J. Child Lang. **43** (2016), 407–424. © Cambridge University Press 2015
doi:10.1017/S0305000915000276

High proficiency across two languages is related to better mental state reasoning for bilingual children

KATHERINE R. GORDON*
University of Iowa, USA

(*Received* 6 *April* 2013 – *Revised* 9 *December* 2013 – *Accepted* 6 *May* 2015 – *First published online 16 June 2015*)

ABSTRACT

Past research suggests that bilingualism positively affects children's performance in false belief tasks. However, researchers have yet to fully explore factors that are related to better performance in these tasks within bilingual groups. The current study includes an assessment of proficiency in both languages (which was lacking in past work) and investigates the relationship between proficiency and performance in a variety of mental state tasks (not just false belief). Furthermore, it explores whether the relationship between language proficiency and performance in mental state tasks differs between bilingual and monolingual groups. Twenty-six Spanish–English bilingual and twenty-six English monolingual preschool-age children completed seven mental state tasks. Findings provide evidence that high proficiency in English is related to better performance in mental state tasks for monolinguals. In contrast, high proficiency in both English and Spanish is related to better performance in mental state tasks for bilinguals.

INTRODUCTION

In the past, bilingualism has been thought to significantly hamper a child's language and cognitive development (see Arsenian, 1937; Darcy, 1953; Macnamara, 1966, for reviews). However, recent research has provided evidence that bilingualism offers some cognitive advantages such as greater cognitive flexibility and better selective attention abilities (see Bialystok,

[*] I would like to thank Dr Maria Sera, Dr Stephanie Carlson, Dr Melissa Koenig, and Dr Kathryn Kohnert for their guidance throughout this project, and to thank Taryn Halverson for help with data collection. I would also like to thank Dr Aaron Taylor and Dr Rahul Parsa for help with the statistical analyses and Dr Karla McGregor and Dr Julie Gros-Louis for their feedback on revisions. Funding from the Institute of Child Development made this project possible and is much appreciated. Address for correspondence: University of Iowa – Communication Sciences and Disorders, 250 Hawkins Drive, Iowa City, IA 52241, United States. e-mail: mail@katiegordon.us

2001, 2009, for reviews). These recent findings have lead to the perception, both in academia as well as general culture, that bilingualism fosters cognitive advantages over monolingualism (de Bruin, Treccani, & Della Sala, 2014). However, other recent studies have shown no differences between the performance of bilinguals and monolinguals on various cognitive tasks (Gathercole *et al*., 2014; Paap & Greenberg, 2013). Through a meta-analysis, de Bruin and colleagues (2014) found that studies that show a clear bilingual advantage on executive function tasks are more likely to be published than studies that show mixed results, null results, or a disadvantage for bilinguals. This publication bias has likely led to a misrepresentation of the effect that bilingualism has on cognitive development. In light of these mixed findings, instead of focusing on the question of whether bilinguals outperform monolinguals on various cognitive tasks, a better way to frame the question would be to explore why bilinguals outperform monolinguals in some cases and not others. Furthermore, researchers should explore whether various factors, such as language abilities and executive function abilities, interact differently throughout development for bilingual and monolingual populations.

One area of research that would benefit from this type of exploration is work that has compared bilingual and monolingual children's performance on mental state (i.e., Theory of Mind) tasks. Findings by Berguno and Bowler (2004), Bialystok and Senman (2004), Goetz (2003), Kovács (2009), and Nguyen and Astington (2014) all suggest that bilingualism positively affects children's ability to reason about false beliefs. Furthermore, Rubio-Fernandez and Glucksberg (2011) found that bilingual adults respond less egocentrically than monolingual adults on a false belief task. A closer examination of this research reveals that while bilinguals showed an outright advantage over monolinguals in four studies (Berguno & Bowler, 2004; Goetz, 2003; Kovács, 2009; Rubio-Fernandez & Glucksberg, 2011), they only showed an advantage after proficiency in the testing language (English) was controlled in two studies (Bialystok & Senman, 2004; Nguyen & Astington, 2014). Many researchers have interpreted results that show a bilingual advantage after proficiency in the testing language is controlled as evidence that bilingualism does in fact foster cognitive advantages (see Carlson & Meltzoff, 2008). However, it is not fully understood how children's language skills are related to performance on various cognitive tasks, and whether controlling for the testing language in this way reveals an effect of bilingualism on various cognitive processes. Furthermore, based on de Bruin and colleagues' (2014) meta-analysis, it seems likely that some researchers found no evidence of a bilingual advantage, or even found a bilingual disadvantage on mental state tasks, but the work was never published. Thus, a more thorough exploration of factors that affect bilingual children's comparative performance to monolingual children on mental state tasks is needed.

An important first step to understanding why bilingual children sometimes show better and sometimes show comparable performance to monolingual children on mental state tasks would be to look at bilingual children's level of proficiency across both languages. Researchers have found substantial evidence of a relationship between the level of proficiency across languages and the magnitude of various cognitive effects among bilingual groups (Bialystok, 1988, 2001; Bialystok & Fang, 2009; Carlson & Meltzoff, 2008; Cummins, 1976; Ricciardelli, 1992). Specifically, bilingual children with high proficiency in both languages tend to outperform their monolingual peers on specific cognitive tasks, while bilingual children with low levels of proficiency in one or both languages tend to perform the same as or worse than their monolingual peers. For example, Ricciardelli (1992) found that only bilingual children who had high proficiency in both French and English outperformed English monolingual children on a variety of cognitive tasks. Cummins (1978) has called this the 'Threshold Hypothesis', in that children must reach a threshold of proficiency in both languages before they start showing cognitive advantages over their monolingual peers.

Exploring the relationship between language proficiency and performance on mental state tasks is important not only because proficiency across languages has been linked to cognitive advantages among bilinguals, but also because many researchers have found evidence of a relationship between language development and the development of mental state reasoning among monolinguals (see Astington & Baird, 2005). While there is quite a bit of disagreement about the exact nature of this relationship, the evidence that language skills are related to the development of mental state reasoning is fairly strong. For example, through a meta-analysis, Milligan, Astington, and Dack (2007) found that a variety of language abilities (i.e., general language abilities, semantics, receptive vocabulary, syntax, memory for complements) at a younger age predicted later false belief task performance, and Slade and Ruffman (2005) found a bi-directional relationship between language proficiency and understanding false beliefs across development. Additionally, deaf children who are not exposed to sign language from birth are delayed in their performance on both standard and low verbal false belief tasks (P. de Villiers & de Villiers, 2012; Figueras-Costa & Harris, 2001; Peterson & Siegal, 2000). Furthermore, there is evidence that learning mental state language (e.g., 'think', 'know') enhances children's ability to reason about mental states (Harris, De Rosnay, & Pons, 2005; Montgomery, 2005), and learning certain syntactic constructions (e.g., complement syntax such as, "He

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

thought he saw a unicorn.") is important for developing an understanding of false beliefs (J. de Villiers & de Villiers, 2009; J. de Villiers & Pyers, 2002).

While this past research makes a fairly strong case for a relationship between monolingual children's language proficiency and their ability to reason about mental states, it is unclear how language abilities affect the development of mental state reasoning among bilingual children. This is a limitation to theories about the development of mental state reasoning, given that bilingual individuals make up a substantial percentage of the world's population. For example, the US census reports that 21% of the total US population speaks a language other than English in their home and 78% of those individuals reported that they speak English either "well" or "very well" (US Census Bureau, 2011). Rates of bilingualism are estimated to be even higher in other regions of the world, such as an estimate of close to 50% in Europe (Tabouret-Keller, 2004). Thus, to explain development more broadly, these theories should address how variations in language proficiency both within and across languages affect the development of children's ability to reason about mental states.

With the exception of Nguyen and Astington (2014), researchers who have assessed bilingual children's performance on mental state tasks did not include reliable measures of proficiency in both languages. Specifically, Bialystok and Senman (2004) assessed bilingual children's English language proficiency by using the Peabody Picture Vocabulary Test (PPVT: Dunn, Dunn, & Williams, 1997), but they did not measure children's proficiency in their non-English language. Goetz (2003) assessed children's proficiency in both English (with the PPVT) and Chinese, but the measure of Chinese proficiency was created by the researcher, and thus not standardized. Kovács (2009) and Berguno and Bowler (2004) verified bilingual status by asking parents and/or teachers about the children's exposure to and use of each language, and Rubio-Fernandez and Glucksberg (2011) asked adult participants about their use of each language. Lastly, although Nguyen and Astington (2014) did assess children's level of proficiency across languages using standardized measures, they did not look at how children's relative level of proficiency across languages affected their performance on the false belief tasks.

Given findings that higher proficiency across languages leads to better performance on specific cognitive tasks for bilinguals, and findings that language development is related to mental state reasoning in monolinguals, the purposes of the current study are: (i) to investigate how proficiency across languages is related to bilingual children's performance on mental state tasks; and (ii) to investigate whether the relationship between language proficiency and mental state task performance differs between bilingual and monolingual groups. In the current study, language proficiency was measured through standardized tests of receptive

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

vocabulary, the PPVT in English (Dunn *et al.*, 1997) and the Test de Vocabulario en Imágenes Peabody in Spanish (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986). This decision was made for two reasons. First, it is easier to compare results to past work because in past studies researchers measured language proficiency by assessing children's receptive vocabulary. Second, in their meta-analysis, Milligan and colleagues (2007) found that receptive vocabulary (as measured through the PPVT or the British Picture Vocabulary Scale) was significantly related to performance on several mental state tasks (including change-of-location, unexpected-identity, deception, and belief-emotion) across fifty-three studies (effect size = ·34). Thus, it seems likely that receptive vocabulary will also be related to performance on the mental state tasks used in the current study.

An additional aim of the current study is to compare bilingual and monolingual children's performance on a variety of mental state tasks, not just false belief tasks. Past research comparing bilingual and monolingual groups has focused almost exclusively on false belief tasks. While Goetz (2003) did include two perspective-taking tasks in her study, the results are difficult to interpret given that she found that English–Chinese bilinguals only outperformed Chinese, but not English, monolinguals in one of the two testing sessions. Although there is a long history of using false belief tasks (especially the unexpected location or Sally/Ann task, unexpected contents task, and appearance/reality task) as a litmus test for mental state reasoning abilities, this use has been criticized in recent years. Bloom and German (2000) cite two main criticisms of this approach, namely: (i) to pass false belief tasks, children need more than an ability to think about mental states, as variations in memory abilities, inhibition, attention, and language abilities are all related to performance on these tasks; (ii) there are many other aspects of mental state reasoning (such as understanding the emotions of others) that are not captured by false belief tasks. Thus, to more fully understand children's ability to reason about the thoughts, beliefs, and desires of others, a number of tasks that vary in both task demands as well as the aspect of mental state reasoning that they assess should be used. To this end, Wellman and Liu (2004) conducted a meta-analysis through which they investigated children's likelihood of passing a variety of mental state tasks at different ages. They identified seven tasks that reliably assess children's understanding of different aspects of mental state reasoning which were scaled in the following order: diverse desires, diverse beliefs, knowledge access, contents false beliefs, explicit false beliefs, beliefs and emotion, and the difference between real and apparent emotions. Children typically pass the tasks listed first (e.g., diverse desires and diverse beliefs) around the age of three, and pass all of the tasks around the age of five. In addition to the meta-analysis, Wellman and Liu conducted an experiment assessing the age at which children pass each task, which confirmed the scale that they had developed. Thus, to assess whether bilingualism only affects children's performance on false belief tasks, or if bilingualism affects children's mental state reasoning abilities more broadly, the current study will compare bilingual and monolingual children's performance on the seven tasks in Wellman and Liu's scale.

To achieve the aims of the current study, there are several methodological challenges with comparing bilingual and monolingual groups that must be considered. First, bilingual children vary in systematic ways from monolingual children because of the conditions that lead to bilingualism in the first place (e.g., immigration, international adoption, etc.). Thus, it is often difficult to match bilingual and monolingual children on age, SES, and a variety of other factors simultaneously (Bialystok, 2001). This is a legitimate concern, as some researchers have found no evidence of a bilingual advantage after they more carefully controlled for differences in ethnicity and SES (Morton & Harper, 2007). To address this challenge, the current study included Spanish–English bilingual children who were matched to English monolingual children on biological sex, age, and parental education (see 'Participants' section below for details about matching). In this way, variables other than the primary variable of interest (i.e., exposure to more than one language) were controlled as much as possible. Spanish–English bilingual children, in particular, were selected for several reasons. First, standardized versions of language measures exist in English and Spanish. Second, this population is fairly representative of bilingual children in the US, as it has been estimated that 60% of the individuals in the US who speak a language other than English in their homes speak Spanish (August & Hakuta, 1997; US Census Bureau, 2011).

Another major methodological challenge is selecting the language that will be used when testing children. Testing all of the children in the same language can be problematic because bilingual children are typically less proficient than monolingual children in each of the languages that they speak (Bialystok & Fang, 2009; Oller & Eilers, 2002; Perani *et al.*, 2003). However, translating the tasks to test children in their most proficient language is problematic as word choice in each language can affect children's performance on mental state tasks (Lee, Olson, & Torrance, 1999; Sera, Bales, & del Castillo Pintado, 1997). Some researchers have addressed this challenge by developing low-verbal or non-verbal versions of mental state tasks (Schick, de Villiers, de Villiers, & Hoffmeister, 2007), but, as yet, such versions do not exist for all of the tasks in Wellman and Liu's (2004) scale. Another solution includes testing each bilingual child twice (once in each language) and testing each monolingual child twice in the same language. This was the strategy employed by Goetz (2003), but is likely not the best method as the monolingual children's performance improved from the first to the second testing session (likely due to practice effects), while the bilingual children's performance did not. To address this challenge, children in the current study were administered Wellman and Liu's (2004) tasks in English. This eliminated confounds that are introduced through translating the tasks or administering the tasks in both languages. However, this decision did raise the possibility that bilingual children would perform worse because they are likely to have lower English language proficiency than their monolingual peers. Thus, the relationship between language proficiency and performance on the mental state tasks will be explored to investigate whether lower English proficiency was related to poorer performance on the mental state tasks for the bilingual group.

With regards to predictions for the current study, given that greater proficiency across languages has been shown to be related to more advantages for bilingual individuals, it is likely that bilingual children with higher proficiency across two languages will demonstrate better performance on the mental state tasks than their bilingual age-mates who have lower proficiency in one or both languages. Furthermore, bilingual children with high proficiency in both languages are more likely than bilingual children with low proficiency in one or both languages to outperform monolingual children on false belief tasks. Given past work that shows a bilingual advantage on false belief tasks, it seems likely that the bilinguals will outperform the monolinguals on these tasks. As for the other mental state tasks, it is difficult to predict whether bilingual children will show an advantage over monolingual children, because bilingual and monolingual children's performance on a wider variety of mental state tasks has not been compared previously.

METHODS

Participants

Twenty-six Spanish–English bilingual children (mean age = 4;6, range = 3;0–6;2, females = 14) and twenty-six English monolingual children (mean age = 4;5, range = 3;0–6;4; females = 14) participated in the study (see Table 1). Children were recruited through local schools, the University of Minnesota's Student Housing, the Institute of Child Development Participant Pool, personal contact by the researcher, and referrals by other research participants. To find bilingual children who had an adequate level of proficiency in both languages to show evidence of bilingual effects, teachers and parents of participants were asked to identify children who were proficient in both English and Spanish. Each bilingual child was matched with a monolingual child (from a larger pool of 39 children) who was the same gender and as close in age as possible. Each bilingual child

TABLE 1. *Information about participants*

	Bilingual group	Monolingual group	Statistical comparison
Sex	females = 14 males = 12	females = 14 males = 12	–
Age	mean = 4;6 range = 3;0–6;2	mean = 4;5 range = 3;0–6;4	$t(50) = 0.85$, $p = .40$, Cohen's $d = 0.05$
Maternal education (number of years)	mean = 15·15 (4·15), range = 3–20	mean = 15·92 (2·15), range = 12–20	$t(50) = 0.41$, $p = .68$, Cohen's $d = 0.23$
Paternal education (number of years)	mean = 15·73 (3·30), range = 9–20	mean = 16·38 (1·88), range = 12–20	$t(46) = 0.41$, $p = .68$, Cohen's $d = 0.24$
English vocab (PPVT-standardized score)	mean = 98 (10·66), range = 66–117	mean = 112 (11·72), range = 90–130	$t(50) = 4.56$, $p < .001$, Cohen's $d = 1.29$
Spanish vocab (TVIP-standardized score)	mean = 102 (17·03), range = 59–139	–	–

was typically only a few months older or younger than the matched monolingual child (mean difference in months = 2·98, SD = 1·67, range = 0–5 months), and the overall ages of the two groups were not significantly different ($t(50) = 0.85$, $p = .40$, Cohen's $d = 0.05$) (see Table 1). Parents were asked on a questionnaire to indicate the number of years of schooling that each parent or guardian completed or highest degree obtained. When highest degree obtained was indicated, the researcher estimated the number of years of schooling (High School = 12, Undergraduate = 16, etc.). Based on responses, number of years of schooling for the mother and father were compared across the bilingual and monolingual groups (see Table 1). Number of years of maternal education did not differ between monolingual and bilingual groups ($t(50) = 0.41$, $p = .68$, Cohen's $d = 0.23$). Data were not given for four of the fathers in the bilingual group, but for the remaining sample, number of years of paternal education did not differ between monolingual and bilingual groups ($t(46) = 0.41$, $p = .68$, Cohen's $d = 0.24$).

Children's receptive vocabulary was assessed through the Peabody Picture Vocabulary Test-Third Edition (Dunn *et al.*, 1997) and through the Test de Vocabulario en Imágenes Peabody, an equivalent Spanish version (Dunn *et al.*, 1986). Children were excluded from the study if they scored 22 or below on the PPVT or 3 or below on the TVIP, the lowest scores that can be obtained on each test respectively. Data from these children were excluded because the researcher was unable to obtain an accurate measure of the children's proficiency in both languages. Based on questionnaire responses, twelve of the bilingual children had parents who both spoke

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

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.

Five additional bilingual children were excluded from the study for the following reasons: their raw score on the PPVT was at 22 or below ($n = 1$), their raw score on the TVIP was at 3 or below ($n = 2$), they did not complete the second session ($n = 1$), or experimenter error ($n = 1$). Four additional monolingual children were excluded from the study for the following reasons: experimenter error ($n = 2$) and regularly being exposed to a language other than English ($n = 2$).

Procedure

The bilingual children completed the PPVT and the TVIP during one session and the seven mental state tasks during a second session a week later. During the first session, bilingual children were allowed to select whether they wanted to complete the English book (PPVT) or Spanish book (TVIP) first. In this way, each child was likely to select the language he/she was most comfortable with, reducing the possibility that performance on the vocabulary tests was affected by not understanding the task. Each monolingual child completed the PPVT and the seven mental state tasks during one session. The parents of both the bilingual and monolingual children were given a brief questionnaire about their educational level and the child's exposure to English as well as other languages. 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 seven mental state tasks were administered to all children in English following the procedures outlined in Wellman and Liu (2004), except that in some cases the character names or toy animals used were changed. The diverse desires task assessed children's understanding that a character could have a different desire (i.e., different food preference) than their own. The diverse beliefs task assessed children's understanding that they could have one belief (i.e., about the location of a cat) while a character could have a different belief. The knowledge access task assessed children's understanding that they might know what was inside an unmarked box, but a character that had never seen inside the box wouldn't know its contents. The contents false belief task assessed children's understanding that they might know the contents of a marked box (i.e., Band-Aid box containing a giraffe), but a character that had never seen inside the box would have a false belief about the contents of that box. The explicit false belief task assessed children's understanding that even though an object was hidden in one location (i.e., gloves hidden in a backpack), a character would look in a different location if that was where he believed that the object was. The belief–emotion task assessed children's understanding that a character could feel happy if he believed that a box contained something desirable (i.e., cheerios), even though the child knew that the box contained something undesirable (i.e., rocks). The real–apparent emotions task assessed children's understanding that an individual could have a happy expression on his face even though he felt sad on the inside. The knowledge access, belief–emotion, and real–apparent emotions tasks included memory check questions that assessed children's understanding of the task thus far (see Wellman & Liu, 2004). If the child answered one of the questions incorrectly, the researcher provided the correct answer to remind the child of a critical component of the task.

To reduce the likelihood that children would pass the mental state tasks by chance, a second version of each task was created. For six of the tasks, the second version followed the exact format of Wellman and Liu (2004) except different pictures or containers were used. For the real–apparent emotions task, the second version of the task included the story of a boy who received socks for his birthday from a friend, but didn't want to hurt his friend's feelings so he hid how he felt. The tasks were administered to each child in one of seven different randomized orders using a 7×7 Latin square design with the two versions of each task always being presented one after the other. Based on their performance on the tasks, each child was assigned a score between 0 and 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 and 14 for their overall performance.

RESULTS

Vocabulary scores

The average standardized PPVT score of the bilingual group was 98 ($SD = 10{\cdot}66$, range 66–117) and the average standardized TVIP score was 102 ($SD = 17{\cdot}03$, range 59–139). The average standardized PPVT score for the monolingual group was 112 ($SD = 11{\cdot}72$, range 90–130). A paired samples t-test revealed that the bilingual children's performance on the PPVT and TVIP (using the standardized scores of these measures) did not significantly differ ($t(25) = 1{\cdot}09$, $p = {\cdot}29$, Cohen's $d = 0{\cdot}28$). Raw PPVT and TVIP scores cannot be compared because they are on different scales. An independent samples t-test was conducted to compare the bilingual and monolingual children's performance on the PPVT. The monolingual children's English vocabulary scores were significantly higher than the bilingual children's for both raw ($t(50) = 2{\cdot}64$, $p = {\cdot}01$, Cohen's $d = 0{\cdot}75$) and standardized scores ($t(50) = 4{\cdot}56$, $p < {\cdot}001$, Cohen's $d = 1{\cdot}29$).

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

Comparing the two versions of the mental state tasks

A series of chi-square analyses were conducted to assess whether children's performance (pass vs. fail) on Wellman and Liu's (2004) version of each mental state task was significantly related to children's performance (pass vs. fail) on the second version of each task. These analyses revealed a significant relationship between children's performance on version 1 and version 2 of each task, indicating that both versions are similar (diverse desires: χ^2 (1, $N = 52$) = 10·84, $p = {\cdot}001$, $\phi = {\cdot}456$; diverse beliefs: χ^2 (1, $N = 52$) = 9·25, $p = {\cdot}002$, $\phi = {\cdot}422$; knowledge access χ^2 (1, $N = 52$) = 28·84, $p < {\cdot}001$, $\phi = {\cdot}745$; contents false-belief χ^2 (1, $N = 52$) = 40·98, $p < {\cdot}001$, $\phi = {\cdot}888$; explicit false-belief χ^2 (1, $N = 52$) = 34·67, $p < {\cdot}001$, $\phi = {\cdot}817$; belief–emotion χ^2 (1, $N = 52$) = 33·68, $p < {\cdot}001$, $\phi = {\cdot}805$; real–apparent emotion χ^2 (1, $N = 52$) = 9·57, $p = {\cdot}002$, $\phi = {\cdot}429$).

Comparing the bilingual and the monolingual children's performance

To compare the two groups' performance on the mental state tasks, an independent samples t-test was conducted with language group (bilingual, monolingual) as the independent variable and the composite mental state task score as the dependent variable. This analysis revealed that there was not a significant difference between the two groups ($t(50) = 0{\cdot}225$, $p = {\cdot}823$, Cohen's $d = 0{\cdot}123$). To compare the groups' performance on the individual tasks, a series of ordinal logistic regressions were conducted with language group (bilingual, monolingual) as the independent variable and task score as the dependent variable. Ordinal regressions were used in this case because the children's scores were either 0, 1, or 2 in the individual tasks, so OLS regression would not be appropriate. These regressions revealed significant effects for the diverse desires and explicit false belief tasks. For the diverse desires task, the bilingual group scored significantly higher than the monolingual group ($B = 1{\cdot}81$, odds ratio [OR] = 6·10 [95% CI = 1·18–31·53], Wald $\chi^2[1] = 4{\cdot}67$, $p = {\cdot}03$). For the explicit false belief task, the bilingual group scored significantly lower than the monolingual group ($B = -1{\cdot}23$, OR = 0·29 [95% CI = 0·09–1·00], Wald $\chi^2[1] = 3{\cdot}87$, $p = {\cdot}05$). None of the other tasks showed a significant difference between groups.

Relationship between language proficiency and mental state tasks performance

To explore whether English language proficiency predicts performance on the mental state tasks for the monolingual and bilingual groups, several regression analyses were conducted with English proficiency (using PPVT standardized scores) as the independent variable and children's composite mental state task score as the dependent variables. For the monolingual group, English proficiency significantly predicted overall performance on the mental state tasks ($B = 0{\cdot}164$, $t = 2{\cdot}656$, $p = {\cdot}014$). However, this was

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

not the case for the bilingual group ($B = 0{\cdot}060$, $t = 0{\cdot}946$, $p = {\cdot}353$). To further explore the relationship between English language proficiency and mental state task performance within the monolingual group, a regression was conducted for each mental state task. These analyses revealed a significant relationship between English proficiency and performance on the knowledge access task ($B = 0{\cdot}038$, $t = 2{\cdot}401$, $p = {\cdot}024$) and the belief–emotion task ($B = 0{\cdot}037$, $t = 2{\cdot}578$, $p = {\cdot}017$). There was not a significant relationship between English proficiency and performance on any of the other mental state tasks for the monolingual group.

To further explore the relationship between language proficiency and performance on the mental state tasks for the bilingual group, a regression analyses with English and Spanish proficiency (using PPVT and TVIP standardized scores) as the independent variables and children's composite mental state task score as the dependent variable was conducted (see Table 2). This analysis revealed an interaction effect of English proficiency and Spanish proficiency ($B = 0{\cdot}013$, $t = 2{\cdot}403$, $p = {\cdot}025$). This interaction was followed up with tests of simple slopes to find the effect of English proficiency at different levels of Spanish proficiency (see Aiken & West, 1991). For Spanish proficiency, at its mean or one standard deviation below its mean, English proficiency was not significantly predictive of composite mental state score (see Table 2). For Spanish proficiency one standard deviation above its mean, however, English proficiency was significantly predictive (simple slope = $0{\cdot}265$, $t = 2{\cdot}627$, $p = {\cdot}015$). Similar regressions were conducted to explore the effect of English and Spanish proficiency on performance on each mental state task. These analyses revealed a significant interaction effect for the diverse desires, ($B = 0{\cdot}005$, $t = 3{\cdot}914$, $p = {\cdot}001$) and explicit false beliefs tasks ($B = 0{\cdot}002$, $t = 2{\cdot}402$, $p = {\cdot}025$). Simple slopes analyses found that for Spanish proficiency one standard deviation below the mean, English proficiency had a negative association with diverse desires (simple slope = $-0{\cdot}035$, $t = -2{\cdot}676$, $p = {\cdot}014$), but for Spanish proficiency one standard deviation above the mean, English proficiency had a positive association with diverse desires (simple slope = $0{\cdot}045$, $t = 4{\cdot}117$, $p < {\cdot}001$). For Spanish proficiency at the mean, English proficiency was not significantly associated with diverse desires. For the explicit false belief task, simple slopes analyses found that English proficiency had a significant positive association with explicit false belief for Spanish proficiency one standard deviation above the mean (simple slope = $0{\cdot}064$, $t = 4{\cdot}222$, $p < {\cdot}001$) or Spanish proficiency at the mean (simple slope = $0{\cdot}030$, $t = 3{\cdot}299$, $p = {\cdot}003$). For Spanish proficiency one standard deviation below the mean, English proficiency was not significantly associated with explicit false belief (see Table 2).

TABLE 2. *Simple slopes analyses following significant Spanish proficiency × English proficiency interactions in predicting task performance for the bilingual group*

	Association between English proficiency and task score at different levels of Spanish proficiency		
	Spanish proficiency 1 SD below mean	Spanish proficiency at mean	Spanish proficiency 1 SD above mean
Diverse desires	$B = -0{\cdot}035$, $t = -2{\cdot}676$, $p = {\cdot}014$	$B = 0{\cdot}005$, $t = 0{\cdot}725$, $p = {\cdot}476$	$B = 0{\cdot}045$, $t = 4{\cdot}117$, $p < {\cdot}001$
Explicit false belief	$B = -0{\cdot}005$, $t = -0{\cdot}249$, $p = {\cdot}805$	$B = 0{\cdot}030$, $t = 3{\cdot}299$, $p = {\cdot}003$	$B = 0{\cdot}064$, $t = 4{\cdot}222$, $p < {\cdot}001$
Composite mental state score	$B = -0{\cdot}192$, $t = -1{\cdot}561$, $p = {\cdot}133$	$B = 0{\cdot}037$, $t = 0{\cdot}613$, $p = {\cdot}546$	$B = 0{\cdot}265$, $t = 2{\cdot}627$, $p = {\cdot}015$

DISCUSSION

The aims of the current study were:(i) to examine the relationship between proficiency across languages and performance on mental state tasks for bilinguals; and (ii) to explore whether the relationship between language proficiency and performance on mental state tasks differed for monolingual and bilingual groups. For the monolingual children, higher English proficiency predicted a better overall mental state task score. This finding coincides with a large body of past research showing a positive relationship between language proficiency and mental state reasoning among monolingual children (J. de Villiers & Pyers, 2002; Lohmann & Tomasello, 2003; Milligan *et al.*, 2007). In contrast, English proficiency was not related to overall mental state task performance for the bilingual children. Instead, proficiency in English was only related to a higher overall mental state task score when children had a high level of proficiency in Spanish. These results coincide with past research with bilinguals in that a higher level of proficiency across languages predicts better performance on specific cognitive tasks (Bialystok, 1988, 2001; Bialystok & Fang, 2009; Carlson & Meltzoff, 2008; Cummins, 1976).

While the current study does provide evidence of a relationship between higher proficiency across languages and better performance on mental state tasks for bilinguals, the question still remains of the mechanism through which this occurs. Many researchers have suggested that bilingualism leads to better inhibitory control skills because bilinguals must inhibit one language while activating the other in various sociolinguistic situations (Costa, Miozzo, & Caramazza, 1999). Thus, given evidence that bilingualism fosters better inhibitory control skills (see Bialystok & Craik, 2010, for a review) and that inhibitory control skills are related to children's performance on mental state tasks, at least among monolinguals

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

(Carlson, Mandell, & Williams, 2004; Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Frye, Zelazo, & Palfai, 1995; Hughes, 1998; Perner & Lang, 1999), it is possible that inhibitory control is the mechanism through which language proficiency affects performance on mental state tasks (Rubio-Fernandez & Glucksberg, 2011). However, Nguyen and Astington (2014) found that bilingual children's performance on a working memory task, and not their performance on an inhibitory control task, mediated performance on false belief tasks (but only after language proficiency was controlled). Furthermore, there is evidence that, in some populations, language abilities are better predictors of performance on mental state tasks than inhibitory control abilities. For example, de Villiers and de Villiers (2012) found that for deaf children who were not exposed to sign language, language abilities (both expressive vocabulary and syntax comprehension), and not executive function abilities, predicted performance on false belief tasks. Thus, perhaps a more parsimonious explanation for the current findings is that measuring bilingual children's abilities in both languages gives a more accurate picture of their overall language skills, and that language skills aid mental state task performance regardless of the language they are acquired in. So, for example, if learning mental state terms and complement syntax are important for developing an understanding of mental states, perhaps children only need to learn those terms or constructions in one language to start to show better mental state reasoning skills regardless of the testing language. To investigate these various explanations, future research should include measures of inhibitory control as well as more comprehensive tests of language abilities (both within and across languages) when investigating bilingual children's performance on mental state tasks. Specific areas of language development, such as knowledge of mental state terms and complement syntax, that have been shown to be related to mental state reasoning in monolingual populations, should be explored in bilingual populations as well.

With regard to the bilingual and monolingual children's comparative performance on the individual tasks, interestingly, bilingual children did not show an advantage over monolingual children on the false belief tasks included in Wellman and Liu's (2004) scale. In fact, bilingual children showed a disadvantage on the Explicit False Belief task, while Kovács (2009) and Rubio-Fernandez and Glucksberg (2011) found evidence of an outright bilingual advantage on this task. Given evidence that higher proficiency across languages was related to better performance on the mental state tasks, it is possible that there was too much variation in language proficiency in the current sample of bilingual children (range of English PPVT standardized scores = 66–117; range of Spanish TVIP standardized scores = 59–139) to show an advantage over monolingual children on the false belief tasks. It is compelling, however, that despite the wide range of

English proficiency scores within the bilingual group, and the fact that the mental state tasks were administered in English, English proficiency alone was not significantly related to overall mental state task performance for the bilingual group as it was for the monolinguals. This suggests that the relationship between language proficiency and performance on mental state tasks does indeed differ between bilingual and monolinguals.

With regard to children's performance on the other mental state tasks, bilinguals were shown to outperform monolinguals on one task (i.e., diverse desires), perform worse than monolinguals on one task (i.e., explicit false belief), and perform similarly to monolinguals on the remaining mental state tasks. As discussed in the 'Introduction', there is a tendency to publish results that show a bilingual advantage, while failing to publish results that show mixed or null results. In fact, de Bruin, Treccani, and Della Sala (2014) posit that one reason for this bias is that researchers submit for publication results from tasks that show a bilingual advantage, but do not include results from tasks that show similar performance between bilinguals and monolinguals. Given de Bruin, Treccani, and Della Sala's findings, it seems likely that the results from the current study reflect a more accurate picture of bilingual and monolingual individuals' relative performance on cognitive tasks. Namely, bilinguals outperform monolinguals in some cases and not others. Unfortunately, the differing levels of performance of the monolinguals and bilinguals on the various mental state tasks included in the current study are difficult to interpret. Theoretically, researchers need to more fully explore whether passing a variety of mental state tasks requires the development of a single cognitive process that improves over time, or whether distinct, but somewhat inter-related skills, are needed to pass each task (Peterson, Wellman, & Liu, 2005; Peterson, Wellman, & Slaughter, 2012; Schick *et al.*, 2007; Wellman, Fang, & Peterson, 2011; Wellman & Liu, 2004). Furthermore, given that the vast majority of research exploring the relationship between children's language abilities and their ability to reason about mental states has focused on performance on false belief tasks, it is unknown how language abilities differentially affect performance on each of Wellman and Liu's (2004) tasks. Thus, future research is needed to gain a better understanding of the relationship between mental state reasoning and language proficiency (within and across languages) to clarify why bilinguals sometimes perform differently than monolinguals on some mental state tasks and not others.

One final consideration is the role of intelligence in the bilingual children's performance on the mental state and language tasks. While the results suggest that higher proficiency across languages fosters bilingual children's ability to reason about mental states, it is possible that children with higher proficiency in both languages are generally more intelligent, and thus likely to perform better on the mental state tasks than children with lower language proficiency. However, if higher intelligence was driving the performance of the highly proficient bilinguals, it is likely that they would have performed better than the less proficient bilinguals on a wider variety of mental state tasks, particularly the most difficult tasks in the scale. Instead, the highly proficient bilinguals showed an advantage on the easiest task in the scale (i.e., diverse desires) as well as one of the middle tasks in the scale (i.e., explicit false belief). Thus, the explanation that intelligence is driving the differences observed within the bilingual group seems unlikely. However, further research could be conducted to better understand the role of intelligence in children's performance on these tasks.

In sum, results from the current study support the prevalent theoretical perspective that language proficiency is related to mental state reasoning abilities among monolinguals. However, results also suggest that the relationship between language development and the development of mental state reasoning functions somewhat differently for bilingual individuals. Namely, proficiency across languages, not English proficiency alone, predicted performance on the mental state tasks for the bilingual group. Because a relatively high proportion of the world's population is bilingual, at least to some degree, the current theoretical framework must be expanded to explain how language proficiency across multiple languages is related to children's developing understanding of mental states. Furthermore, efforts should be made to develop a more comprehensive theoretical framework to explain the skills required to reason about different mental states, and how language skills relate to understanding those mental states for both bilingual and monolingual individuals.

REFERENCES

Aiken, L. S. & West, S. G. (1991). *Multiple regression: testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.

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

Astington, J. W. & Baird, J. A. (2005). *Why language matters for theory of mind*. New York: Oxford University Press.

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

Berguno, G. & Bowler, D. M. (2004). Communication interactions, knowledge of a second language, and theory of mind in young children. *Journal of Genetic Psychology: Research and Theory on Human Development* **165**(3), 293–309.

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

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

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism* **12**, 3–11.

Bialystok, E. & Craik, F. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science* **19**, 19–23.

https://doi.org/10.1017/S0305000915000276 Published online by Cambridge University Press

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. & Senman, L. (2004). Executive processes in appearance-reality tasks: the role of inhibition of attention and symbolic representation. *Child Development* **75**, 562–79.

Bloom, P. & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* **77**(1) B25–31.

Carlson, S. M., Mandell, D. J. & Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Developmental Psychology* **40**(6), 1105–22.

Carlson, S. M. & Meltzoff, A. (2008). Bilingual experience and executive functioning in young children. *Developmental Science* **11**, 282–98.

Carlson, S. M. & Moses, L. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development* **72**(4), 1032–53.

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

Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language* **41**(3), 365–97.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism* **9**, 1–43.

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of Cross-Cultural Psychology* **9**, 131–49.

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.

de Bruin, A., Treccani, B. & Della Sala, S. (2014). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science* **26**(1), 99–107.

de Villiers, J. & de Villiers, P. (2009). Complements enable representation of the contents of false beliefs: the evolution of a theory. In S. Foster-Cohen (ed.), *Language acquisition* (pp. 169–195). Basingstoke: Palgrave-McMillian Press.

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**, 1037–60.

de Villiers, P. & de Villiers, J. (2012). Deception dissociates from false belief reasoning in deaf children: implications for the implicit versus explicit theory of mind distinction. *British Journal of Developmental Psychology* **30**, 188–209.

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.

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**, 92–102.

Frye, D., Zelazo, P. & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development* **10**(4), 485–527.

Gathercole, V. C., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., ... Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology* **5**(11), 1–14.

Goetz, P. (2003). The effects of bilingualism on theory of mind development. *Bilingualism* **6**, 1–15.

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.

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *British Journal of Developmental Psychology* **16**(2), 233–53.

Kovács, A. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science* **12**, 48–54.

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**, 1130–44.

Macnamara, J. (1966). *Bilingualism and primary education*. Edinburgh: Edinburgh University Press.

Milligan, K., Astington, J. W. & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development* **77**, 622–46.

Montgomery, D. E. (2005). The developmental origins of meaning for mental terms. In J. Astington & J. Baird (eds), *Why language matters for theory of mind* (pp. 106–122). New York: Oxford University Press.

Morton, J. & Harper, S. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science* **10**(6), 719–26.

Nguyen, T. & Astington, J. W. (2014). Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism: Language and Cognition* **17**(2), 396–409.

Oller, D. K. & Eilers, R. (2002). *Language and literacy in bilingual children*. Clevedon: Multilingual Matters.

Paap, K. & Greenberg, Z. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* **66**(2), 232–58.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F. & Fazio, F. (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**, 170–82.

Perner, J. & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences* **3**(9), 337–44.

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. & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development* **76**, 502–17.

Peterson, C. C., Wellman, H. & Slaughter, V. (2012). The mind behind the message: advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or asperger syndrome. *Child Development* **83**(2), 469–85.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research* **21**, 301–16.

Rubio-Fernandez, P. & Glucksberg, S. (2011). Reasoning about other people's beliefs: bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **38**(1), 211–7.

Schick, B., de Villiers, P., de Villiers, J. & Hoffmeister, R. (2007). Language and theory of mind: a study of deaf children. *Child Development* **78**, 376–96.

Sera, M. D., Bales, D. W. & del Castillo Pintado, J. (1997). Ser helps Spanish speakers identify 'real' properties. *Child Development* **68**, 820–31.

Slade, L. & Ruffman, T. (2005). How language does (and does not) related to theory of mind: a longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology* **23**, 117–41.

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.

US Census Bureau. (2011). *Language use in the United States*. Online: <http://www.census.gov/content/dam/Census/library/publications/2013/acs/acs-22.pdf>.

Wellman, H., Fang, F. & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: longitudinal perspectives. *Child Development* **82**(3), 780–92.

Wellman, H. & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development* **75**(2), 523–41.