his/her mother (question 7), the language that the child uses with his/her father (question 7), the language the child uses with his/her siblings (question 9), and the language the child uses with his/her friends (question 10). For example, if the parent answered either "English" or "Spanish" on one of these questions then 1 point was be added to that child's score. However, if the parents answered "English and Spanish" on one of these questions then no points were added to that child's score. This score was positively correlated with children's performance on the Real-Apparent Emotion task when age (in months) was controlled $r = .476$, $p = .016$. In other words, children who were more likely to use only one language with these various people/groups of people were doing better on the Real-Apparent Emotion task. This score was not correlated with performance in any of the other mental state tasks.

Summary of Results

From these analyses, there are several main findings. First of all, the monolingual group's English vocabulary level was significantly higher than the bilingual group's. Secondly, the order of difficulty of the mental state tasks for the monolingual group replicated the order of difficulty found in Wellman and Liu (2004). However, the bilingual group passed the two tasks that measured their ability to reason about emotions at a higher rate than would be expected based on Wellman and Liu's (2004) findings. Third, while age and English vocabulary score were highly predictive of differences in mental state task performance for the monolingual group, this was less true for the bilingual group. Fourth, effect sizes suggest that the bilingual group outperformed the

monolingual group in the Diverse Desires task and the monolingual group outperformed the bilingual group in the Explicit False Belief task when English language proficiency was not controlled. However, the bilingual group significantly outperformed the monolingual group in the Diverse Desires when English language proficiency was controlled. Furthermore, effect sizes suggest that the bilingual group outperformed the monolingual group in the Knowledge Access, the Belief-Emotion task, and the Real-Apparent Emotion task when English language proficiency was controlled. Finally, findings from the questionnaire data suggest that relative interaction with monolingual versus other bilingual speakers is related to children's understanding of the emotions of others. The implications of these findings will be discussed in the following chapter.

Chapter 8: Discussion

The purpose of the current study was to investigate whether bilingualism has a broader effect on mental state reasoning than has been previously found. Although past research has provided evidence that bilingual children outperform monolingual children in false belief tasks (Bialystok & Senman, 2004; Goetz, 2003; Kovács, 2009), it was unknown whether this advantage would extend to other areas of mental state reasoning. In order to address this question, Spanish-English bilingual children were compared to English monolingual children in Wellman and Liu's (2004) seven scaled mental state tasks. Additionally in order to address limitations of past research, the level of vocabulary comprehension was assessed in both Spanish and English and used as an index of language proficiency. Findings from the current study provide evidence that bilingualism has a broader impact on mental state reasoning than has been previously found. Furthermore, findings from the current study offer some insights into how differences in sociolinguistic experiences and language proficiency might be related to mental state reasoning in bilingual groups. Each of these findings will be discussed in detail below.

Language Proficiency of the Monolingual and Bilingual Groups

As discussed in Chapter 3, bilingual children tend to show a lower vocabulary level than their monolingual peers when you consider their skills in each language separately (Bialystok & Fang, 2009; Oller & Eilers, 2002; Perani, et al., 2003; Portocarrero, et al., 2007). Thus, because the bilingual and monolingual groups in the current study were matched on age, gender, and the educational level of their parents, it is not surprising that the bilingual group had a significantly lower English vocabulary level

than the monolingual group. Additionally, it is not surprising that age and vocabulary score were highly correlated in the monolingual group, but not as correlated in the bilingual group as both the PPVT and the TVIP are normed with monolingual samples. In other words, because the bilingual children's vocabulary abilities are spread across two languages, there was more variation in how the bilingual children scored on the vocabulary tests thus leading to less of a correlation between age and vocabulary scores for each test individually.

As discussed in Chapter 3, in order to get a true picture of bilingual children's language proficiency, we need to look at their language skills across both languages instead of looking at each language separately. Although considering the bilingual children's standardized scores in both the PPVT and the TVIP in the current study allowed us to do that to some extent, we must remember that both of these measures were designed to be used with monolingual samples. Thus, combining the two measures does not give us a true picture of the bilingual children's overall vocabulary comprehension abilities. Furthermore, while skills in vocabulary comprehension were assessed in the current study, other language skills could be measured in order to examine how language proficiency is related to performance in mental state tasks. For example, measures of language production as well as measures of other aspects of language such as syntax could be assessed in bilingual groups. Despite the limitations in the current methods for assessing language proficiency in the bilingual group, the findings from the current study did offer insight into how level of language proficiency across both languages influences mental state understanding. This will be discussed in more detail later in this chapter.

The Effect of Bilingualism on Mental State Reasoning Abilities

The current study provides two pieces of evidence that bilingualism has an effect on the development of mental state reasoning. First of all, the regression analysis revealed differences in how much variance in mental state reasoning was accounted for by various factors in the monolingual and bilingual groups. For the monolingual group, age and English vocabulary score were highly correlated with each other and significantly predicted the variance in the mental state task scores. These findings are not surprising given past findings that language proficiency is highly correlated with performance in mental state tasks (Cutting & Dunn, 1999; Pons, Lawson, Harris, & de Rosnay, 2003). Furthermore, given the high correlation between age and vocabulary score it is not surprising that age and language proficiency explained very little of the unique variance on their own.

While age and vocabulary score were highly predictive of mental state task performance in the monolingual group, these variables did not significantly predict the variance in performance in the mental state tasks for the bilingual group. These findings suggest that additional factors beyond age and language proficiency in each individual language are contributing to differences in mental state reasoning in the bilingual group. As discussed previously, differences in language proficiency *across* the two languages (i.e. degree of bilingualism), sociolinguistic experiences, and sociocultural experiences in bilingual groups may all contribute to these differences. For example, the results from the questionnaire data suggest that a child's relative interaction with either bilingual or monolingual speakers may impact that child's ability to reason about the emotions of

others. Clearly, further research is needed to explore which aspects of bilingualism are contributing to this variance in mental state reasoning abilities.

In addition to the regression analyses, differences between the bilingual and monolingual groups' performance in each of Wellman and Liu's (2004) mental state tasks provides evidence that bilingualism impacts mental state reasoning abilities. Specifically, the bilingual children outperformed the monolingual children in the Diverse Desires task when English vocabulary level was controlled. Furthermore, effect sizes suggest that the bilingual children also outperformed the monolingual children in the Knowledge Access, the Belief-Emotion, and the Real-Apparent Emotion tasks when English vocabulary level was controlled. These finding suggests that bilingualism contributes to a broader effect on mental state reasoning than has been found in the past. In addition to overall differences between the bilingual and monolingual groups in Wellman and Liu's (2004) tasks, there are two specific findings that warrant further discussion. First of all, while the order of difficulty found by Wellman and Liu (2004) was replicated in the current monolingual sample, the bilingual children passed the two emotions tasks at a higher rate than would be expected based on Wellman and Liu's findings. Secondly, the three tasks that bilingual children failed to show an advantage in were the Diverse Beliefs, the Contents False Belief, and the Explicit False Belief tasks which all require children to reason about the beliefs of others. Each of these specific trends will be discussed below.

Bilingual Children's Performance in the Emotions Tasks

In the current study, the bilingual children showed a particular advantage in the two mental state tasks that tested their understanding of other's emotions: the Belief-

Emotion task, and the Real-Apparent Emotion task. This was shown by two pieces of evidence. First of all, bilingual children passed the two emotions tasks at a higher rate than they passed the Contents False Belief and Explicit False Belief tasks, while the opposite was true for the monolingual group. Additionally, for the monolingual group, performance on the Belief-Emotion task was more highly correlated with their performance in the other mental state tasks than it was for the bilingual group. This suggests that perhaps other factors are influencing the development of the ability to understand the emotions of others beyond the factors that lead to a general ability in mental state reasoning in the bilingual group.

Two possible explanations for why bilingual children showed a particular advantage in the emotions tasks are cross-linguistic and the cross-cultural differences in the way in which emotions are conceptualized, expressed, and understood. First of all, researchers have found evidence in monolingual samples of a relationship between children's knowledge of the semantics and syntax used to talk about emotion and their understanding of emotion (J. Dunn, Brown, & Beardsall, 1991). Similar to cross-linguistic differences in how shape, color, and number are expressed, there are also cross-linguistic differences in how various linguistic communities talk about and interpret terms that express emotion (Pavlenko, 2006). For example, Sachs and Coley (2006) presented a series of stories to Russian monolingual, English monolingual, and Russian-English bilingual adults. In half of the stories (the envy stories) a character wanted something that his/her friend possessed. In the other half of the stories (the jealousy stories) a character suspected that his/her romantic partner was being unfaithful. Participants from all three language groups were asked whether the term *envious*

(*zaviduet* in Russian) or the term *jealous* (*revnuet* in Russian), among other emotion terms, were appropriate to describe how the protagonist felt in the story. The Russian monolingual participants rated *zaviduet* (*envious*) as being much more appropriate to describe how the protagonist felt in the envy stories and *revnuet* (*jealous*) as being much more appropriate to describe the how the protagonist felt in the jealousy stories. Similar to the Russian monolinguals, the English monolinguals rated the term *jealous* as being more appropriate to describe how the protagonist felt in the jealousy stories, although the effect was less strong than it was for the Russian monolinguals. However, the English monolinguals rated the terms *envious* and *jealous* as being equally appropriate to describe how the protagonist felt in the envy stories. Interestingly, Russian-English bilingual participants responded similarly to the Russian monolinguals when they were told the stories in Russian, and similarly to the English monolinguals when they were told the stories in English. This provides evidence that bilingual individuals interpret or think about emotions differently based on the language that they are using at the time.

Similar to the Sachs and Coley (2006) study, the majority of cross-linguistic research on emotional expression looks at how bilingual speakers interpret and express emotions differently based on the language that they are using at the time. While there is ample evidence for these cross-linguistic differences (Pavlenko, 2008), one question that should be asked is whether cross-linguistic differences in how emotions are interpreted and expressed affect bilingual individuals' general ability to reason about emotions. In other words, perhaps interpreting emotions in two distinct ways based on the language that one is using leads to an advanced understanding of emotional expression.

In addition to cross-linguistic differences, bilingual children may have an advanced ability to reason about emotions because of cross-cultural differences in how emotions are discussed and displayed (Pavlenko, 2008). Similar to other aspects of mental state reasoning, there is both evidence for (P. L. Harris & Gross, 1988; P. L. Harris, Olthof, Terwogt, & Hardman, 1987) and against (Vinden, 1999) a universal timeframe for developing an understanding of emotions across cultures. Thus, the bilinguals in the current study could be outperforming the monolinguals in the emotions tasks for two possible reasons. One would be that Spanish-speaking cultures foster a greater understanding of emotions than English-speaking cultures. However, another explanation could be that being exposed to two different cultures (with different emotional display rules etc.) uniquely contributes to a greater understanding of emotional expression.

The findings from the questionnaire data in the current study appear to support the second explanation. If exposure to more Spanish-speakers (i.e. Hispanic culture) were driving the bilingual children's superior performance in the emotion tasks, then we would predict that their Spanish-exposure score would be related to performance in these tasks, but it was not. However, bilingual children who used only one language as opposed to two languages with more groups of people (i.e. their parents, siblings and friends) were more likely to perform well in the Belief-Emotion task. It is possible that children who interact with more monolingual speakers are simultaneously interacting with more mono-cultural individuals. Thus, these children could be doing better in the Belief-Emotion task because they are more likely to see greater differences in how emotions are displayed and interpreted. In other words, children who interact with more mono-cultural

individuals could have a greater awareness of other's emotions due to the cross-cultural differences in emotional expression.

The finding that performance in the Belief-Emotion task is more highly correlated with performance in the other mental state tasks in the monolingual than the bilingual group suggests that some aspect of bilingualism contributes to the understanding of the emotions that is different from factors that contribute to general mental state reasoning abilities in the bilingual group. Perhaps, cross-cultural differences in emotional expression are the unique factors that are contributing to these differences. While it is difficult to separate the influence of cross-linguistic and cross-cultural differences in bilingual groups, certain methods can be used to try to separate out the effects of these two factors. For example, a more thorough and detailed questionnaire could help us better understand the nature of the individuals with whom the child is interacting. Furthermore, comparing monolingual groups that vary in their language and cultural backgrounds to bilingual groups could offer insights into which factors are impacting the understanding of other's emotions in bilingual groups.

Bilingual Children's Performance in the Others' Beliefs Tasks

The finding that the bilingual children in the current study did not outperform the monolingual children in the Diverse Beliefs, the Contents False Belief, and the Explicit False Belief tasks when English vocabulary level was controlled is surprising for two reasons. First of all, this finding is surprising because effect sizes suggest that bilingual children in the current study outperformed the monolingual children in all other mental state tasks. Secondly, this finding is surprising given that Kovacs (2009), Goetz (2003), and Bialystok and Senman (2004) all found evidence of a bilingual advantage in false

belief tasks. These diverse findings may be due, in part, to the culturally specific aspects of the tasks or the testing language used to test the children in the current study, or they may be due to differences in the degree of bilingualism of the groups included in the various studies. Each of these possibilities will be discussed in turn.

First of all, because the tasks in Wellman and Liu's (2004) scale were originally validated with an American monolingual English-speaking sample, it is possible that some aspects of the tasks are culturally specific. For example, one explanation for the bilingual children's performance in the Contents False Belief task is that this task included a Band-aid box and a Cherrios box that may have been less familiar to some of the bilingual children. Wellman and Liu's scale has been used to test children in different cultures. For example, Australian children (Peterson, Wellman, & Liu, 2005), and Chinese children (Wellman, Fang, Liu, Zhu, & Liu, 2006) have both been tested using the scale. Results from these studies indicate that the Australian children passed the tasks in the same order as the American children (Peterson, et al., 2005). However, while the Chinese children passed the tasks in the same general order as the American children, they passed the Knowledge Access task at a higher rate than the Diverse Beliefs task. These results suggest that there are some cross-cultural differences in the order of the development of mental state reasoning abilities. Furthermore, in these studies some aspects of the tasks were altered in order to be culturally relevant to the groups being tested. For example, when the Chinese sample was tested (Wellman, et al., 2006), a potato chip tube was used instead of a band-aid box for the Contents False Belief task. Thus, it is possible that the bilingual children in the current sample did not outperform the monolingual sample in the other's beliefs tasks because of cultural differences between

the bilingual and monolingual groups or because of the culturally specific items used in the tasks. These hypotheses could be tested by comparing monolingual children from Hispanic countries to the bilingual and monolingual groups in the current study, or by changing some of the culturally specific aspects of the tasks.

Another factor that may explain the bilingual children's performance in the other beliefs task are the specific English mental state words used in these tasks. For example, the Diverse Beliefs, the Contents False Belief, and the Explicit False Belief tasks all include the word *think* when talking about someone else's thoughts or beliefs. As discussed in Chapter 5, word choice can impact how children perform in various tasks. For example, Shatz and her colleagues (2003) found evidence that Puerto Rico Spanish speakers performed better in false belief tasks when a verb that explicitly marks for false belief (*creer-se*) was used than when the unmarked form of the verb (*creer*) was used. Additionally, she found that monolingual Spanish-speaking children outperformed monolingual English-speaking children in a false belief task, but only when the marked forms were used in Spanish. Thus, the bilingual children in the current study may have had a difficult time identifying that someone's belief was false in tasks that included the word *think* because different amounts of certainty cannot be marked linguistically in English as they are in Spanish. However, this explanation must be taken with caution as the study with Shatz and her colleagues (2003) was conducted with Puerto Rican Spanish speakers and it is unclear how this word choice would impact the performance of speakers of other dialects of Spanish. However, it is possible that in the current study the bilingual children had a more advanced understanding of false belief then they were able to demonstrate because they were tested in English. This hypothesis could be tested by

assessing Spanish-English bilingual children in Spanish versions of the belief tasks or by testing them in low-verbal versions of the tasks.

In addition to the specific wording, there is evidence that the degree of bilingualism mattered in these tasks. When the bilingual children were split into subgroups based on their level of vocabulary comprehension in both English and Spanish, performance on the Explicit False Belief task was shown to differ reliably between the groups. Post-hoc analyses revealed that the group of children who had a high level of proficiency in both English and Spanish outperformed every other group in this task. Because previous research did not include reliable measures of the children's language proficiency in both languages, it is difficult to compare this finding to past work on false belief tasks. However, these findings suggest that the magnitude of the effect of bilingualism on understanding the beliefs of others is related to the level of language proficiency across both languages.

One question that should be asked at this point is why language proficiency in both languages was related to performance in the Explicit False Belief task and the Real-Apparent Emotion task, but not any of the other tasks. One possible explanation is that degree of bilingualism affects performance in a wide variety of mental state reasoning tasks, but there was not enough variation in degree of bilingualism in the current study to show this broader effect. Another explanation is that degree of bilingualism has a larger effect on some aspects of mental state reasoning than others. For example, it is possible that degree of bilingualism affects the development of inhibition skills and thus impacted performance in the mental state reasoning tasks to various degrees based on the level of

inhibition that was needed in each task. Further research will be needed to test these and other hypotheses.

Does Bilingualism Lead to an Advantage in Mental State Reasoning Abilities?

While both the regression analyses and performance in the mental state tasks suggest that bilingualism impacts mental state reasoning, one question we should ask is whether bilingualism really leads to an *advantage* in mental state reasoning. While bilingual groups have been shown to outperform monolingual groups in mental state and other tasks when language proficiency is controlled, bilingual and monolingual children typically perform similarly when language proficiency is not controlled. As pointed out by Carlson and Meltzoff (2008), this similar performance is notable in and of itself because based on bilingual children's lower level of language proficiency, one would expect them to perform significantly *worse* than monolingual children in these tasks. This is true of the current study given that the bilingual group's English vocabulary scores were significantly lower than the monolingual group's and that researchers have found a strong correlation between children's language proficiency and performance in mental state tasks (P. L. Harris, 2006). However, when language proficiency was not controlled, there was only evidence of a monolingual advantage (based on effect size) in the Explicit False Belief task. Additionally, without controlling for language proficiency there was evidence of a bilingual advantage (based on effect size) in the Diverse Desires task. This finding is important because it suggests that while bilingualism may have negative impacts in some areas (e.g. language proficiency in each individual language) it has positive effects on others (e.g. ability to reason about mental states). In other words,

the current findings suggest that bilingualism is in some way compensating for a lower level of language proficiency.

Based on these findings, can we conclude that bilingualism leads to an advantage in mental state reasoning? As discussed in Chapter 3, the magnitude of the bilingual effect on various cognitive abilities varies based on particular nature of the bilingual sample being studied. Thus, across studies one would expect to find variations in the comparative performances of bilingual and monolingual groups. Because of these variations, the interesting question is not whether bilingualism leads to an *advantage* in mental state reasoning abilities, but whether bilingualism affects the development of mental state reasoning abilities (the current study suggests that it does). Because there is evidence that bilingualism affects mental state reasoning abilities we should strive to identify the specific aspects of bilingualism that impact mental state reasoning abilities, and how variations in those factors influence the magnitude of the effect. While past findings suggest that level of language proficiency (both within and across languages), sociolinguistic factors, and sociocultural factors may all play a role, further research is needed to tease apart the relative impact of these different factors as well as their interactions.

Implications and Suggestions for Further Research

Overall, the current study provides evidence that bilingualism impacts the ability to reason about the mental states of others that extends beyond an ability to reason about false beliefs. This finding is notable for several reasons. First of all, past research has found that bilingual groups outperform monolingual groups particularly in tasks that include conflicting or misleading cues. Thus, it is not surprising that evidence has been

found of a bilingual effect on performance in false belief tasks because a key component of those tasks is the conflicting and misleading cues (P. L. Harris, 2006). However, it was unknown whether this superior performance indicated a greater ability to reason about mental states, or simply reflected a greater ability to negotiate misleading cues. Because effect sizes indicate that the bilingual children outperformed the monolingual children in a wider variety of mental state tasks when language proficiency was controlled, the current study provides evidence that bilingualism does impact the ability to reason about mental states that extends beyond an ability to negotiate misleading cues. Another interesting finding in the current study is that bilingualism has a particular effect on children's ability to reason about the emotional states of others. This finding is notable because it suggested that not only does bilingualism affect differences in children's cognitive development, but it might affect differences in their social and emotional development as well.

While the current study provides evidence that bilingualism has a broader effect on mental state reasoning abilities than was previously found, there are some limitations to the current set of results. First of all, these findings are based on research with Spanish-English bilingual children within the United States. As discussed previously, there may be particular linguistic or cultural differences between this specific bilingual group and other bilingual groups that would impact the magnitude of the effect of bilingualism on mental state reasoning abilities. Secondly, the children in the current study came from homes with highly educated parents suggesting that they came from a relatively high socio-economic group. Thus, it is unknown whether we would find similar results if we studied children from a wider variety of SES backgrounds. Third,

while the current study included bilingual children with a range of levels of language ability across the two languages, there may not have been enough variation in degree of bilingualism to show the effect of this factor on the development of mental state reasoning abilities. Additionally, language proficiency was only measured through tests that assessed vocabulary comprehension. Thus, through including a broader range of language abilities across the two languages and including measures of a wider variety of language abilities, we could gain a better understanding of how degree of bilingualism is related to the ability to reason about the mental states of others.

Based on the results of the current study, there are some obvious next steps. First of all, it would be useful to test bilingual children in Spanish and/or non-verbal versions of these tasks. In this way, we could get a better idea of how the linguistic aspects of the tasks are influencing children's performance, and whether the English versions of the tasks underestimated children's ability to reason about the mental states of others. Another obvious next step would be to study bilingual children from a wider variety of language, cultural, and SES backgrounds, and bilingual children with a wider variety of levels of language proficiency. By systematically varying these factors, we could gain a better understanding of how each factor uniquely contributes to the effect of bilingualism on the development of mental state reasoning abilities as well as how these factors contribute to the magnitude of that effect.

In addition to these specific questions, there are also some broader questions that can be asked. One question is how skills in cognitive flexibility are related to differences in children's understanding of mental states. A second question involves how differences in the bilingual children's sociolinguistic experiences contribute to different cognitive

outcomes. Because this and other studies suggest that there is more than one cognitive difference between bilingual and monolingual groups, it is highly likely that there is more than one cause for these various cognitive outcomes. As mentioned previously, Costa and colleagues (2009) suggest that bilinguals who interact with more bilingual speakers may have more practice in switching back and forth between the two languages thus leading to greater advantages in cognitive flexibility. However, the current findings suggest that bilingual individuals who interact with more monolingual speakers have a greater understanding of the emotions of others. A lot more research needs to be done in order to better understand how variations in these factors contribute to these different outcomes.

Finding answers to these questions and subsequent questions that arise are important for both theoretical as well as practical reasons. Theoretically this research will help us better understand the relationship between language and cognitive processes in the developing child. From a practical standpoint, it is important for parents and educators to know that exposing children to two languages early in life does not lead to overall cognitive and language deficits. Additionally, further research that identifies specific aspects of the bilingual experience that contribute to various outcomes could inform parents and educators about how to help children achieve the best outcomes possible.

Table 1. Means, standard deviations (in parentheses), and ranges for the monolingual and bilingual group's age (in months), PPVT-raw, PPVT-standardized, PPVT-age equivalent (in months), TVIP-raw, TVIP-standardized, and TVIP-age equivalent (in months) scores.

	Age	**PPVT-Raw**	**PPVT-Standardized**	**PPVT-Age Equivalent**	**TVIP-Raw**	**TVIP-Standardized**	**TVIP-Age Equivalent**
Monolingual Group	Mean= 53.42 (13.59) Range= 36-76	Mean=72.96 (28.15) Range= 34-127	Mean=112.19 (11.72) Range= 90-142	Mean=67 (23.98) Range= 33-116	-	-	-
Bilingual Group	Mean= 54.12 (11.90) Range= 36-74	Mean= 55.77 (17.70) Range=27-84	Mean= 98.04 (10.66) Range= 66-117	Mean= 52 (14.92) Range= 26-75	Mean= 29.65 (15.71) Range= 6-71	Mean= 102.00 (17.03) Range= 59-139	Mean= 58 (18.31) Range= 32-108

Table 2. The mental state tasks are ordered from easiest to most difficult for the monolingual and bilingual groups based on group means. The numbers listed to the left of each task are based on the order found by Wellman and Liu (2004) and the group means are listed in parentheses.

Monolingual Group	**Bilingual Group**
1- Diverse Desires (1.54)	1- Diverse Desires (1.88)
2- Diverse Beliefs (1.35)	2- Diverse Beliefs (1.19)
3- Knowledge Access (1.04)	3- Knowledge Access (1.15)
4- Contents False Belief (1.00)	6- Belief-Emotion (1.00)
5- Explicit False Belief (.81)	7- Real-Apparent Emotion (.88)
6- Belief-Emotion (.73)	4- Contents False Belief (.73)
7- Real-Apparent Emotion (.58)	5- Explicit False Belief (.35)

Table 3. Mean differences between the monolingual group's performance in each individual mental state task.

Monolingual Group	**Diverse Desires**	**Diverse Beliefs**	**Knowledge Access**	**Contents False Belief**	**Explicit False Belief**	**Belief-Emotion**	**Real-Apparent Emotion**
Diverse Desires	-						
Diverse Beliefs	.192	-					
Knowledge Access	.500	.308	-				
Contents False Belief	.538	.346	.038	-			
Explicit False Belief	.731*	.538	.231	.192	-		
Belief-Emotion	.808*	.615	.308	.269	.077	-	
Real-Apparent Emotion	.962**	.769*	.462	.423	.231	.154	-

* < .05, ** < .01, *** < .001

Table 4. Mean differences between the bilingual group's performance on each individual mental state task.

Bilingual Group	**Diverse Desires**	**Diverse Beliefs**	**Knowledge Access**	**Contents False Belief**	**Explicit False Belief**	**Belief-Emotion**	**Real-Apparent Emotion**
Diverse Desires	-						
Diverse Beliefs	0.692*	-					
Knowledge Access	0.731*	0.038	-				
Contents False Belief	1.154***	0.462	0.423	-			
Explicit False Belief	1.538***	0.846**	0.808*	0.385	-		
Belief-Emotion	0.885**	0.192	0.154	0.269	0.654	-	
Real-Apparent Emotion	1.000***	0.308	0.269	0.154	0.538	0.115	-

* < .05, ** < .01, *** < .001

Table 5. Correlations between age (in months), PPVT raw score, and scores on the mental state tasks for the monolingual group.

Monolingual Group	**Age**	**PPVT (Raw)**	**Diverse Desires**	**Diverse Beliefs**	**Know-ledge Access**	**Contents False Belief**	**Explicit False Belief**	**Belief-Emotion**	**Real-Apparent Emotion**
Age	-								
PPVT (Raw)	.891***	-							
Diverse Desires	.371	.361	-						
Diverse Beliefs	-.193	-.112	.392*	-					
Knowledge Access	.829***	.739***	.366	-.233	-				
Contents False Belief	.758***	.658***	.506**	-.016	.909***	-			
Explicit False Belief	.157	.240	.525**	.213	.264	.361	-		
Belief-Emotion	.743***	.700***	.356	.025	.622***	.705***	.216	-	
Real-Apparent Emotion	.185	.286	.443*	-.014	.234	.320	.443*	.289	-

* < .05, ** < .01, *** < .001

Table 6. Correlations between age (in months), PPVT raw score, TVIP raw score, and scores on the mental state tasks for the bilingual group.

Bilingual Group	**Age**	**PPVT (Raw)**	**TVIP (Raw)**	**Diverse Desires**	**Diverse Beliefs**	**Know-ledge Access**	**Contents False Belief**	**Explicit False Belief**	**Belief-Emotion**	**Real-Apparent Emotion**
Age	-									
PPVT (Raw)	.684***	-								
TVIP (Raw)	.567**	.477*	-							
Diverse Desires	.322	.410*	.047	-						
Diverse Beliefs	-.013	.051	.104	.267	-					
Knowledge Access	.666***	.539**	.464*	.364	.113	-				
Contents False Belief	.566**	.387	.525**	.221	.357	.795***	-			
Explicit False Belief	.018	.434*	.537**	.129	.316	.342	.317	-		
Belief-Emotion	.185	.344	.327	.284	.228	.278	.311	.383	-	
Real-Apparent Emotion	.235	.432*	.130	.285	-.022	.129	-.041	.375	.095	-

* < .05, ** < .01, *** < .001

Table 7. Differences between the correlations of the monolingual and bilingual groups. Positive values indicate that the monolingual group has a higher correlation than the bilingual group. Negative values indicate that the bilingual group has a higher correlation than the monolingual group.

	Age	**PPVT (Raw)**	**Diverse Desires**	**Diverse Beliefs**	**Knowledge Access**	**Contents False Belief**	**Explicit False Belief**	**Belief-Emotion**	**Real-Apparent Emotion**
Age	-								
PPVT (Raw)	.207*	-							
Diverse Desires	.049	-.049	-						
Diverse Beliefs	.180	.061	.125	-					
Knowledge Access	.163	.200	.002	.120	-				
Contents False Belief	.192	.271*	.285*	-.341**	.114	-			
Explicit False Belief	.139	-.194	.396**	-.103	-.078	.044	-		
Belief-Emotion	.558**	.356**	.072	-.203*	.344**	.394**	-.167	-	
Real-Apparent Emotion	.050	-.146	.158	-.008	.105	.279*	.068	.194	-

*Difference > .20, **Difference > .30

Table 8. Regression analysis (for each task) with age (in months), language group, and PPVT raw score as the independent variables and each mental state task score as the dependent variable.

	R^2	**F-values** F (3, 48)	**Age**		**Language Group**		**PPVT (Raw)**	
Diverse Desires	.218	F = 4.46 p = .008**	β = .128	t = .59 p = .558	β = -.381	t = -2.48 p = .017*	β = .273	t = 1.18 p = .243
Diverse Beliefs	.024	F = .385 p = .764	β = -.193	t = .80 p = .429	β = .046	t = .27 p = .789	β = .124	t = .48 p = .633
Knowledge Access	.583	F = 22.37 p = .000***	β = .648	t = 4.10 p = .000***	β = -.095	t = -.85 p = .402	β = .144	t = .86 p = .396
Contents False Belief	.459	F = 13.56 p = .000***	β = .651	t = 3.62 p = .001***	β = .110	t = .86 p = .393	β = .025	t = .13 p = .898
Explicit False Belief	.212	F = 4.32 p = .009**	β = -.400	t = -1.84 p = .072	β = .178	t = .18 p = .860	β = .656	t = 2.83 p = .007**
Belief-Emotion	.309	F = 7.14 p = .000***	β = .104	t = .509 p = .613	β = -.308	t = -2.13 p = .038*	β = .480	t = 2.21 p = .032*
Real-Apparent Emotion	.153	F = 2.89 p = .045*	β = -.170	t = -.757 p = .453	β = -.367	t = -2.30 p = .026*	β = .495	t = 2.06 p = .045*
Composite Score	.336	F = 8.10 p = .000***	β = .204	t = 1.03 p = .310	β = -.112	t = -.79 p = .432	β = .428	t = 2.01 p = .050*

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 9. Regression analysis for the monolingual group with age (in months), and PPVT raw score as the independent variables and each mental state task score as the dependent variable.

Monolingual Group	**R^2**	**F-values** F (2, 23)	**Age**		**PPVT (Raw)**	
Diverse Desires	.142	F = 1.91 p = .171	β = .240	t = .56 p = .578	β = .148	t = .35 p = .731
Diverse Beliefs	.055	F = .66 p = .525	β = -.451	t = -1.01 p = .323	β = .290	t = .65 p = .522
Knowledge Access	.688	F = 25.33 p = .000***	β = .827	t = 3.22 p = .004*	β = .003	t = .01 p = .991
Contents False Belief	.576	F = 15.61 p = .000***	β = .832	t = 2.78 p = .011*	β = -.083	t = -.28 p = .784
Explicit False Belief	.073	F = .91 p = .416	β = -.276	t = -.625 p = .538	β = .486	t = 1.10 p = .282
Belief-Emotion	.559	F = 14.57 p = .000***	β = .578	t = 1.90 p = .071	β = .185	t = .61 p = .550
Real-Apparent Emotion	.105	F = 1.36 p = .278	β = -.338	t = -.78 p = .444	β = .587	t = 1.35 p = .189
Composite Score	.441	F = 9.06 p = .001**	β = .368	t = 1.07 p = .295	β = .315	t = .92 p = .369

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 10. Regression analysis for the bilingual group with age (in months), PPVT raw score, and TVIP raw score as the independent variables and each mental state task score as the dependent variable.

Bilingual Group	**R^2**	**F-values** F (3, 22)	**Age**		**PPVT (Raw)**		**TVIP (Raw)**	
Diverse Desires	.214	F= 1.99 p= .144	β = .192	t = .69 p = .498	β = .399	t = 1.52 p = .142	β = -.252	t = -1.09 p = .288
Diverse Beliefs	.023	F= .17 p= .916	β = -.160	t = -.51 p = .613	β = .087	t = .30 p = .769	β = .154	t = .59 p = .559
Knowledge Access	.465	F= 6.37 p= .003**	β = .510	t = 2.21 p = .038*	β = .138	t = .64 p = .530	β = .108	t = .57 p = .577
Contents False Belief	.383	F= 4.56 p= .012*	β = .427	t = 1.73 p = .098	β = -.052	t = -.22 p = .825	β = .308	t = 1.50 p = .149
Explicit False Belief	.647	F= 13.46 p= .000***	β = -.834	t = -4.46 p = .000***	β = .677	t = 3.86 p = .001**	β = .687	t = 4.42 p = .000***
Belief-Emotion	.175	F= 1.55 p= .229	β = -.220	t = -.77 p = .450	β = .361	t = 1.34 p = .193	β = .280	t = 1.18 p = .251
Real-Apparent Emotion	.198	F= 1.81 p= .175	β = -.081	t = -.29 p = .777	β = .523	t = 1.98 p = .061	β = -.074	t = -.32 p = .755
Composite Score	.201	F= 1.85 p= .168	β = .156	t = .55 p = .585	β = .362	t = 1.37 p = .185	β = -.069	t = -.30 p = .770

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 11. Mean composite mental state task score for the monolingual and bilingual groups (range 0-14) with the standard deviation (in parenthesis). Results from the ANOVA and ANCOVA with language group (bilingual, monolingual) as the independent variable, the composite mental state task score as the dependent variable, and PPVT score as the covariate (in the MANCOVA).

	ANOVA				**ANCOVA**			
	Monolingual Means	**Bilingual Means**	**F (1, 50)**	**p- and d-values**	**Monolingual Adjusted Means**	**Bilingual Adjusted Means**	**F (1, 49)**	**p- and d-values**
Composite Score	7.04 (4.02)	6.81 (3.35)	F = .05	p = .823 d = .06	6.27	7.57	F = 2.04	p = .160 d = .42

* $p < .05$, ** $p < .01$
* $d < .50$, ** $d < .80$

Table 12. Mean and adjusted mean score for each mental state task for the monolingual and bilingual groups (range 0-2) with the standard deviation (in parenthesis). Results from the MANOVA and MANCOVA with language group (bilingual, monolingual) as the independent variable, the mental state task scores as the dependent variables, and PPVT score as the covariate (in the MANCOVA).

	MANOVA				MANCOVA			
	Monolingual Means	**Bilingual Means**	**F (1, 50)**	**p- and d-values**	**Monolingual Adjusted Means**	**Bilingual Adjusted Means**	**F (1, 49)**	**p- and d-values**
Diverse Desires	1.54 (.71)	1.88 (.43)	F = 4.55	p = .038* d = .60*	1.46	1.97	F = 9.75	p = .003** d = .91**
Diverse Beliefs	1.35 (.75)	1.19 (.90)	F = .45	p = .504 d = .19	1.36	1.18	F = .52	p = .473 d = .33
Knowledge Access	1.04 (1.00)	1.15 (.88)	F = .20	p = .661 d = .12	.81	1.38	F = 7.23	p = .010** d = .78*
Contents False Belief	.92 (1.02)	.73 (.92)	F = .51	p = .478 d = .20	.73	.93	F = .67	p = .416 d = .24
Explicit False Belief	.81 (.94)	.35 (.75)	F = 3.85	p = .055 d = .56*	.71	.44	F = 1.29	p = .262 d = .33
Belief-Emotion	.73 (.92)	1.00 (.98)	F = 1.04	p = .312 d = .28	.54	1.19	F = 7.23	p = .010** d = .78*
Real-Apparent Emotion	.58 (.76)	.88 (.86)	F = 1.87	p = .178 d = .39	.48	.98	F = 4.87	p = .032* d = .64*

* $p < .05$, ** $p < .01$
* $d < .50$, ** $d < .80$

Table 13. Number and percent (in parentheses) of children from the bilingual group whose standardized scores in the PPVT and the TVIP fell above the standardized mean (100) or below the standardized mean. The mean PPVT and TVIP standardized scores and standard deviation (in parentheses) appears for each group.

Bilingual Group	**High English**	**Low English**
High Spanish	n = 7 (27%) English = 108.29 (6.16) Spanish = 117.43 (14.75)	n = 8 (30%) English = 91.63 (6.76) Spanish = 107.50 (6.63)
Low Spanish	n = 5 (19%) English = 103.00 (2.55) Spanish = 82.00 (17.33)	n = 6 (23%) English = 90.50 (12.16) Spanish = 93.33 (3.88)

Table 14. Number and percent (in parenthesis) of children from the monolingual group whose standardized scores in the PPVT fell above the standardized mean (100) or below the standardized mean. Mean PPVT standardized scores and standard deviation (in parentheses) appears for each group.

Monolingual Group	**High English**	**Low English**
English	n = 24 (92%) English =113.96 (10.34)	n = 2 (8%) English = 91.00 (1.41)

References

Arsenian, S. (1937). *Bilingualism and Mental Development.* New York: Teachers College Press.

Astington, J. W., & Baird, J. A. (2005). Introduction: Why language matters. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 3-25). Oxford: Oxford University Press.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 5*, 1311-1320.

August, D., & Hakuta, K. (1997). *Improving schooling for langugae-minority children: A research agenda.* Washington, DC: National Academy Press.

Baker, C. (2006). *Foundations of bilingual education and bilingualism.* Tonawanda, NY: Multilingual Matters Ltd.

Barke, E. M., & Perry-Williams, D. E. (1938). A further study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *British Journal of Educational Psychology*(8), 63-77.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition, 21*, 37-46.

Ben-Zeev, S. (1977). The influence of bilngualism on cognitive strategy and cognitive development. *Child Development, 48*, 1009-1018.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57*, 498-510.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Child Development, 57*, 498-510.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Developoment, 70*(3), 636-644.

Bialystok, E. (2001). *Bilingualism in development: language, literacy, and cognition.* Cambridge, UK: Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology, 60*, 68-79.

Bialystok, E. (2007). Cognitive Effects of Bilingualism: How Linguistic Experience Leads to Cognitive Change. *Int. Jnl of Bilingual Edu. & Bilingualism, 10*(3), 210. doi: 10.2167/beb441.0

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism, 12*(01), 3. doi: 10.1017/S1366728908003477

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859-873.

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(6), 1341-1354. doi: 10.1037/0278-7393.32.6.1341

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*, 290-303.

Bialystok, E., & Fang, F. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language, 109*, 93-100.

Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics, 19*, 69-85.

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24*, 27-44.

Bialystok, E., Martin, M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9*(1), 103-119. doi: 10.1177/13670069050090010701

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in biligual children: evidence from the dimensional change card sort task. *Developmental Science, 7*, 325-339.

Bialystok, E., & Senman, L. (2004). Executive Processes in Appearance-Reality Tasks: The Role of Inhibition of Attention and Symbolic Representation. *Child Development, 75*(2), 562-579.

Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: the effect of bilingualism on reversing ambiguous figures. *Developmental Sci, 8*(6), 595-604.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for biilngual children in two cultures. *Cognition, 112*, 494-500.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, B25-B31.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 21*, 82-86.

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 76-86.

Brown, R. (1976). Reference: in memorial tribute to Eric Lenneberg. *Cognition*, 125-153.

Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70*(2), 381-395.

Carlson, S. M., & Choi, H. P. (2009). *Bilingual and bicultural: Executive function in Korean and American children.* Paper presented at the Society of Research in Child Development, Denver, Colorado.

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more. *Pscyholgoical Science, 16*(8), 609-616.

Carlson, S. M., & Meltzoff, A. (2008). Bilingual experience and executive functioning in young children. *Developmental Sci, 11*(2), 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and

working memory. *Inf. Child Develop., 11*(2), 73-92. doi: 10.1002/(ISSN)1522-7219

Caskey-Sirmons, L. A., & Hickerson, N. P. (1977). Semantic shift and bilingualism: variation in the color terms of five languages. *Anthropological Linguistics, 19*(8), 358-367.

Chomsky, N. (1967). Recent contributions to the theory of innate ideas. *Synthese, 17*(1), 2-11.

Colzato, L., Bajo, M., Van Den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 302-312. doi: 10.1037/0278-7393.34.2.302

Costa, A., Hernández, A., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*, 135-149.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59-86. doi: 10.1016/j.cognition.2006.12.013

Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics, 20*, 1-20.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9*.

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of Cross-Cultural Psychology, 9*, 131-149.

Curtiss, S., Fromkin, V., Krashen, S., Rigler, D., & Rigler, M. (1974). The linguistic development of Genie. *Language, 50*(3), 528-554.

Cutting, A. L., & Dunn, J. (1999). Theory of Mind, Emotion Understanding, Language, and Family Background: Individual Differences and Interrations. *Child Development, 70*(4), 853-865.

Darcy, N. T. (1953). A review of the literature on the effects of bilingualism upon the measurement of intelligence. *Journal of Genetic Psychology, 82*, 21-57.

Darcy, N. T. (1963). Bilingualism and the measurement of intelligence: Review of a decade of research. *Journal of Genetic Psychology, 103*, 259-282.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a ston-age tribe. *Nature, 398*, 203-204.

de Villiers, J. (2007). The interface of language and Theory of Mind. *Lingua, 117*(11), 1858-1878. doi: 10.1016/j.lingua.2006.11.006

de Villiers, J., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development, 17*(1037-1060).

Diaz, R. M. (1983). Thought and two languages: The impact of bilingualism on cognitive development. *Review of Research in Education, 10*(23), 23-54.

Dunn, J., Brown, J. B., & Beardsall, L. (1991). Falmily talk about feeling states and children's later understanding of other's emotions. *Developmental Psychology, 27*(3), 448-455.

Dunn, L. W., Dunn, L. M., & Williams, K. T. (1997). *Peabody Picture Vocabulary Test-Third edition*. Circle Pines, MN: American Guidance Service.

Dunn, L. W., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario en Imágenes Peabody*. Circle Pines, MN: American Guidance Service.

Feldman, C., & Shen, M. (1971). Some language-related cognitive advantages of bilingual five-year-olds. *Journal of Genetic Psychology, 118*, 235-244.

Figueras-Costa, B., & Harris, P. L. (2001). Theory of mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education, 6*(2), 92-102.

Flavell, J. H. (1986). The development of children's knowledge about the appearance-reality distinction. *American Psychologist, 41*, 418-425.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology, 15*(95-120).

Fodor, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Gelman, R., & Gallistel, C. R. (2004). Language and the origin of numerical concepts. *Science, 306*(5695), 441-443.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind* (pp. 195-235). Cambridge, MA: Massachusetts Institute of Technology.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225-277). Cambridge: Cambridge Univesity Press.

Goetz, P. (2003). The effects of bilingualism on theory of mind development. *Bilingualism, 6*(1), 1-15. doi: 10.1017/S1366728903001007

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*, 26-37.

Gordon, P. (2004). Numerical cognition without words: Evidence from amazonia. *Science, 306*, 496-499.

Grabo, R. P. (1931). A study of comparative vocabularies of junior high school pupils from English and Italian speaking homes. *Bulletin No. 13*. Washington, D.C.: U.S. Office of Education.

Grainger, J., & Dijkstra, A. (1992). On the representation and use of langugae information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207-220). Asterdam: Elsevier.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-81.

Gumperz, J. J., & Levinson, S. C. (1997). Rethinking linguistic relativity. *Current Anthropology, 32*, 613-623.

Guttentag, R. E., Haith, M. M., Goodman, G. S., & Hauch, J. (1984). Semantic processing of unattended words in bilinguals: a test of the input switch mechanism. *Journal of Verbal Learning and Verbal Behavior, 23*, 178-188.

Hakuta, K. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: a critical discussion and some new longitudinal data. *Children's Language-book, 5*, 329-344.

Harris, C. W. (1948). An exploration of language skill patterns. *Journal of Educational Psychology, 32*, 351-364.

Harris, P. L. (2006). Social cognition. In W. Damon, R. Lerner, D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology* (6th ed., Vol. 2: Cognition, Perception and Language, pp. 811-858). New York: John Wiley.

Harris, P. L., De Rosnay, M., & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science, 14*(2), 69-73.

Harris, P. L., & Gross, D. (1988). Children's understanding of real and apparent emotion. In J. Astington, P. L. Harris & D. R. Olson (Eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Harris, P. L., Olthof, T., Terwogt, M. M., & Hardman, C. E. (1987). Children's knowledge of the situations that provoke emotion. *International Journal of Behavioral Development, 10*, 319-344.

Holowka, S., Brosseau-Lapre, F., & Petitto, L. A. (2002). Semantic and conceptual knowledge underlying young bilingual babies' signs and words. *Language Learning, 55*(2), 205-262.

Ianco-Warrall, A. (1972). Bilingual and cognitive development. *Child Development, 43*, 1390-1400.

Jensen, J. V. (1962). Effects of childhood bilingualism. *Elementary English, 39*, 132-143.

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging, 23*, 70-78.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist, 86*, 65-79.

King, K., & Fogle, L. (2006). Raising bilingual children: Common parental concerns and current research. *CALdigest*, 2.

Kovács, A. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Sci, 12*(1), 48-54. doi: 10.1111/desc.2009.12.issue-1

Kovács, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, 106*(16), 6556-6560.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. d. G. a. J. F. Kroll (Ed.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 169-199). Mahwah, NJ: Erlbaum.

Kuczaj, S. A., & Hendry, J. L. (2003). Does language help animals think? In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the Study of Language and Thought* (pp. 237-276). Cambridge, MA: MIT Press.

Langer, J. (2000). The descent of cognitive development. *Developmental Science, 3*, 361-378.
Lee, K., Olson, D. R., & Torrance, N. (1999). Chinese children's understanding of false beliefs: the role of language. *Journal of Child Language, 26*, 1-21.
Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Development, 74*(4), 1130-1144.
Macnamara, J. (1966). *Bilingualism and primary education.* Edinburgh: Edinburgh University Press.
MacNamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior, 10*, 480-487.
Martin-Rhee, M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism, 11*(1), 81-93.
Morton, J., & Harper, S. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Sci, 10*(6), 719-726. doi: 10.1111/desc.2007.10.issue-6
Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology, 21*, 258-270.
Oller, D. K., & Eilers, R. (2002). *Language and literacy in bilingual children.* Clevedon, UK: Multilingual Matters.
Özgen, E., & Davies, I. R. L. (2002). Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General, 131*(4), 477-493.
Pavlenko, A. (2006). Bilingual Selves. In A. Pavlenko (Ed.), *Bilingual minds: emotional experience, expression and representation.* Clevedon, UK: Multilingual Matters.
Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*(2), 147-164.
Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs, 76*, 1-26.
Pearson, B. Z., Fernández, S., & Oller, D. K. (1993). Lexical development in bilingual babies and toddlers: Comparison to monolingual norms. *Language Learning, 43*(1), 93-120.
Penfield, W., & Roberts, L. (1959). Speech and brain mechanisms. *Princeton, NJ: Princeton University Press.*
Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping, 19*, 170-182.
Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *Bristish Journal of Developmental Psychology, 5*, 125-137.
Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. In M. Coltheart & M. Davies (Eds.), *Pathologies of belief* (pp. 123-145). Malden, MA: Blackwell Publishers.
Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*(2), 502-517.

Petitto, L. A., & Holowka, S. (2002). Evaluating attributions of delay and confusion in young bilinguals: Special insights from infants acquiring a signed and spoken language. *Sign Language Studies, 3*(1), 4-33.

Petitto, L. A., Katerelos, M., Levi, B., Gauna, K., Tatrault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language, 28*(2).

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 499-503.

Pinker, S. (1994). *The Language Instinct.* New York: W. Morrow & Co.

Pons, F., Lawson, J., Harris, P. L., & de Rosnay, M. (2003). Individual differences in children's emotion understanding: Effects of age and language. *Scandinavian Journal of Psychology, 44*, 347-353.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*, 415-422.

Premack, D. (1983). Animal cognition. *Annual Review of Psychology, 34*, 351-362.

Rattermann, M. J., & Gentner, D. (1998). The effect of language on similarity: The use of relational labels improves chidlren's performance in a mapping task. In K. Holyoak, D. Gentner & B. Kokinov (Eds.), *Advances in analyogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 274-282). Sophfia: New Bulgarian University.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Pscyholinguistic Research, 21*(4), 301-316.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*, 378-411. doi: 10.1016/j.cogpsych.2004.10.001

Roberson, D., Davidoff, J., & Shapiro, L. (2002). Squaring the circle: The cultural relativity of 'good' shape. *Journal of Cognition and Culture, 2*(1), 29-51.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernández, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology, 16*, 635-645.

Rosch, E. H. (1973). Natural Categories. *Cognitive Psychology, 4*, 328-350.

Sachs, O. S., & Coly, J. D. (2006). Envy and jealousy in Russian and English: Labeling and conceptualization of emotions by monolinguals and bilinguals. In A. Pavlenko (Ed.), *Bilingual Minds: Emotional Experience, Expression and Representation*. Clevedon, UK: MPG Books Ltd.

Saer, D. J. (1923). The effect of bilingualism on intelligence. *The British Journal of Psychology, 14*, 25-38.

Sapir, E. (1929). The status of linguistics as a science. *Language, 5*(4), 207-214.

Sera, M. D., Bales, D. W., & del Castillo Pintado, J. (1997). Ser helps Spanish speakers identify "real" properties. *Child Development, 68*(5), 820-831.

Shatz, M., Diesendruck, G., Martinez-Beck, I., & Akar, D. (2003). The influence of language and socioeconomic status on children's understanding of false belief. *Developmental Psychology, 39*(4), 717-729. doi: 10.1037/0012-1649.39.4.717

Tabouret-Keller, A. (2004). Bilingualism in Europe. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 662-688). Malden, MA: Blackwell Publishing.

Tyack, P. L. (1993). Animal language research needs a broader comparative and evolutionary framework. In H. L. Roitblat, L. M. Herman & P. E. Nachtigall (Eds.), *Language and communication: comparative perspectives* (pp. 115-152). Hillsdale, NJ: Erlbaum.

U.S. Census Bureau. (2003). *Language use and English-speaking ability: 2000.* Washington, DC: U.S. Census Bureau.

Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Ollie, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63*, 1012-1020.

Vinden, P. (1999). Children's understanding of mind and emotion: a multi-culture study. *Cognition and Emotion, 13*(1), 19-48.

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understanding in chinese children. *Psychological Science, 17*(12), 1075-1081.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 20.

Whorf, B. J. (1956). *Language, Thought and Reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.

Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development, 73*, 768-778.

Yang, S., & Lust, B. (2004). *Testing effects of bilingualism on executive attention: comparison of cognitive performance on two non-verbal tests*. Cornell University.

Zelazo, P., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive funciton in children. *Current Directions in Psychological Science, 7*, 121-126.

Appendix A:

Language Background Questionnaire

1. How many years has the child's immediate family been in the United States?_____
2. What is the birthplace of the parents (guardians)?

   Mother: _____

   Father: _____
3. What is the child's birthplace? _____
4. In what language did each parent/guardian receive most of his/her education?

   Mother: _____

   Father: _____
5. How many years of schooling did each parent (guardian) complete (or highest degree attained)?

   Mother: _____

   Father: _____
6. What language do the parents (guardians) speak at home most of the time?

   Mother: _____

   Father: _____
7. What language does your child speak with his/her parents?

   Mother:_____

   Father: _____
8. How many siblings does your child have? _____
9. What language does your child most often speak with his/her brothers and sisters? (List each sibling and language)

Sibling/Relationship	**Language**
_____	_____
_____	_____
_____	_____
_____	_____

10. What language does your child most often speak with his/her friends or playmates?

    _____

11. In which language are television and radio programs most often received in the home?_____

12. In which language is most print media (books, magazines, newspapers) in the home?

    _____

13. Does the family receive a daily newspaper? If so, in what language? _____

14. How much time per week is your child read to?

    None 15 minutes 30 minutes 45 minutes one hour or more

    In what language(s)?_____

15. How many hours per week does your child observe his/her parents reading? _____

    In what language(s)? _____

16. Has your child learned to read yet? No Limited Yes

    In what language(s)? _____

17. Prior to kindergarten, had your child ever attended a preschool or playgroup conducted in a language other than English? If so, please specify. _____

    _____

18. How many months of schooling has your child received in a language other than English?

    _____

19. Please describe any specific experiences your child has had with relatives, close friends (including exchange students) or caregivers who speak languages other than English. _____

    _____

    _____

    _____

Appendix B:

Language Background Questionnaire

1. ¿Por cuántos años ha estado en los Estados Unidos la familia inmediata del niño/a?___
2. ¿Cuál es el lugar de nacimiento de los padres (guardianes)?
   Madre: _____
   Padre: _____
3. ¿Cuál es el lugar de nacimiento del niño/a? _____
4. ¿En qué idioma recibieron los padres/guardianes la mayoría de su educación?
   Madre: _____
   Padre: _____
5. ¿Cuántos años de educación cumplieron los padres/guardianes?
   Madre: _____
   Padre: _____
6. ¿En qué idioma hablan los padres usualmente en casa?
   Madre: _____
   Padre: _____
7. ¿En qué idioma habla su niño/a con los padres?
   Madre: _____
   Padre: _____
8. ¿Cuántos hermanos tiene su niño/a? _____
9. ¿En qué idioma habla su niño/a con sus hermanos?
   (Liste cada hermano/a y el idioma)

Hermano/a / Relación	**Idioma**
_____	_____
_____	_____
_____	_____
_____	_____

10. ¿En qué idioma habla su niño/a con más frequencia con sus amigos?

    _____

11. ¿En qué idioma son recibidos los programas de radio y la televisión con más frequencia en su casa?_____

12. ¿En qué idioma es la mayoría de los materiales escritos (libros, revistas, periódicos) en su casa?_____

13. ¿Recibe la familia un periódico diariamente? En qué idioma?_____

14. ¿Cuánto tiempo cada semana alguien lee a su niño/a?

    Ningún 15 minutos 30 minutos 45 minutos 1 hora o más

    ¿En cuál(es) idioma(s)? _____

15. ¿Por cuántas horas cada semana les observa su niño/a a sus padres mientras están leyendo?

    _____

    ¿En cuál(es) idioma(s)? _____

16. ¿Ya ha aprendido su niño/a a leer? No Limitado Sí

    ¿En cuál(es) idioma(s)? _____

17. Antes de la guardería, ¿ha asistido su niño/a una escuela o grupo de jugar que era conducido en un idioma que no era inglés? ¿Cuál idioma?_____

    _____

18. Cuántos meses de educación en otro idioma que no sea inglés ha recibido su niño/a?

    _____

19. Por favor, describa las experiencias que su niño/a ha tenido con los parientes o amigos (incluya los estudiantes del intercambio) quien hablan otro idioma que no sea inglés. _____

    _____

    _____

    _____