Child Development, November/December 2003, Volume 74, Number 6, Pages 1709–1719

False Belief Reasoning and the Acquisition of Relative Clause Sentences

Mark Smith, Ian Apperly, and Victoria White

Perner (1991) has claimed that the linguistic structures and reasoning tasks mastered by 4-year-olds share a requirement to handle metarepresentation. In contrast, de Villiers (2000) has argued that they share a requirement to handle misrepresentation. In the current study, a correlation is observed between success on false belief tasks and the acquisition of relative clause sentences. This correlation is not predicted by de Villiers's account because such sentences do not require the handling of misrepresentation, but it is consistent with Perner's account because such sentences do require the handling of metarepresentation. It is proposed that only an account that integrates the accounts of both de Villiers and Perner can explain extant data on language and cognition in 4-year-olds.

Accounts of the relation between language and thought typically argue either that language does not determine thought or that differences between individual languages give rise to cognitive differences between speakers of different languages (Bloom & Keil, 2001). In contrast to these two approaches, de Villiers (2000) has argued that the acquisition of a linguistic structure common to all languages, namely, the embedded complement sentence, determines the development of a cognitive ability found among speakers of all languages, namely, the ability to handle misrepresentation. For de Villiers, embedded complement sentences are critical to the development of this ability because they ''uniquely allow the representation of false propositions'' (p. 90). Consider, for instance, the following sentences:

1. Bob was convinced that *the Russians landed on the moon first*.
2. The postman ran away from the dog that *bit him*.

Sentence 1 is an embedded complement sentence with an (italicized) complement clause embedded inside a matrix clause. In contrast, Sentence 2 is a relative clause sentence with an (italicized) relative clause embedded inside a matrix clause. Crucially, although a speaker may be correct to assert an embedded complement sentence whose complement clause represents a false proposition, a speaker cannot be correct to assert a relative clause sentence whose relative clause represents a false proposition (de Villiers & de Villiers, 2000, p. 198). Thus, a speaker would be correct to assert Sentence 1 if Bob had actually been convinced that the Russians landed on the moon first but a speaker would be wrong to assert Sentence 2 if the dog had not actually bitten the postman. For this reason, embedded complement sentences, unlike other sentence types such as relative clause sentences, require an understanding that propositions can misrepresent reality (de Villiers & Pyers, 2002, p. 1038). According to de Villiers (2000), it is as a result of this unique property that the acquisition of embedded complement sentences gives rise to the ability to handle misrepresentation.

In a test of de Villiers's theory, de Villiers and Pyers (2002) tracked both the acquisition of embedded complement sentences and relative clause sentences and the development of the ability to handle misrepresentation in false belief tasks. As de Villiers and Pyers stated, such a design allowed them to ''separate the critical feature, namely sentential complements with mental/communication verbs, from other forms of complex sentence which play no role in our theoretical argument, such as relative clauses'' (p. 1044). In line with de Villiers's theory, they found that although the acquisition of embedded complement sentences correlated with and predicted the development of false belief reasoning, the acquisition of other sentence types such as relative clause sentences did not.

Mark Smith, Department of Psychology, University of Warwick; Ian Apperly and Victoria White, Department of Psychology, University of Birmingham.

This research was supported by Grants R000237932 and R000239002 from the Economic and Social Research Council, England. The authors would like to thank Sarah Beck, Jill Boucher, Gordon Brown, Teresa McCormack, Peter Mitchell, Kevin Riggs, Liz Robinson and Glyn Thomas for advice and support. Thanks also to Kate Broomfield for running a pilot study.

Correspondence concerning this article should be addressed to Mark Smith, Department of Psychology, University of Warwick, Coventry CV4 7AL England. Electronic mail may be sent to psscz@dredd.csv.warwick.ac.uk.

In contrast to de Villiers, Perner (1991, 1995) has argued that the linguistic structures and reasoning tasks mastered by 4-year-olds are linked by a shared requirement to handle metarepresentation. Metarepresentation arises when a representation of an event is embedded inside a representation of an event. Consequently, a metarepresentation represents both an event and a representation of an event (Perner, 1991; Stoichita, 1997). Thus, as Stoichita (1997) observed, many 17th-century Dutch paintings are metarepresentational because they embed representations of events (usually in the form of paintings, letters, or maps) inside other events (typically, everyday activities occurring in household rooms). Similarly, the unexpected transfer test is metarepresentational because it embeds a (mis)representation of an event (i.e., Maxi's belief regarding the storing of the chocolate) inside an event (i.e., Maxi's search for the chocolate in the kitchen). Three-year-olds appear to fail this task because they represent only the event and not the (mis)representation that it contains. Crucially, Perner (1991) has argued that the ability to handle misrepresentation is one of the abilities that emerges as a result of the development of the ability to handle metarepresentation by 4-year-old children. By subsuming the ability to handle misrepresentation under the ability to handle metarepresentation in this way, Perner was able to predict an observed correlation between the development of the ability to pass tasks requiring the handling of both metarepresentation and misrepresentation (such as false belief tasks) and tasks requiring the handling of metarepresentation but not misrepresentation.

As an example of the latter type of task, Perner (1991) cited evidence that 3-year-olds (in contrast to 4-year-olds) are unable to grasp that a person viewing a drawing of a turtle from different angles may give different interpretations of it (e.g., Flavell, Everett, Croft, & Flavell, 1981; Masangkay et al., 1974). Although success on such a task does not require the handling of misrepresentation (because both interpretations are valid), it does require the handling of metarepresentation. Thus, the task requires children to represent an event (i.e., in which they and another person look at a drawing of a turtle) in which a representation of an event (i.e., the other person's representation of the turtle's position) is embedded. As in the false belief task, children fail such a task when they represent only the event and not the representation that it contains. Crucially, although Perner's account predicted that success on Masangkay et al.'s (1974) task would correlate with success on false belief tasks, de Villier's (2000) account, because it confines itself to the handling of misrepresentation, it does not.

Perner's (1991) account also predicted de Villiers and Pyers's (2002) finding that the development of false belief reasoning correlates with the acquisition of embedded complement sentences because such sentences involve the handling not just of misrepresentation but of metarepresentation also. Consider, for instance, the following embedded complement sentence:

3. Jane claimed that the Americans landed on the moon first.

The complement clause of Sentence 3 asserts that a moon landing event occurs at an initial time (T1) and its matrix clause asserts that a claiming event occurs at a later time (T2). However, because the complement clause is embedded in the matrix clause, the sentence also asserts that the complement clause event recurs as a part of the matrix clause event at T2. Clearly, it cannot be that the complement clause event is recurring at T2 in the form of an actual event because this would entail the absurd implication that the complement clause event actually occurs twice (i.e., once at T1 and once at T2). Instead, the sentence asserts that the complement clause event recurs at T2 in the form of a claim that represents the moon landing event. Crucially, because it recurs not in the form of an actual event but in the form of a representation of an event, the complement clause event can form a part of the matrix clause event without actually occurring twice. Also, however, because Sentence 3 represents both an event and a representation of an event, it meets the definition of a metarepresentation and we should therefore expect that the acquisition of such a sentence should correlate with the development of false belief reasoning on Perner's (1991) account. As such, Sentence 3 contrasts with the following coordinated clause sentence:

4. Jane claimed and the Americans landed on the moon first.

Because Sentence 4 does not embed the clause representing the moon landing event within the clause representing the claiming event, it does not assert that the claiming event involves a claim about the moon landing event. Thus, although it represents a moon landing event it does not also represent a representation of it and does not meet the definition of a metarepresentation. Consequently, on Perner's (1991) account we should not expect its acquisition to correlate with the development of false belief reasoning.

Crucially, such an analysis can also be applied to relative clause sentences and suggests that they meet the definition of a metarepresentation too. Thus, consider again Sentence 2:

2. The postman ran away from the dog that *attacked him*.

The relative clause of Sentence 2 asserts that an attacking event occurs at an initial time (T1) and its matrix clause asserts that a fleeing event occurs at a later time (T2). However, because the relative clause is embedded in the matrix clause, the sentence also asserts that the relative clause event recurs as a part of the matrix clause event at T2. Clearly, it cannot be that the relative clause event is recurring at T2 in the form of an actual event because this would entail the absurd implication that the relative clause event actually occurs twice (i.e., once at T1 and once at T2). Instead, the sentence asserts that the relative clause event recurs at T2 in the form of a property that represents the attacking event. Thus, the sentence asserts that the matrix clause event involves not just a postman running away from a dog but rather a postman running away from a dog that has the property of having attacked him. Crucially, because it recurs not in the form of an event but in the form of a representation of an event, the relative clause event can form a part of the matrix clause event without actually occurring twice. Also, however, because Sentence 2 represents both an event and a representation of an event, it meets the definition of a metarepresentation and we should expect a correlation between the acquisition of such a sentence and false belief reasoning on Perner's (1991) account.

Whether the ability to handle a sentence that embeds a relative clause within a matrix clause develops at the age of 4 as Perner's (1991) account predicts has been investigated by Hamburger and Crain (1982). Hamburger and Crain observed that children often misinterpret a relative clause sentence such as Sentence 5 as a coordinated clause sentence such as Sentence 6:

5. The cow bumped the horse that tickled the cat.
6. The cow bumped the horse and the horse tickled the cat.

Because neither of the two clauses of a coordinated clause sentence are embedded within the other, to misinterpret a relative clause sentence as a coordinated clause sentence is to treat its relative clause as if it were unembedded rather than embedded within a matrix clause. To determine the age at which children treat relative clauses as embedded clauses, Hamburger and Crain (1982) got children to act out the meaning of spoken relative clause sentences such as Sentence 5 with toys. Crucially, the events represented by the leftmost clauses of a coordinated clause sentences do not happen after the events represented by their rightmost clauses (hence, the oddity of "Harry ran out of the house and jumped out of bed," as Townsend & Ravelo, 1980, observed)—a fact that children are aware of from the age of 2 years 6 months and onward (e.g., Bloom, 1991). Consequently, when children misinterpret a sentence such as Sentence 5 as a coordinated clause sentence, they tend to avoid acting out the bumping of the horse after the tickling of the cat. It follows therefore, as Hamburger and Crain argued, that children signal that they are interpreting a sentence such as Sentence 5 as a relative clause sentence rather than as a coordinated clause sentence when they act out the bumping of the horse after the tickling of the cat because relative clause events can happen before matrix clause events. Children also signal such an interpretation when they act out only the matrix clause event. Specifically, such a response indicates an awareness that acting out the relative clause event is unnecessary because it is already contained within the matrix clause event as a property of one of its objects. As Hamburger and Crain observed, the correct response to the request "Show me the pen you wrote your dissertation with" is not to rewrite your dissertation but simply to show a pen. Crucially, Hamburger and Crain found that these two signals were given by 3-year-olds on 27% of trials, by 4-year-olds on 56% of trials, and by 5-year-olds on 90% of trials, thereby indicating that the ability to treat relative clauses as the embedded clauses of relative clause sentences rather than as the unembedded clauses of coordinated clause sentences tends to emerge between the ages of 4 and 5.

Hamburger and Crain's (1982) study has been reinforced by other studies showing either that 3-year-olds misinterpret relative clause sentences as coordinated clause sentences or that the handling of relative clause sentences by 3-year-olds is very poor and is significantly better in 4- and 5-year-olds (Brown, 1971; Clancy, Lee, & Zoh, 1986; de Villiers & Roeper, 1995; de Villiers, Tager-Flusberg, Hakuta, & Cohen, 1979; Goodluck & Tavakolian, 1982; Hakuta, 1982; Lebeaux, 2000; Sheldon, 1974; Tavakolian, 1981). Nevertheless, contrary to Hamburger and Crain, studies have claimed that 3-year-olds can handle relative clause sentences. Thus, Hakansson and Hansson (2000) found that a group of children

with a mean age of 3 years 5 months performed above chance in acting out relative clause sentences such as Sentence 7 and Lempert and Kinsbourne (1980; cf. McKee, McDaniel, & Snedeker, 1998) found that a group of children with a mean age of 2 years 10 months successfully acted out relative clause sentences such as Sentence 8 on 90% of trials:

7. The girl chases the dog that is big.
8. It's the horse that kisses the cow.

Crucially, the relative clause sentences employed in such studies represent only a single event and are semantically equivalent to an unembedded single clause sentence. Thus, Sentence 7 is equivalent to ''The girl chases the big dog'' and Sentence 8 is equivalent to ''The horse kisses the cow'' (on this point, cf. Diessel & Tomasello, 2000; Hamburger & Crain, 1982). As a result, these sentences do not allow children to manifest either of the two behaviors that (as Hamburger and Crain, 1982, have shown) indicate that they are interpreting relative clauses as embedded clauses rather than unembedded clauses. Thus, in such studies children can neither act out the relative clause event before the matrix clause event nor act out the matrix clause event but not the relative clause event. As a result, such studies have failed to demonstrate that 3-year-olds treat relative clause sentences as sentences featuring an embedded relative clause rather than as sentences whose clauses are unembedded.

Overall, then, language acquisition studies indicate that the ability to embed a relative clause inside a matrix clause emerges between the ages of 4 and 5. Because such an ability requires the handling of metarepresentation and because the age at which it emerges is similar to that of false belief reasoning, such studies provide indirect support for Perner's (1991) claim that the linguistic structures and reasoning tasks mastered by 4-year-olds are linked by a requirement to handle metarepresentation. Crucially, de Villiers and Pyers (2002) may have failed to observe a correlation between the onset of false belief reasoning and the acquisition of relative clause sentences because they did not apply the criteria of Hamburger and Crain (1982) for determining whether children are treating relative clause sentences as sentences with embedded relative clauses (see also Hale & Tager-Flusberg, 2003). In the present experiment, we tested for a correlation between the development of false belief reasoning and the acquisition of relative clause sentences using a test of language comprehension (namely, the truth value judgment task of Crain & Thornton, 1998) which uses one of Hamburger and Crain's criteria. In this task, children judge whether events acted out by an experimenter match a spoken sentence. The following sentences were included in this test:

10. The girl kicked the man that jumped over the wall.
11. The girl kicked the man that is wearing a hat.
12. The girl jumped on the chair and the pig chased the ball.

Sentence 10 was termed a double-event relative clause sentence because it refers to a sequence of two temporally discrete events and requires the experimenter to act out two separate events one after the other. Sentence 11 was termed a single-event relative clause sentence because it refers to a single event occurring at a single point in time and requires the experimenter to act out only a single event. Sentence 12 was termed a coordinated clause sentence because it refers to two separate events and is unembedded.

To test whether children could handle relative clause sentences correctly, the experimenter acted out the event denoted by the relative clause in Sentence 10 before the event denoted by its matrix clause. It was reasoned, following Hamburger and Crain (1982), that if children were misinterpreting the sentence as a coordinated clause sentence they should judge that the sentence did not describe the sequence of events acted out by the experimenter and should judge that the sentence did not correspond to the act out scenario. Also, it was reasoned, following Hamburger and Crain, that if children were interpreting Sentence 10 as a relative clause sentence, they should judge that the sentence did describe the sequence of events acted out by the experimenter and should judge that the sentence did correspond to the act out scenario. On the account of Perner (1991), then, we should expect successful judgments on trials involving double-event relative clause sentences such as Sentence 10 to correlate with success on false belief tasks because both require the ability to handle metarepresentation. Also, Perner's account would not predict such a correlation in sentences such as Sentence 11 or Sentence 12 because successful judgments on trials involving these sentences do not require an ability to handle metarepresentation. In contrast, the account of de Villiers (2000) does not predict a correlation between false belief reasoning and the handling of double-event relative clause sentences because such sentences do not require the handling of misrepresentation.

Method

Participants

Having excluded 7 children for refusing to pay attention, to participate, or to give more than one type of answer, 56 children composed of 28 three-year-olds (8 boys and 20 girls) aged between 3 years 3 months and 4 years 0 months (*M* age = 3 years 9 months) and 28 four-year-olds (12 boys and 16 girls) aged between 4 years 3 months and 5 years 2 months (*M* age = 4 years 9 months) were tested. The children were recruited from a primary school in London and were mostly from middle-class families. The racial composition of the sample was 80% White, 16% Black, and 4% other. All of the children spoke English as their first language in the home.

Procedure

In tests of children's handling of relative clause sentences, design features that increase task difficulty can mask linguistic competence (Crain & Thornton, 1998). To guard against this, we designed an experiment free from features that increased task difficulty. To minimize sentence difficulty, we used only OS relative clause sentences (such as Sentences 10 and 11) in which the object of the matrix clause is the subject of the relative clause because it has been shown that such sentences are easier for children to handle than many other relative clause sentence types (Hakuta, 1982). Sentence difficulty was also reduced by employing relative clauses featuring only a single animate entity because children are better at handling relative clauses that feature a single animate entity rather than two animate entities (Goodluck & Tavakolian, 1982). Sentence difficulty was also minimized through the use of very familiar nouns (e.g., car, dog, hat, table, etc.) and verbs (e.g., kick, touch, jump over, bite, etc.), and before the experiment, we checked that each child knew the correct name for all of the toys and actions. We avoided using act out, production, picture comprehension, and elicited imitation tasks as our pilot studies and previous studies had suggested they increase task difficulty (e.g., Crain & Thornton, 1998; de Villiers & Roeper, 1995; Hakansson & Hansson, 2000). Instead, we chose the truth value judgment task in which children assess whether a sentence matches a scene acted out by an experimenter with toys because both our pilot studies and Crain and Thornton (1998) indicated that it delivers a precise and reliable estimate of the child's linguistic competence.

Because children find it easier to judge toys than adults, we also employed a design in which children judged descriptions of enacted scenes that an experimenter repeated aloud after ''Harry the hippo'' had whispered them to her. To emphasize that Harry could be either right or wrong, the experimenter stated that Harry could be either right or wrong and performed two trials in which she correctly judged Harry to be right and wrong. This was also emphasized in two pre-experimental trials in which the child was required to judge Harry to be right and wrong. We guarded against the effects of fatigue by dividing the experiment into two short testing sessions separated by an hour. In the initial session, children were tested on the easier single-event relative clause sentences rather than the harder double-event relative clause sentences to increase their confidence in their ability to do the task.

It has also been argued that performance is facilitated by including a second toy similar to the toy referred to by the head of the relative clause (i.e., the noun of which the relative clause forms a property, such as *man* in Sentences 10 and 11) in the act-out scene (Crain & Thornton, 1998). It is argued that this allows the relative clause to fulfill its pragmatic function of contrastively referring, that is, of discriminating a target entity from a set of similar entities on the basis of a property. Thus, it is argued that including an extra toy man during the acting out of Sentence 10 would allow a property denoted by the relative clause (i.e., of jumping over the wall) to discriminate a target entity (i.e., the toy man denoted by the head of the relative clause) from a set of similar entities (i.e., the bystanding man). Because no previous study has directly compared children's handling of contrastively referring sentences with their handling of sentences that do not contrastively refer, it is uncertain whether contrastive reference facilitates performance. To assess this, the experiment featured conditions in which the sentences did and did not contrastively refer.

Three types of sentence were employed in the language comprehension tasks. Instances of these sentence types are provided in Examples 10, 11, and 12:

10. *Double-event relative clause*: The girl kicked the man that jumped over the wall.
11. *Single-event relative clause*: The girl kicked the man that is wearing a hat.
12. *Coordinated clause*: The girl kicked the man and the dog chased the ball.

The relative clause sentences were always of the OS type. The double-event relative clause sentences

featured two verbs in the past tense (i.e., kicked and jumped) and the single-event relative clause sentences featured a past tense verb in the matrix clause (i.e., kicked) and a present tense verb in the relative clause (i.e., is wearing). Also, they featured an animate noun as the subject of the matrix clause (i.e., girl), an animate noun as the subject of the relative clause (i.e., man), and an inanimate noun as the object of the relative clause (i.e., wall and hat). The coordinated clause sentences featured two past tense verbs (i.e., kicked and chased) and four distinct nouns, at least one of which was inanimate (i.e., ball).

In trials featuring double-event relative clause sentences, the experimenter first acted out the event referred to by the relative clause and then acted out the event referred to by the matrix clause. Thus, in Sentence 10, the experimenter first acted out an event in which a man jumps over a wall using a toy man and a toy wall and subsequently acted out an event in which a girl kicks a man using a toy girl and the same toy man. In trials featuring single-event relative clause sentences, the experimenter acted out a single event. Thus, in Sentence 11, the experimenter acted out an event in which a girl kicks a man wearing a hat using a toy girl and a toy man wearing a toy hat. In trials featuring coordinated clause sentences, the experimenter first acted out the event referred to by the leftmost clause and subsequently acted out the event referred to by the rightmost clause. Thus, in Sentence 12, the experimenter first acted out an event in which a girl kicks a man using a toy girl and a toy man and subsequently acted out an event in which a dog chases a ball using a toy dog and a toy ball. Rather than work from a fixed set of sentences, the experimenter freely combined toys and actions to create sentences featuring new combinations of nouns and verbs for each child. This ensured that, although each child was tested on each noun or verb only once, each noun or verb occurred (approximately) equally often in each of the sentence types over the entire set of children, thereby making a balanced contribution to the results.

Each child was tested on four trials featuring double-event relative cause sentences and four trials featuring single-event relative clause sentences where the events acted out matched the sentence spoken by the experimenter and the correct response was right. In both the single- and double-event relative clause sentences, half of these trials featured contrastive sentences and half featured noncontrastive sentences. Thus, in the noncontrastive version of Sentence 10, the events were acted out using a toy man, girl, and wall, but in the contrastive version of the sentence a second toy man was present in the act-out event but did not take part in the action. The contrastive toy was always as prototypical an example of the noun as the toy to which the noun referred. There were also eight trials (four featuring double-event relative clause sentences and four featuring single-event relative clause sentences) in which the correct answer was always wrong to guard against false positives from children always answering ''right.'' Thus, the experimenter might act out an event in which a horse jumped over a car and then act out an event in which the same horse kicked a cow and utter the incorrect description: ''The horse kicked the cow that jumped over a car.'' We had two reasons for not scoring these wrong trials. First, such an action could be described by SS (subject-subject) relative clause sentences in which the subject of the matrix clause is the subject of the relative clause (e.g., ''The horse that jumped over a car kicked the cow''). Because these sentences are easier for children to comprehend than OS relative clause sentences (e.g., Hakuta, 1982; Tavakolian, 1981), it may be that the task of judging an incorrect trial such as Sentence 16 to be incorrect is easier than the task of judging a correct trial to be correct. Second, it was impossible to be sure that children were rejecting wrong sentences on account of their syntax rather than for other reasons. There were four trials featuring coordinated clause sentences.

At the end of the first session children were tested on an unexpected transfer task (see Mitchell, 1996, for details). In this task, the experimenter narrated a story and acted out the events of the story using toys. In the story a young boy places a biscuit in a tin and his mother transfers the biscuit to a cupboard while he is absent from the room. Children were first asked memory check questions (''Where did Johnny put the biscuit? Where is it now?'') and then asked a false belief question (''Where will Johnny look first for the biscuit?''). At the end of the second session they were tested on a deceptive box task (see Mitchell, 1996, for details). In this task, children were shown a Smarties candy box and asked what it contained. They were then shown that it contained pencils rather than Smarties. They were then told that a specific classmate (e.g., Adam) would be brought into the room and they were asked to predict what the classmate would think was in the box (''What will Adam think is in the box?'').

Results

First, we considered children's performance on the relative clause sentence comprehension task. Children were scored only for sentences where the

correct answer was ''right.'' Thus, each child heard two examples of each kind of relative clause sentence (single-event/contrastive reference, single-event/noncontrastive reference, double-event/contrastive reference, double-event/noncontrastive reference) and were given scores of 0, 1, or 2 according to the number of correct judgments in each case. A repeated measures analysis of variance (ANOVA) was computed with event type (single vs. double) and reference type (contrastive vs. noncontrastive) as within-subject factors, and age group (3-year-olds and 4-year-olds) as a between-subject factor. There was a significant main effect of event type, $F(1, 54) = 16.5$, $p < .001$. Collapsing across age and reference type, the mean for single-event sentences was 1.57 (.06) and the mean for double-event sentences was 1.22 (.08).

There was also a main effect of reference type, $F(1, 54) = 8.3$, $p < .01$. Collapsing across age and event type, the mean for contrastive sentences was 1.30 (.07) and the mean for noncontrastive sentences was 1.5 (.07). Age group had a significant overall effect, $F(1, 54) = 16.8$, $p < .001$, with 3-year-olds performing less well than 4-year-olds. There were no within- or between-subject interactions of any kind (all $ps > .1$).

Although 3-year-olds performed consistently less well than 4-year-olds, from the means reported in Table 1 it is unclear whether their performance was well above a zero baseline or close to a chance baseline. We therefore conducted chi-square analyses comparing the frequency with which children answered 0, 1, or 2 of each question type correctly, with the frequency expected on the basis of chance performance. These tests revealed that the performance of the 4-year-olds was above chance for all four types of relative clause sentence (all $ps < .005$). For the 3-year-olds it was above chance for the single-event/noncontrastive relatives ($p < .005$), and there was a strong trend toward an above-chance level of performance for the single-event/contrastive relatives ($p < .05$). In contrast, the 3-year-olds' performance did not differ significantly from chance for the two types of double-event relatives (both $ps > .25$). Such results are compatible with the view that, although 4-year-olds can handle both double- and single-event relative clause sentences, 3-year-olds are only above chance baseline performance on single-event relative clause sentences.

It is important to know whether the effect of event type in the relative clause sentences can be attributed solely to the greater cognitive demands of handling sentences describing two events. Thus, we next compared children's performance on double-event relative clause sentences (e.g., The horse kicked the cow that jumped over a car) with their performance on the coordinated clause sentences, which described one event following another event (e.g., The man touched the dog and the horse kicked the chair). In one ANOVA, scores on the coordinated clause sentences and the double-event/contrastive relative clause sentences (which, like the coordinated clause sentences, featured the use of four toys in the act-out scene) were entered as repeated measures and age was entered as a between-subject factor. There was a significant main effect of sentence type, $F(1, 54) = 15.6$, $p < .001$. Collapsing across age, the means for double-event/contrastive relative clause sentences and coordinated clause sentences were 1.14 (.09) and 1.57 (.08), respectively. There was a significant effect of age, $F(1, 54) = 19.3$, $p < .001$, but no significant interaction between age and sentence type ($p > .99$). A similar ANOVA was computed to compare scores on the coordinated and double-event/ noncontrastive relative clause sentences. Once again there was a significant main effect of sentence type, $F(1, 54) = 5.31$, $p < .05$. Collapsing across age, the means for double-event/noncontrastive relative clause sentences and coordinated clause sentences were 1.30 (.1) and 1.57 (.08), respectively. There was a significant effect of age, $F(1, 54) = 14.7$,

Table 1

Three- and 4-Year Olds' Mean (Standard Deviation) Out of a Possible Two Correct Responses on the False Belief Task, Coordinate Clause Sentences, and Four Types of Relative Clause Sentences

	Relative clause sentences					
	Single event		Double event			
Age	Noncontrastive reference	Contrastive reference	Noncontrastive reference	Contrastive reference	Coordinate clause sentences	False belief tasks
3 ($N = 28$)	1.57 (.57)	1.18 (.82)	1.07 (.66)	0.86 (.65)	1.29 (.76)	0.89 (.74)
4 ($N = 28$)	1.82 (.39)	1.71 (.60)	1.54 (.80)	1.43 (.69)	1.86 (.36)	1.50 (.75)

$p < .001$, but no interaction between age and sentence type ($p = .65$). Clearly, the simple fact that the double-event relative clauses described two events could not explain the problems they posed for children.

Finally, we looked at scores on the two false belief tasks. Children were scored between 0 and 2 depending on the number of correct answers they gave. Consistent with prior studies, the 4-year-old group performed significantly better than the 3-year-old group, $t(53) = 3.1$, $p < .01$, (see Table 1). Thus, children's ability to reason about false beliefs improves across the same time scale as their ability to comprehend all of the sentence types examined in this experiment. To test whether double-event relative clauses were uniquely predictive of children's performance on the false belief tasks, we carried out a stepwise multiple regression analysis with scores on the false belief tasks as the dependent variable. In the first step, age (in months) was entered, resulting in a significant change in R^2: $\Delta R^2 = .255$, $F(1, 54) = 18.4$, $p < .001$. In the second step, children's total score with the single-event relative clauses (contrastive+noncontrastive) was added with no significant change in R^2: $\Delta R^2 = .007$, $F(1, 53) = .499$, $p = .48$. In the third step, children's score (out of 2) with the coordinated clause sentences was added with no significant change in R^2: $\Delta R^2 = .007$, $F(1, 52) = .47$, $p = .50$. Finally, children's total score with the double-event relative clauses was added, resulting in a significant increase in the predictive value of the model: $\Delta R^2 = .106$, $F(1, 51) = 8.6$, $p < .01$. Thus, children's aptitude with double-event relative clauses predicted their belief reasoning ability independent of their ability to understand other sentences describing two events and their aptitude with relative clauses per se.

Discussion

Perner (1991) has claimed that the linguistic structures and reasoning tasks mastered by 4-year-olds are linked by a requirement to handle metarepresentation. In contrast, de Villiers (2000) has argued that they are linked by a requirement to handle misrepresentation. To test between these two accounts, the current study tested for a correlation between the development of false belief reasoning and the acquisition of relative clause sentences. As argued in the introduction, relative clause sentences do not require the handling of misrepresentation but they do require the handling of metarepresentation because they embed a relative clause event inside a matrix clause event. Consequently, testing the age at which children are able to treat such sentences as embedded sentences constitutes a test of when children acquire linguistic structures that require the ability to handle metarepresentation but not misrepresentation. In the current study, it was found that 4-year-olds succeeded on a task that required them to treat a relative clause sentence as a sentence with an embedded clause whereas 3-year-olds did not. Moreover, success on this task correlated with successful performance on false belief tasks. Reinforcing this, it was observed that both 3- and 4-year-olds succeeded on tasks that did not require them to treat either a relative clause sentence or a coordinate clause sentence as a sentence with an embedded clause. Moreover, success on these tasks did not correlate with successful performance on false belief tasks.

Such a pattern of results is neither predicted nor explained by the account of de Villiers (2000). Thus, because this account confines itself to linguistic structures and reasoning tasks that require the handling of misrepresentation, it does not predict and cannot explain the finding that the acquisition of a linguistic structure that does not require the handling of misrepresentation (namely, relative clause sentences) should correlate with the development of false belief reasoning. As de Villiers and Pyers (2002, p. 1044) stated, relative clause sentences "play no role" in de Villiers's account. Consequently, only a modification of de Villiers's account would allow it to predict and explain the current results. In contrast, such a pattern of results is consistent with the claim of Perner (1991) that the linguistic structures and reasoning tasks mastered at around age 4 are linked by a shared requirement to handle metarepresentation. Specifically, because the embedding of a relative clause event within a matrix clause event requires the handling of metarepresentation, we should expect the acquisition of this ability to correlate with the development of false belief reasoning on Perner's account.

Crucially, various aspects of the results allow a number of alternative explanations to be ruled out. Thus, the fact that there were more affirmative responses to single-event relative clause sentences and coordinate clause sentences than to double-event relative clause sentences indicates that the pattern of responses cannot be explained as the product of an affirmative bias because it is implausible that children would find single-event relative clause sentences and coordinate clause sentences more confusing than double-event relative clause sentences. Similarly, because 4-year-olds gave more affirmative responses than 3-year-olds, accounting

for the results in terms of an affirmative bias leads to the implausible conclusion that 4-year-olds found the task more confusing than 3-year-olds. Moreover, the fact that 3-year-olds could handle coordinate clause sentences indicates that the difference in the performance of the 3- and 4-year-olds cannot be explained in terms of a contrast in their ability to handle sentences representing two events. Furthermore, the fact that 3-year-olds could handle single-event relative clause sentences indicates that the differing performance of 3- and 4-year-olds cannot be reduced to a purely formal syntactic contrast because there are no extant formal analyses that discriminate between single- and double-event relative clause sentences. Moreover, it is unlikely that 3-year-olds' poor handling of double-event relative clause sentences is the result of an inability to handle sentences where the linear order of the clauses mismatches the temporal order of events because studies of before–after sentences have shown that 3-year-olds' ability in this regard is broadly equivalent to that of 4-year-olds (e.g., Townsend & Ravelo, 1980). It is also implausible that the study has underestimated 3-year-olds' ability to handle relative clause sentences as a result of employing a design that increased task difficulty. Thus, the experiment employed all of the design features noted in the previous literature for minimizing the difficulty of relative clause sentence comprehension tasks to avoid underestimating children's competence. Finally, it is doubtful that the estimate of the age at which children handle relative clause sentences obtained in the current study is an idiosyncratic result unrepresentative of typical performance. Thus, as argued earlier, many previous studies have shown that relative clause sentences are acquired by 4-year-olds.

The current study also has implications for whether the development of false belief reasoning depends on the prior acquisition of linguistic structures such as relative clause sentences and, in particular, for the account of this issue developed by Perner, Sprung, Zauner and Haider (2003). In this study, Perner et al. argued against de Villiers's (2000) claim that the acquisition of embedded complement sentences triggers false belief reasoning. In support of this view, they cited evidence that there is a significant developmental lag between the expression of false belief and the expression of desires and pretense even when they are expressed through formally similar structures (Custer, 1996; Tardiff & Wellman, 2000), and they provided new evidence demonstrating that German children use embedded complement sentences to express desire long before they use them to express false propositions. On the basis of such evidence, Perner et al. concluded contrary to de Villiers that "mastery of the grammatical structures that are required for communicating about particular mental states is not the determining factor … for developing an understanding of … these mental states" (p. 186). An obvious problem for such a view is that it is unable to account for the growing body of evidence that the development of false belief reasoning does depend on the prior acquisition of specific linguistic structures (Astington & Jenkins, 1999; de Villiers & Pyers, 2002; Gale, de Villiers, de Villiers, & Pyers, 1996; Ruffman, Slade, & Crowe, 2002; Tager-Flusberg, 2000). Crucially, the current data suggest a possible means of reconciling these two opposing sets of data. Thus, the current study suggests that although 3-year-olds already possess much of the formal machinery of relative clause sentences (as is evidenced by their ability to handle single-event relative clause sentences), the ability to handle certain nonformal, semantic aspects of such sentences (and in particular, the ability to handle metarepresentation through such sentences) only becomes apparent much later at around the time that children are developing false belief reasoning. As such, the study naturally suggests the possibility that false belief reasoning depends on the prior acquisition of nonformal rather than formal aspects of linguistic structures. Such a view is compatible with the data of studies such as Custer (1996), Perner et al. (2003), and Tardiff and Wellman (2000), which show only that the development of false belief reasoning is not dependent on the prior acquisition of formal aspects of linguistic structures. Also, because such a view assumes that the development of false belief reasoning depends on the acquisition of the ability to handle metarepresentational aspects of linguistic structures such as embedded complement and relative clause sentences, it is compatible with the data from studies such as de Villiers and Pyers (2002).

The view that 4-year-olds' successful performance on reasoning tasks requiring the handling of metarepresentation depends on the ability to handle metarepresentational aspects of particular linguistic structures constitutes a fusion of the accounts of Perner (1991) and de Villiers (2000). Thus, such a view follows Perner in affirming that 4-year-olds acquire the ability to handle metarepresentation. Also, such a view follows de Villiers in affirming that the cognitive transition that 4-year-olds undergo depends on changes in their handling of specific linguistic structures. Crucially, fusing the two accounts of de Villiers and Perner in this way

produces an account that accommodates a range of data that the two accounts in isolation cannot. Thus, in contrast to Perner, such an account both explains and predicts data suggesting that false belief reasoning depends on the prior acquisition of embedded complement sentences (e.g., de Villiers & Pyers, 2002). Also, in contrast to de Villiers, such an account both explains and predicts data suggesting that 4-year-olds acquire the ability to handle not just misrepresentation but also metarepresentation (e.g., Masangkay et al., 1974) and data indicating that the development of false belief reasoning correlates with the acquisition of double-event relative clause sentences (e.g., the current study). Finally, in contrast to both Perner and de Villiers, such an account both explains and predicts data demonstrating that the handling of embedded sentences other than embedded complement sentences such as relative clause sentences and before–after sentences correlates with their performance on false belief tasks at later time points (e.g., Ruffman et al., 2002).

Clearly, however, much further work investigating such a view of the relation between language acquisition and cognitive development in 4-year-olds is still required. Most obvious, although the data from the current study is compatible with the view that false belief reasoning depends on the ability to handle metarepresentational aspects of linguistic structures such as relative clause sentences, it does not provide direct support for such a claim because it only provides evidence of a correlation. As a result, a direct test of this claim in the form of a longitudinal version of the current study needs to be conducted. Also, work is required to determine whether the various types of embedded sentences acquired by 4-year-olds make differing contributions to the development of false belief reasoning and other types of reasoning ability. Even if it were to be shown, for instance, that relative clause sentences do play some role in triggering the development of false belief reasoning, it may still be that embedded complement sentences contribute more to such cognitive development because they require the handling of both misrepresentation and metarepresentation. Finally, work is required to probe the relation between language acquisition and the mastery of reasoning tasks that do not require the handling of metarepresentation by 4-year-olds. Thus, Frye, Zelazo, Brooks, and Samuels (1996) have shown that success on causal reasoning tasks that do not require the handling of metarepresentation first emerges in 4-year-olds. If a link can be demonstrated between the development of the ability to handle such tasks and the ability to handle sentences with embedded clauses, it would necessitate an even broader account of the relation between language acquisition and cognitive development in 4-year-olds than the one postulated here.

References

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*, 1311–1320.

Bloom, L. (1991). *Language development from two to three*. New York: Cambridge University Press.

Bloom, P., & Keil, F. C. (2001). Thinking through language. *Mind and Language*, *16*, 351–367.

Brown, H. D. (1971). Children's comprehension of relativised English sentences. *Child Development*, *42*, 1923–1936.

Clancy, P., Lee, H., & Zoh, M. (1986). Processing strategies in the acquisition of relative clauses: Universal principles and language specific realizations. *Cognition*, *24*, 225–262.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Custer, W. L. (1996). A comparison of young children's understanding of contradictory representations in pretense, memory and belief. *Child Development*, *67*, 678–688.

de Villiers, J. G. (2000). Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 83–123). Oxford, England: Oxford University Press.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and false belief. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Hove, England: Psychology Press.

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

de Villiers, J. G., & Roeper, T. (1995). Relative clauses are barriers to Wh- movement in young children. *Journal of Child Language*, *22*, 389–404.

de Villiers, J. G., Tager-Flusberg, H. B., Hakuta, K., & Cohen, M. (1979). Children's comprehension of relative clauses. *Journal of Psycholinguistic Research*, *8*, 499–518.

Diessel, H., & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics*, *11*, 131–151.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*, 99–103.

Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology, 32*, 120–131.

Gale, E., de Villiers, P., de Villiers, J., & Pyers, J. (1996). Language and theory of mind in oral deaf children. *Proceedings of the twentieth annual Boston University Conference on Language Development*. Sommerville, MA: Cascadilla Press.

Goodluck, H., & Tavakolian, S. (1982). Competence and processing in children's grammar of relative clauses. *Cognition*, *7*, 85–91.

Hakansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, *27*, 313–333.

Hakuta, K. (1982). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition*, *9*, 197–236.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, *6*, 346–359.

Hamburger, H., & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Ed.), *Language development, Vol. 1.: Syntax and semantics* (pp. 245–274). Hillsdale, NJ: Erlbaum.

Harris, P. (1996). Desires, beliefs and language. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 200–222). Cambridge, England: Cambridge University Press.

Lebeaux, D. (2000). *Language acquisition and the form of the grammar*. Amsterdam: Benjamins.

Lempert, H., & Kinsbourne, M. (1980). Preschool children's sentence comprehension: Strategies with respect to word order. *Journal of Child language*, *7*, 371–379.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, *45*, 357–366.

McKee, C., McDaniel, D., & Snedeker, J. (1998). Relatives children say. *Journal of Psycholinguistic Research*, *27*, 573–596.

Mitchell, P. (1996). *Acquiring a conception of mind: A review of psychological research and theory*. Hove, England: Psychology Press.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Perner, J. (1995). The many faces of belief: Reflections on Fodor's and the child's theory of mind. *Cognition*, *57*, 241–269.

Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). *Want that* is understood well before *say that, think that* and false belief: A test of de Villiers' linguistic determinism hypothesis. *Child Development*, *74*, 179–188.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*, 734–751.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, *13*, 272–281.

Stoichita, V. I. (1997). *The self-aware image: An insight into early modern meta-painting*. Cambridge, England: Cambridge University Press.

Tager-Flusberg, H. (2000). Language and understanding other minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 124–150). Oxford, England: Oxford University Press.

Tardiff, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin- and Cantonese-speaking children. *Developmental Psychology, 36*, 25–43.

Tavakolian, S. (1981). The conjoined clause analysis of relative clauses. In S. Tavakolian (Ed.), *Language acquisition and linguistic theory* (pp. 167–187). Cambridge, MA: MIT Press.

Townsend, D. J., & Ravelo, N. (1980). The development of complex sentence processing strategies. *Journal of Experimental Child Psychology, 29*, 60–73.