SOCIAL NEUROSCIENCE, 2006, 1 (3–4), 334–348

Psychology Press
Taylor & Francis Group

Intact first- and second-order false belief reasoning in a patient with severely impaired grammar

Ian A. Apperly, Dana Samson, Naomi Carroll, Shazia Hussain, and Glyn Humphreys
The University of Birmingham, Birmingham, UK

The retention of first-order theory of mind (ToM) despite severe loss of grammar has been reported in two patients with left hemisphere brain damage (Varley & Siegal, 2000; Varley, Siegal, & Want, 2001). We report a third, and more detailed, case study. Patient PH shows significant general language impairment, and severe grammatical impairment similar to that reported in previous studies. In addition we were able to show that PH's impairment extends to grammatical constructions most closely related to ToM in studies of children (embedded complement clauses and relative clauses). Despite this, PH performed almost perfectly on first-order false belief tasks and on a novel nonverbal second-order false belief task. PH was also successful on a novel test of "ToM semantics" that required evaluation of the certainty implied by different mental state terms. The data strongly suggest that grammar is not a necessary source of structure for explicit ToM reasoning in adults, but do not rule out a critical role for "ToM semantics." In turn this suggests that the relationship observed between grammar and ToM in studies of children is the result of an exclusively developmental process.

The relationship between language and theory of mind (ToM) is of interest for at least two reasons. First, for the large body of researchers interested in ToM, it seems plausible that language is a crucial source of informational input and representational structure, both in the development of ToM abilities, and in the on-line use of these abilities in the mature system (Astington & Baird, 2005; Astington & Jenkins, 1999; de Villiers & Pyers, 2002). Second, the relationship between language and ToM may be an important case study in a much larger debate about the relationship between language and complex conceptual thought (de Villiers & de Villiers, 1999; Carruthers, 2002). Most studies of the relationship between language and ToM have been conducted with children[1] and reveal consistent evidence of a strong and possibly causal relationship, with continuing debate about the relative importance of different components of the linguistic input (e.g., Astington & Baird, 2005). However, these developmental data are fundamentally ambiguous about the role of language in ToM because a developmental relationship could exist for two broad categories of reason: because language is constitutively involved in mature ToM abilities that children are developing; or because language is critical for the *developmental* process by which

[1] There is also an important literature demonstrating ToM abilities in nonverbal infants and non-human animals (e.g., Csibra, Gergely, Biro, Koos, & Brockbank, 1999; Emery, 2005; Onishi & Baillargeon, 2005). A discussion about the relationship between these abilities and those responsible for explicit reasoning using mental state concepts (beliefs, desires, intentions, etc.) is a crucial topic, but beyond the scope of the current article.

Correspondence should be addressed to: Ian Apperly, School of Psychology, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: i.a.apperly@bham.ac.uk

The research reported in this paper was supported by grants from the Leverhulme Trust, the MRC and the Stroke Association. We are very grateful to PH for his kind participation.

www.psypress.com/socialneuroscience DOI:10.1080/17470910601038693

the mature system is acquired. Data from adults who have impaired language following brain injury may help resolve this ambiguity by informing us directly about what, if any, aspects of language are necessary for mature ToM. To date just two such patients have been studied, revealing evidence of spared false belief reasoning abilities despite substantial general grammatical impairment (Varley & Siegal, 2000; Varley, Siegal, & Want, 2001). However, these studies fall short of ruling out a need for grammar in ToM and do not speak to the relationship between verbal semantic knowledge and ToM. In the current paper we present a third case of a patient who has been assessed on a broader range of ToM tasks, and tests of grammar and semantics designed specifically to assess the relationship between language and ToM.

Relationships between language and ToM in children

The developmental literature provides a wealth of information about the complex relationship between children's language, social communication and ToM (e.g., Astinton & Baird, 2005). Some authors suggest that the critical role of language is to provide a particularly rich medium for interpersonal understanding, within which ToM develops (e.g., Dunn & Brophy, 2005; Harris 2005). Others focus on the relative importance of syntax and semantics, as measured on standard developmental tests (Astington & Jenkins, 1999; Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). Finally, some authors have highlighted the importance of specific aspects of syntax or semantics. For example, there is evidence that 3- to 5-year-old children's performance on false belief tasks is specifically related to their ability to evaluate the certainty implied by different mental state terms (e.g., ''know'' is more certain than ''think''; Moore, Pure, & Furrow, 1990), to comprehend complex relative clause sentences (e.g., The woman pushed the man that opened the box; Smith, Apperly, & White, 2003), and to comprehend embedded complement clause sentences (e.g., John said that Aldrin was the first man on the moon; de Villiers & Pyers, 2002). de Villiers in particular suggests that acquiring the syntax of complementation may provide a critical representational basis for understanding beliefs in both children and adults (e.g., de Villiers & de Villiers, 2002; de Villiers & Pyers, 2002). However, for current purposes we take all three sets of findings as distinct indicators about which aspects of language might be most closely related to ToM in adults. It is noteworthy that neither of the existing studies of ToM in patients with aphasia specifically examined these aspects of language.

Existing studies of patients with aphasia

Varley and Siegal (2000) report the case of SA, whose grammatical abilities in both comprehension and production were significantly disrupted following a large left hemisphere brain lesion. On the PALPA (a standardized test battery for language assessment in aphasia) SA was significantly impaired on a variety of tests that are sensitive to grammatical impairment. He was 51% accurate on spoken sentence–picture matching (chance = 33%; PALPA subtest 55), 50% accurate on written sentence–picture matching (chance = 33%; PALPA subtest 56) and 52% accurate on auditory comprehension of verbs (chance = 50%; PALPA subtest 57). Besides his grammatical processing impairment, SA showed poor single word production (low intelligibility for spoken words and impaired written word production as assessed by a written picture naming task, PALPA subtest 54) but he showed relatively spared comprehension of single words (he scored above 87% correct on a spoken and written word–picture matching task, PALPA subtests 47 and 48, and he scored 71% correct on a task testing auditory comprehension of adjectives, subtest 57).

To assess ToM, SA was tested on false belief tasks in which he was shown containers with unexpected contents (e.g., a pill bottle containing buttons) and was asked to judge what a naïve person would think was inside, as well as to report what was really inside. To support SA's comprehension, he was pretrained to associate judgments about what another person thought with a cue card with the printed words ''X thinks,'' and judgments about reality with a cue card printed with the word ''really.'' In the test phase these cue cards were presented together with verbally presented questions: ''What does X think is in the [container]?'' and ''What is really in the [container]?'' SA was tested on multiple trials, each using a different container, which sometimes had unexpected contents (so the naïve person would have a false belief), and on other occasions had their expected contents (so the naïve person would have a true belief). SA's overall performance was significantly above chance on this task, leading Varley and Siegal (2000) to conclude that false belief reasoning (a critical component of ToM) can remain intact despite severe grammatical impairment.

The second case, reported by Varley, Siegal, and Want (2001) concerns patient MR, who, like SA, showed severe grammatical impairment following left hemisphere brain damage. Here also, the patient was impaired in tasks testing grammatical processing. MR scored 55% and 57% correct for spoken and written comprehension of reversible sentences, i.e., sentences in which the agent and patient relation cannot be understood by simply assigning the agent to the first noun encountered in the sentence (e.g., the man was killed by the lion). He scored 54% correct for the auditory comprehension of adjectives and verbs (PALPA subtest 57). The patient could still judge the grammaticality of non-embedded sentences (scoring 100% and 85% correct for auditory and written stimuli, respectively), but showed some difficulties for judging the grammaticality of embedded sentences (scoring 85% and 65% for auditory and written stimuli, respectively). Similarly to SA, single word production was also impaired (MR scored 65% and 5% correct in a spoken and written picture naming task, PALPA test 54) but single word comprehension was relatively spared (MR scored above 85% correct in a spoken and written synonym judgment task, he scored above 77% correct in a spoken and written word–picture matching task, and he scored 88% correct on the picture version of the Pyramids and Palm Trees task (Howard & Patterson, 1992).

To assess ToM, MR was tested on a picture-based task that required content to be supplied for the thought bubble of a cartoon character who had a true or a false belief. For example, one picture depicted a fisherman who has caught something on his fishing hook, and in the true belief condition the participant can see that he has actually hooked a fish, whereas in the false belief condition the participant can see that he has actually hooked a boot. MR scored 5/5 correct on both true and false belief trials.

These studies provide suggestive evidence about the independence of ToM from grammatical abilities, but the nature of the ToM assessments limits the certainty with which conclusions can be drawn. The ToM task in the study of SA involved repeated judgments about what a naïve person would think was inside containers the contents of which either corresponded with the outside appearance of the container (true belief condition) or were unexpected (false belief condition). Thus, in order to respond correctly on true or false belief trials, all SA had to do was report the appearance of the box whenever the cue-card ‘‘X-thinks’’ was displayed, and report the content of the box whenever the cue-card ‘‘reality’’ was displayed. It is by no means certain that SA did employ such a strategy, but this cannot be excluded. The ToM task in the study of MR could not be solved with this strategy. However, although 5/5 correct responses is clearly consistent with MR having no difficulty on this task, 5 test trials are insufficient for a reliable statistical comparison against chance performance. Thus, conclusions based upon this case must be cautious. Also, as noted above, neither patient was tested on the specific aspects of syntax and semantics that are most closely associated with ToM in studies of children.

The current study

In the current study we employ ToM tasks that were designed to minimize the chances that a patient could give correct answers by adopting a response strategy that did not involve ToM, and which allowed statistical comparison against chance for an individual patient. The test phase of our tasks was entirely nonverbal, so eliminating any suspicion that language could be necessary merely for comprehending the task or the test questions. For the first time, we were able to examine both first-order and second-order ToM using nonverbal tasks. We took explicit reasoning about false beliefs to be a good test case for whether language has a constitutive role in ToM. If language is unnecessary for the relatively complex problem of explicit belief reasoning, then it seems unlikely that language has a necessary role in less complex and less explicit ToM processes. The current study extends the tests of language ability beyond those examined by Varley et al. Like Varley et al. we included standardized language assessments from the PALPA test battery but in addition we examined specific aspects of syntax and semantics that the developmental literature suggests may be particularly important for ToM.

CASE REPORT

PH is a right-handed man with a degree in law. PH was 33 years old at the time of testing (2004–2005). Six years prior to the testing, PH suffered a left hemisphere stroke, which affected the left medial and superior temporal gyri as well as the left inferior and middle frontal gyri (see Figure 1). The most obvious consequences of his stroke included a left upper and lower limb hemiplegia as well as aphasia. Language problems included severe anomia, deep dyslexia, deep dysgraphia and difficulties in grammatical processing (see Kumar & Humphreys, 2006, for a full report on PH's language abilities; Table 1 provides a summary of data from background neuropsychological tests). Errors in picture naming were primarily semantic or "no response," and small percentages of semantic errors also arose in reading and in writing. The reading and writing of nonwords was particularly difficult. Single word comprehension showed a mild impairment for concrete items (e.g., on the Picture–word match task from PALPA) and a more severe deficit for low imagery stimuli (Table 1). Auditory lexical decision was spared, but there were problems in immediate repetition, which tended to be particularly severe for function words (4/15 functors correct vs. 24/45 summed across nouns, verbs and adjectives; $\chi^2(1) = 3.21$, $p = .07$). The functors used here included 7 prepositions (on, under) and 8 determiners (either, who). All the errors were omissions. PH was completely unable to repeat back a sentence (0/35, PALPA test 12). Object naming was impaired more for naming actions (using verbs) than for naming nouns (using items from Druks & Masterson, 2000); 138/300 for actions vs. 292/486 for objects, $\chi^2(1) = 14.85$, $p < .01$). He was also impaired at comprehending verbs and adjectives, scoring 55/82 for matching a spoken verb or adjective to an aurally presented sentence (PALPA test 57; controls at ceiling). On a test requiring matching of a locative term to a picture (PALPA test 58) he scored 15/24 (impaired). The problems with verbs, function and low imagery words all suggest that PH should have major problems in grammatical processing.

Verbal long-term and short-term memory was difficult to assess given PH's language impairment. He scored 36/50 on the word version of the Warrington Recognition Memory test, which is just below the mean level expected in patients with left temporal lobe damage (Warrington, 1984). Visual long-term memory was spared (PH score 67/72 on learning meaningless drawing; Violon & Seyll, 1984), and 43/50 on the face version of the Warrington test (within normal limits; Warrington, 1984). He showed a reduced visuo-spatial span (Corsi block = 3) and impaired verbal digit span (2). He was also impaired on a range of tests of executive function (see Table 1).

Grammatical processing

Test 1. Auditory and written sentence comprehension

Method. PALPA test 55 (sentence–picture matching: auditory version, 3 response options). PALPA test 56 (sentence–picture matching: written version, 3 response options).

Results. In the auditory version, Global score = 40/60 (67%). Most errors were made on reversible sentences (50% correct compared to 81% correct for non-reversible sentences) and errors on these sentences consisted of choosing the reverse action to that described in the sentence.

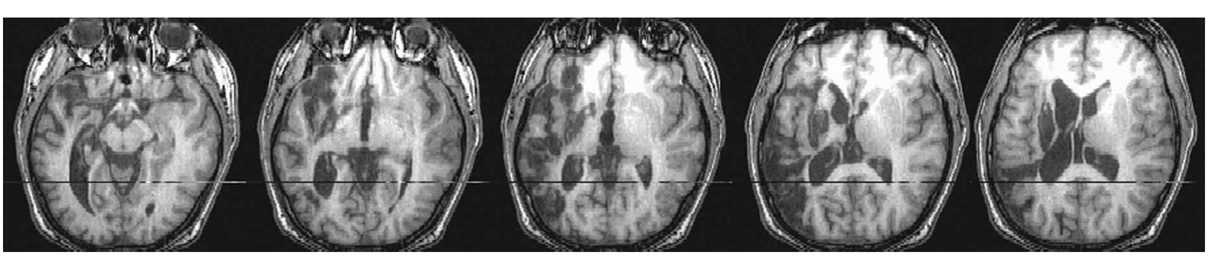

Figure 1. MRI scan results (axial slices) showing PH's left hemisphere lesions.

TABLE 1
Summary of PH's performance on basic neuropsychological tests

Test	*Maximum possible score*	*Obtained score*
Picture naming		
BORB (Riddoch & Humphreys, 1993)	76	9
Laiacona, Barbarotto, Capitani, & Trivelli (1993)	80	26
Word naming		
Regular words (PALPA 35)	30	3
Irregular words (PALPA 35)	30	6
Non words (PALPA 36)	15	0
Imageability and Reading (PALPA 31)		
HI HF	20	10
HI LF	20	6
LI HF	20	0
LI LF	20	1
Synonym matching (PALPA 49 and 50)		
Visual HI	30	27
Visual LI	30	20
Auditory HI	30	25
Auditory LI	30	24
Picture word matching (PALPA 47)	40	34
Pyramids and Palm Trees Test (Howard & Patterson, 1992) (mean 3 tests)	52	48
Executive functions		
Brixton Spatial Anticipation Test (Burgess & Shallice, 1997) (mean 2 tests)	54	31 (average)
WCST (Harris, 1988)	132	54 (no categories) (impaired)
Stroop (Trenerry, Crosson, DeBoe, & Leber, 1989)	–	20 errors on color name relative to word name trials (impaired)
Spelling (words taken from the PALPA, sets of HFHI, HFLI, LFHI, LFLI words, 15 words per set)		
Copy	60	60
Delayed copy (30 seconds delay)	60	44
Dictation	60	0

Thus, PH was not simply guessing, but interpreted the sentences using semantic plausibility (based on individual word meaning) rather than grammatical information. Note that overall, PH's score was higher than one of the previously reported cases (patient SA scored 23/60 and 28/60 on two presentations of PALPA test 55; Varley & Siegal, 2000). However, if we only take into account the reversible sentences (i.e., the sentences that most require taking into account the grammatical information), PH's score of 50% correct is very similar to the previously reported cases (no distinction between reversible and non-reversible sentences is reported for patient SA, Varley & Siegal, 2000; but patient MR scored 55% correct for reversible sentences, Varley et al., 2001).

In the written version, global score = 42/60 (70%). As in the oral version of the task, PH made more errors on reversible compared to non-reversible sentences (55% correct compared to 100% correct), and all but one error on the reversible sentences consisted of choosing the reverse action to the one described in the sentence.

Test 2. Comprehension of sentences with relative clauses, adverbial clauses and co-ordinated clauses

Some of the strongest claims that language has a necessary role for ToM have been made by de Villiers (e.g., de Villiers & de Villiers, 2002; de Villiers & Pyers, 2002), who argues that complement clauses embedded under a main clause (as in "John thinks that [embedded complement clause]," or "John said that [embedded complement clause]") provide a representational structure for thinking about mental states such as beliefs. However, Smith et al. (2003) found a correlation between 3- to 4-year-olds' performance on false belief tasks and their comprehension of relative clause sentences (e.g., John rode the horse that [embedded relative clause]), suggesting that belief reasoning may be related to children's abilities with a broader category of embedded grammatical constructions, not with embedded complement clauses per se. We were therefore interested in PH's ability to understand relative clause sentences.

In order to test comprehension of relative clause sentences (e.g., The woman pushed the man that opened the box) it is necessary to ensure that the participant cannot make correct judgments by using superficial strategies such as matching the order of appearance of the subject and object of the sentence to the order of occurrence of the events described (e.g., a woman pushing a man who then opens a box). To do this, it is possible to exploit one consequence of the relative clause's embedded syntax; that relative clause sentences (e.g., The woman pushed the man that opened the box) may describe an event in the main clause (woman pushing man) that actually occurred *after* the event in the relative clause (man opening box). A participant who has difficulty processing relative clause syntax may not recognize this possibility, and so may fail to see the correspondence between such a relative clause sentence and an event sequence in which the man first opens the box and is then pushed by the woman. In previous research such failure has been taken as diagnostic of difficulty processing embedding in relative clause sentences (e.g., Hamburger & Crain, 1982; Smith et al., 2003).

Method. For each trial PH viewed a short video clip (3 to 5 seconds) showing a simple event sequence, such as a male actor opening a box and a female actor pushing the male actor. This was followed by two still photographs depicting critical features of the video in the order in which they appeared in the video (in this case the man opening the box and the woman pushing the man). A sentence was then presented beneath the photographs and this was read by the experimenter (e.g., "The woman pushed the man who opened the box"). PH was asked to judge whether or not the sentence was consistent with the video. Each sentence appeared twice, once with a corresponding video (as in the foregoing example) and once with discrepant video (for the foregoing sentence the discrepant video showed the woman opening the box and then pushing the man). We considered that a sentence was comprehended if it was both correctly accepted and correctly rejected as descriptions of these videos. The task was presented on a computer using PowerPoint®, with the experimenter controlling the pace of presentation. Four sentence types were presented, with 14 sentences of each type, plus 16 filler trials, giving a total of 128 test trials that were presented over 8 separate blocks. An example of each sentence type is given below[2]:

1. Target relative clause sentences, e.g., "The woman pushed the man that opened the box."
2. Control relative clause sentences, e.g., "The woman pushed the man that was holding the box."
3. Co-ordinated clause sentences, e.g., "The woman pushed the man and the man opened the box."
4. Adverbial clause sentences, e.g., "Before the woman pushed the man, the man opened the box."

In Target relative clause sentences the event described in the relative clause could plausibly have occurred before the event described in the main clause. If PH failed to process the embedded syntax of relative clauses, he might incorrectly reject a sentence such as (1) as a valid description of an event sequence in which the man first opened a box and was later pushed by a woman. Of course, there were also other reasons why PH might incorrectly reject or accept these sentences as descriptions of video sequences. Comprehension of Target relative clause sentences also required assignment of the correct actions to the subject and object of the sentence, and processing of a sentence whose clause order was the reverse of the real order of the described events. PH's ability to meet these demands was assessed in the other three types of sentence.

In Control relative clause sentences the event in the relative clause (e.g., holding box) was continuous during the event in the main clause (e.g., woman pushing man). As a result, even if PH failed to process the embedded syntax of the sentence there would be no reason for him to reject it as a valid description of an event sequence in which a man holding a box was pushed by a woman. In order to have a continuous event described in the relative clause it was necessary to substitute many of the verbs and to alter the verb aspect from that in the Target

[2] Full details of the stimuli are available from the authors on request.

relative clause sentences. However, the Control relative clause sentences were matched to the Target relative clause sentences insofar as the relative clause was subject-extracting, modified the object of the main clause of the sentence, and subject and object assignment was reversible. As a result, the Control relative clause sentences served as a reasonable test of PH's ability to assign the correct actions to the subject and object of relative clause sentences that were relevantly similar to the Target relative clause sentences.

Co-ordinated clause sentences lacked the syntactic embedding of relative clause sentences, but like the Target relative clause sentences, described a discontinuous event in each clause. Thus we were able to use the same verbs with the same aspect as occurred in the Target relative clause sentences. This enabled us to check PH's ability to assign actions to the subject and object of two discontinuous events using the same verbs as for the Target relative clause sentences.

Like the Target relative clause sentences, adverbial clause sentences (based upon the adverb "before") required PH to comprehend a sentence whose clause order is the reverse of the real events described. However, adverbial clause sentences do not have the structural embedding of relative clause sentences, and the order of the events described by the separate clauses is determined by the semantic content of the adverb. Thus, if PH made errors on Target relative clause sentences but not on adverbial clause sentences, this would suggest that his difficulty was with processing the embedded syntax of relative clause sentences. However, if PH made errors both on Target relative clause sentences and on adverbial clause sentences this would suggest that his difficulty was with the more basic problem of interpreting sentences whose clause order does not match the order of events they describe.

Results. PH did not score significantly above chance for target relative clause sentences (7/14), or adverbial clause sentences (7/14), but was above chance for co-ordinated clause sentences (14/14) and control relative clause sentences (12/14).

Summary. On standard tests of grammatical ability taken from the PALPA (Kay, Lesser, & Coltheart, 1992), PH was significantly impaired. It is difficult to make equivalent comparisons between PH and the two patients reported previously because of differences in how these patients' performance was reported. Varley and Siegal (2000) reported the overall scores of SA on oral and written sentence–picture matching (PALPA subtests 55 and 56). PH was significantly impaired on these subtests but made fewer errors than SA. Varley, Siegal, and Want (2001) only reported the performance of patient MR on the reversible sentences from these PALPA subtests, which cannot be solved with simple semantic strategies. PH was significantly impaired on these items, and performed at a similar absolute level to MR. PH showed impaired comprehension of target relative clause sentences, but unimpaired comprehension of control relative clause sentences and co-ordinated clause sentences. This is the same pattern observed among 3- to 4-year-old children who fail standard false belief tasks (Smith et al., 2003). However, on the same task, PH showed impaired comprehension of adverbial clause sentences. This finding suggests that PH's difficulty is not with the embedded nature of relative clauses in particular, but with the more fundamental problem of using grammar (or grammar + semantics in the case of temporal adverbs) to constrain the appropriate interpretation of sentences whose clause order does not correspond directly to the order of the events described. These results were found in the context of a general impairment with function words, verbs and low imagery stimuli. We conclude that PH's grammar was sufficiently disrupted for us to predict disruption on reasoning tasks that are critically dependent upon grammatically structured linguistic representations. In particular, if belief reasoning requires processing of the grammar of embedded complement clauses, then belief reasoning should be impaired.

Nonverbal tests of ToM

PH was assessed on two nonverbal first-order false belief tasks and one nonverbal second-order false belief task. Although linguistic communication was involved in establishing the principles of the tasks with practice stimuli, the testing phase of each task could be conducted entirely without language.

Test 3. Reality-known false belief task

Method. This task was a video-based instantiation of the "unexpected transfer" false belief task, which has become a standard paradigm for testing ToM (e.g., Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). Practice trials established that PH's task was to indicate which of the two boxes a woman in the video would open first in order to find a green object. In false belief trials, a woman watched as a man placed the green object in one of the two boxes. The woman then left the room and, while she was outside, the man moved the green object from one box to the other in full view of the participant. When the woman returned, the video was paused and PH was asked to respond (by pointing to the computer monitor to indicate in which box the woman would look). Twelve false belief trials were mixed with three further trial types (12 items of each type). These included true belief trials and served to make it impossible to be consistently correct by adopting superficial strategies, such as always pointing to the box where the object was first placed. A full description of the task is provided in Samson, Apperly, Kathirgamanathan, and Humphreys (2005).

Results. PH scored 12/12 on false belief trials and 12/12 on all other trial types.

Test 4. Reality-unknown false belief task

Method. This task was a video-based instantiation of the nonverbal false belief task devised by Call and Tomasello (1999). Both the method and the data from PH were reported in Apperly, Samson, Chiavarino, and Humphreys (2004). Like the reality-known false belief task, PH viewed short video sequences involving a man and a woman, only this time practice trials established that PH's task was to use a clue given by the woman to locate a green object hidden in one of two boxes. On false belief trials the woman has a false belief, which must be taken into account when using her clue to locate the object. The crucial difference from the reality-known false belief task is that, at the point when the participant must infer that the woman has a false belief, he or she does not know the object's true location, and so this knowledge cannot interfere with their judgment about the woman's perspective (see Figure 2a). False belief trials were mixed with four further trials (12 items of each type), which served to guard against the adoption of superficial strategies to solve the false belief trials.

Results. PH scored 11/12 on false belief trials (significantly above chance), and 12/12 on all other trial types.

Test 5. Second-order ToM task: False belief about false belief

Second-order false belief tasks require the participant to reason about what one person thinks another person is thinking (Perner & Wimmer, 1985). Such tasks are significantly harder than first-order false belief tasks for children (who do not typically pass until 6 or 7 years of age), for patients with brain damage (e.g., Fine, Lumsden, & Blair, 2001) and for individuals with autism or schizophrenia (e.g., Baron-Cohen, 1989; Pickup & Frith, 2001). They typically require comprehension of linguistically complex stories and questions, making it particularly difficult to distinguish between errors that arise because of the complexity of the ToM reasoning, or the complexity of the associated language of the task. The current study was the first attempt to present a fully nonverbal second-order ToM task.

Conceptually, a second-order task was achieved by having PH predict the behavior of a participant in a version of the reality-unknown false belief task (see Figure 2a) where the participant would sometimes have a false belief about the beliefs of the clue-giver (see Figure 2b).

Method. Videos included 3 instead of 2 characters and used 4 instead of 2 containers (two white cylinders and two purple square boxes). Practice trials established the idea that that one of the women in the video (the pointer) would be asked to point to the box where the green object was located (the pointer was thus in the role of the participant in the reality-unknown false belief task). PH's task was to predict where the pointer would point. Practice trials also established that the other woman in the video (the helper) would place a red marker as a clue to help the pointer find the green object.

In the second-order (false belief about false belief) scenario, PH watched as a man showed the helper in which white cylindrical box the green object was located. Neither the pointer nor PH could see where the green object was. The helper then pretended to leave the room (neither the man nor the pointer could see that the helper was in fact still inside the room). The man then swapped the two white cylindrical boxes. The helper pretended to re-enter the room (by open-

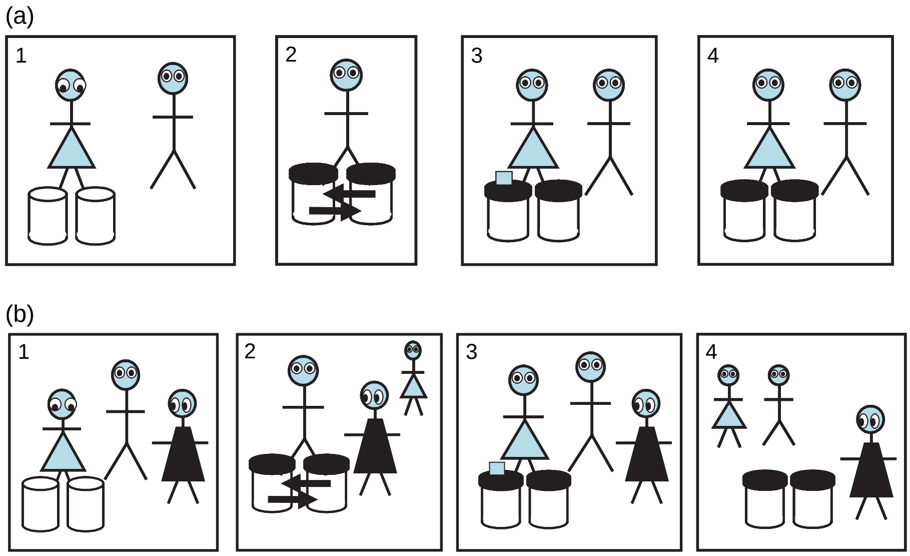

Figure 2. (a) Reality–unknown false belief task. Participants do not know which box contains a hidden object, their task is to work this out from a clue given by the woman. (1) The woman looks in the box. (2) In her absence, the man swaps the boxes. (3) The woman indicates where she thinks the object is. (4) The participant is asked to indicate where the object is located. (b) Second-order false belief task. Participants do not know which box contains a hidden object, their task is to work out where the pointer (woman in black) will think the object is (and so where she will point). (1) The helper (woman in blue) looks in the box. (2) She appears to leave, but in fact sees the man swap the boxes. (3) The helper returns to the scene and indicates where she thinks the object is. (4) The pointer believes falsely that the woman in blue has a false belief, so the participant should judge that she will point to the opposite box from the one indicated by the helper.

ing and closing the door) and pointed to one of the white cylindrical boxes. In order to correctly predict where the pointer would point, PH needed to realize that the pointer falsely believed that the helper had a false belief and that the pointer would thus point to the opposite white cylinder to the one indicated by the helper. Because PH could see that the helper only pretended to leave the room, he could infer that the helper had a true belief about the object's location. The second-order ToM question was whether PH could also work out that the pointer would (wrongly) believe that the helper had a false belief.

In the true belief (about false belief) scenario, the only difference was that the helper really left the room. So here, in order to predict correctly where the pointer would point, PH needed to realize that the pointer would correctly infer that the helper had a false belief and would therefore point to the opposite white cylinder box from the one indicated by the helper. Note that this time, both the participant and the pointer correctly thought that the helper had a false belief.

In an attempt to control for the memory demands of keeping in mind beliefs embedded within beliefs, we included memory control trials where PH had to keep in mind the effect of an initial physical transformation on the effect of a second physical transformation. In these memory control trials, the video started with the man showing to the helper where the green object was located. On half of the trials, the helper pretended to leave the room, and on the other half of the trials, the helper really left the room. While the helper was out of view, the man placed each of the white cylindrical boxes inside two purple square boxes. Half of the time the left cylinder went in the left box, and half the time it went in the right box. This was the first physical transformation. Next, the man swapped the square boxes. This was the second physical transformation. The helper then returned and pointed to the location on the table where the white cylinder containing the green object was located before it was hidden in the purple box. Correctly predicting where the pointer would point thus required PH to reconstruct where the white cylinder referred to by the helper at the end of the trial had ended up after the two location changes. (Note that whether or not the helper left the room was irrelevant because the helper pointed to where she originally saw the white cylinder containing the object. Our purpose in including this was to help match the belief trials for event complexity.)

Finally, the fourth type of trials (true belief about true belief filler trials) was designed to guard against the strategy of always pointing to the opposite location from the one indicated by the helper. For half of these trials, after the helper saw where the green object was located, she really left the room and for the other half of the trials, she only pretended to leave the room. While the helper was outside the room or pretended to be outside the room, the man lifted the white cylinder boxes (there was thus no change of location). PH needed to realize that the helper had a true belief about the green object location and that the pointer knew that the helper had a true belief. The correct prediction thus consisted of pointing to where the helper pointed.

In total there were 48 trials, 12 of each type. The items were split in two blocks. After each trial PH was given feedback in which he saw which location the pointer indicated. The real location of the green object was never shown, primarily to avoid PH confusing the goal of the task and pointing where he thought the green object was located instead of predicting where the pointer would point.

Results. PH scored 12/12 correct on the critical false belief about false belief second-order ToM trials. He also scored above chance for the other three types of trial: True belief about false belief, 11/12; true belief about true belief, 12/12; memory control trials, 10/12.

Summary. PH achieved near perfect scores on nonverbal tests of first-order and second-order belief reasoning. His performance on each individual task was significantly above chance, as was his performance on control trials designed to guard against the development of superficial strategies for success on false belief trials.

Combining Language and ToM

Test 6. Semantics of ToM

Moore et al. (1990) found that children's ability to pass false belief tasks was significantly related to their ability to evaluate the certainty implied by mental state terms. We considered this an interesting test of ToM semantics because success depends upon accessing the meaning of particular lexical items, but does not, in this task, depend upon integrating this information within the broader grammatical context of the sentence. We adapted the method of Moore et al. to make it suitable for administering a larger number of items to an adult participant.

Method. PH was shown two pictures of a man pointing to one of the two boxes (the man points to a different box on each picture) while expressing verbally a certain level of certitude as regards the object location. The man's face was masked and a speech bubble indicated what the man was saying while pointing to the box. Three levels of certitude were expressed with simple verbal utterances. The highest level of certitude was expressed with words denoting sure knowledge (e.g., "I know it's here"; "Certainly, it's here"). The middle level of certitude was expressed with words denoting that the character had some idea of where the object might be (e.g., "I think it's here"; "Probably it's here"). The lowest level of certitude was expressed by words denoting the absence of knowledge as to where the object is located (e.g., "I am guessing it's here"; "Maybe it's here"). Twelve utterances were selected for each level of certitude and the critical words were matched as closely as possible for frequency across the levels of certitude. Each utterance was presented twice, each time paired with an utterance from another level of certitude (for a total of 36 items). Correctly locating the object required PH to choose the box the man pointed at when expressing the highest level of certitude. Feedback as to the location of the object was given after each trial.

Results. PH scored 35/36.

Test 7. Story-based false belief tasks

Method. Apperly et al. (2004) reported PH's performance on 12 story-based false belief tasks. The tasks involved simple six-line stories, which were presented in written form and read aloud twice by the experimenter. Each story was followed by four questions, which were presented orally. One example story follows:

> Jeremy is eating out at a restaurant. Inside the restaurant, Jeremy hangs his coat on the stand by the door, and leaves his bag underneath. The waitress shows Jeremy to his table and tells him about today's special dishes. When she comes back, the waitress notices Jeremy's bag beneath the coat stand by the door. She decides that it is unsafe for the bag to stay there, as it would be

easy for someone to steal. Leaving the coat on the coat stand, she locks the bag in the store cupboard.

1. False belief question: Where does Jeremy think the bag is? On the chair or in the store cupboard?
2. Counterfactual question: What if the waitress had not noticed the bag? Where would the bag be? On the chair or in the store cupboard?
3. Memory control question: Where was the bag at the beginning? On the chair or in the store cupboard?
4. Reality control question: Where is the bag now? On the chair or in the store cupboard?

Results. PH scored above chance on the false belief questions (10/12 correct) and on the counterfactual, memory control and reality control questions (all 12/12 correct).

Test 8. Comprehension of ''thinks''+embedded complement syntax

de Villiers and Pyers (2002) found specific and potentially causal relationships between children's performance on false belief tasks and their comprehension of embedded complement syntax. Comprehension was assessed in a ''memory for complements'' task. Children were told short stories about a character who comes to have a false belief. They were then given an explicit summary about the false belief of the character, e.g., ''He thought he found a ring,'' and were then told what was really the case, ''But really, it was a bottle cap.'' The experimenter then pointed to a picture of the character and asked the crucial test question, ''What did he think?.'' Unlike false belief tasks, this memory for complements task does not require the participant to make any kind of inference about beliefs, but merely to process information about a false belief presented in its conventional grammatical context. Nonetheless, young children often responded incorrectly by reporting reality, rather than the character's false belief. Our task was modeled on de Villiers' paradigm, but with a larger number of trials and trial types.

Method. A sentence expressing the color of balls located in either one or two boxes was shown to PH and read aloud for him. On the ''belief'' and ''reality'' trials the one part of the sentence expressed a man's belief about the color of the ball in a single box and the other part of the sentence expressed the real color of the ball (e.g., ''He thinks there is a red ball in the box but really there is a blue ball in the box''). Half of the sentences started with the ''think'' phrase and half started with the ''really'' phrase. On other trials, information was presented about balls in different places (e.g., ''There is a yellow ball in the box on the table but there is a green ball in the box on the chair'') or different times (e.g., ''Yesterday there was a yellow ball in the box but today there is a green ball in the box''). After PH had read the sentence, two pictures appeared on the screen, and PH was asked to judge which of the two pictures matched the meaning of the sentence. The pictures only represented one part of the sentence. For belief trials (12 in total), the picture would show the content of the box within a thought bubble (e.g., for the sentence, ''He thinks there is a red ball in the box but really there is a blue ball in the box,'' one picture would show a thought bubble with a red ball inside the box and the other picture would show a thought bubble with a blue ball in the box). For the reality trials (12 in total), each picture would show a possible content for the box on the table. For the table, chair, yesterday and today trials (6 of each), each picture would show the possible content of the box on the chair, table, yesterday or today. Yesterday and today were distinguished by tagging the box in the picture ''past'' or ''present'' (see Table 2). Before starting the test, we ensured that the PH understood the symbols used such as the thought bubbles and the word tags.

Results. In this task, PH did not score above chance for belief trials (3/12) but was above chance for reality trials (11/12). He was not above chance for table/chair (8/12) or yesterday/today trials (9/12).

Summary. PH was unimpaired on a test of ToM semantics and he was above chance on a set of first-order story-based false belief tasks that did not entail comprehending embedded complement clause sentences. However, PH showed impaired comprehension of simple belief reports that required him to process the verb ''thinks'' together with an embedded complement clause.

TABLE 2
Illustration of sentence and picture stimuli for each trial type of Test 8: Comprehension of "thinks" + embedded complement syntax

Trial type	*Sentence*	*Forced choice between two pictures*
Belief	He thinks there is a red ball in the box but really there is a blue ball in the box	Picture 1: Thought bubble showing red ball in box
		Picture 2: Thought bubble showing blue ball in box
Reality	He thinks there is a red ball in the box but really there is a blue ball in the box	Picture 1: Red ball in box on table
		Picture 2: Blue ball in box on table
Location (table or chair)	There is a red ball in the box on the table but there is a blue ball in the box on the chair	Picture 1: Red ball in box on table
		Picture 2: Blue ball in box on table
Time (yesterday or today)	Yesterday there was a red ball in the box but today there is a blue ball in the box	Picture 1: "PAST" Red ball in box on table
		Picture 2: "PRESENT" Red ball in box on table

GENERAL DISCUSSION

Across a range of tasks PH showed significant grammatical impairment. On standard measures he has clear difficulty in processing sentences dependent on grammatical information (e.g., with reversible sentences), and he has problems in processing function words, abstract words and verbs. PH's performance on language tasks is comparable to that of the two patients reported by Varley and colleagues (Varley & Siegal, 2000; Varley et al., 2001). Importantly, we were able to investigate PH's ability to comprehend embedded complement clauses and relative clauses, which have been shown to have a particularly close relationship with ToM in studies of children (e.g., de Villiers & Pyers, 2002; Smith et al., 2003). PH was unable to judge above chance for these grammatical constructions, but was able to succeed on some other trial types within the same tasks, suggesting that PH's difficulties were with the grammar of these constructions and not with the tasks in general.

Despite substantial and general grammatical impairment that included the constructions most potentially relevant to ToM, PH performed significantly above chance (indeed almost without error) on two nonverbal tasks that required first-order reasoning about false beliefs, and a nonverbal task that required second-order reasoning about one character's false belief about another character's false belief. It is noteworthy that these false belief tasks, on which PH succeeded, were substantially more complicated than the language assessments on which PH failed, which included simple word–picture and sentence–picture matching. The events of the first-order false belief tasks were played out over 40–50 seconds and required PH to monitor the movements of the boxes in which the object was hidden and the movements of the woman and to co-ordinate this information in order to infer whether the woman had a true or false belief. Second-order false belief tasks additionally required the monitoring of the true or false beliefs of a second character, and the co-ordination of these with the true or false beliefs of the first character. In comparison, tests of PH's comprehension of embedded complement clauses and relative clauses required the evaluation of just one or two sentences against a short video (3 to 5 s) and/or still photographs.

The tasks that combined ToM requirements with language requirements provided stark evidence of PH's strengths and weaknesses by showing that PH succeeded on a test of ToM semantics, and on simple story-based false belief tasks, and only experienced difficulty when grammatical structure was critical for understanding ToM information. In relation to ToM semantics, PH performed well on a test that required him to evaluate the certainty implied by a range of mental state terms. Clearly, this is far from being an exhaustive test of "ToM semantics," but given that children's ability to make such judgments correlates with their ability to reason about false beliefs (Moore et al., 1990), it may be significant that this ability is spared in PH. It is unknown whether such abilities were spared or impaired in the two patients studied by Varley et al. (Varley & Siegal, 2000; Varley et al., 2001).

PH's success on the story-based false belief tasks, including verbal false belief and counterfactual questions, may seem surprising. However, it is important to recognize that PH has relatively unimpaired access to lexical semantics, at least for concrete concepts (see Table 1), and he is highly skilled at constructing plausible meaning from this information and from the context in which it appears. The false belief stories are grammatically quite simple and, in particular, the order in which events occur is the same as the order in which they are described. Thus, the weaknesses that PH demonstrated in using grammar to comprehend reversible sentences on the PALPA or the adverbial or relative clause sentences, might not have prevented him from constructing the gist of the stories from comprehension of specific words combined with their order and context. In addition, the stories did not describe the false beliefs of the story character, so did not require PH to gain any information from the comprehension of embedded complement syntax. The false belief question did involve direct reference to the thoughts of the character, but the embedded complement clause served only to identify the object of the character's thought, and not the propositional content of their false belief. Thus, the meaning of the false belief question was not critically dependent upon its grammar. Nonetheless, with his substantial grammatical impairments, PH's success on these story-based false belief tasks is a testament to his skill at making plausible interpretations from the language processing ability that he still possesses.

In the light of this it is all the more striking that PH performed so poorly on a test of his comprehension of embedded complement syntax that involved direct description of a character's false belief and the corresponding reality, e.g., "Really there's a blue ball in the box, but he thinks that there's a red ball in the box." It is important to recognize that PH's errors were not the result of a simple confusion between information about belief and reality. When asked to select a pictorial representation of reality he was correct on eleven out of twelve trials. However, when asked to select a pictorial representation of the character's belief he answered correctly on only three out of twelve trials. On the remaining nine trials he picked the picture with the colored ball corresponding to reality. This asymmetry might be explained if PH was able to comprehend the report of reality, but his disrupted syntax meant that the embedded complement clause was not doing its job of establishing a clear separation between reality and the content of the character's belief.

Conclusions about language and ToM

The current study strengthens the evidence that, for adults at least, explicit reasoning about beliefs does not depend upon access to grammatical structure. In particular, belief reasoning seems to be possible in the absence of the specific grammatical construction—embedded complement clauses—that is used for representing beliefs in language. Nonetheless, it is important to recognize that, to the extent that the meaning of linguistically reported mental states depends critically upon grammar, grammatical impairments will severely restrict the availability of much everyday input for ToM reasoning. Importantly, the existing data leave open the possibility that access to lexical semantics for ToM concepts does have a necessary role in ToM reasoning. To address this question it will be necessary to examine verbal and nonverbal ToM abilities in patients with disproportionate semantic impairment, as in the case of semantic dementia.

If reasoning about beliefs does not require grammatical structure, this leaves an important question about what representations and processes are in fact necessary. Independence from language has been seen by some authors as evidence that belief reasoning depends upon representations and processes that are specific to the domain of ToM, and perhaps constitute a cognitive module (e.g., Carruthers, 2002). The case of PH does not provide evidence to support this view because PH shows spared reasoning abilities beyond the theory of mind domain. PH passed memory control trials of the second-order false belief task, which required him to reason about the consequences of one physical transformation on another. He also passed the counterfactual reasoning trials of the story-based false belief tasks. Finally, PH passes a nonverbal false photograph task that is closely matched in its general reasoning demands to the reality-unknown false belief task (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, in press). Of course, this is not evidence against the existence of domain-specific or modular ToM processes, but a parsimonious explanation for PH's successful reasoning about beliefs, non-mental misrepresentations, counterfactuals and physical transformations need only appeal to domain-general processes that are independent of grammar.

Finally, if belief reasoning in adults is not dependent upon access to grammatical structure, then what are we to make of the strong, consistent and possibly causal relationships that have been found between grammar and ToM in studies of children (e.g., Astington & Jenkins, 1999; de Villiers & Pyers, 2002; Ruffman et al., 2003; Smith et al., 2003)? Our view is that the accumulating data from adults helps to resolve the inherent ambiguity in studies of children. Children's success on false belief tasks could have been dependent upon their grammatical ability because grammatical structure was a constitutive component of successful belief reasoning. Existing studies of adults suggest that this is not the case. This leaves the interesting possibility that grammar plays a critical role in the *development* of successful belief reasoning, but, in doing this, it helps to form a capacity for belief reasoning that does not depend upon grammar as a source of structure or expression.

REFERENCES

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and left temporo-parietal contributions to theory of mind: Neuropsychological evidence from a false belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*, 1773–1784.

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W., & Humphreys, G. W. (in press). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on nonverbal, "reality-unknown" false belief and false photograph tasks. *Cognition*.

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

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

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, *30*, 285–298.

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests – manual*. Bury St Edmunds, UK: Thames Valley Test Company.

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

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*, 657–674.

Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition*, *72*, 237–267.

de Villiers, J. G., & de Villiers, P. A. (1999). Linguistic determinism and the understanding of false beliefs. In P. Mitchell (Ed.), *Children's reasoning and the mind*. Hove, UK: Psychology press.

de Villiers, J. G., & de Villiers, P. A. (2002). Why not LF for false belief reasoning? *Behavioral and Brain Sciences*, *25*, 682–683.

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

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. Hove, UK: Psychology Press.

Dunn, J., & Brophy, M. (2005). Communication, relationships and individual differences in children's understanding of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind*. Oxford, UK: Oxford University Press.

Emery, N. J. (2005). The evolution of social cognition. In A. Easton & N. J. Emery (Eds.), *The cognitive neuroscience of social behaviour*. Hove, UK: Psychology Press.

Fine, C., Lumsden, J., & Blair, R. J. R. (2001). Dissociation between "theory of mind" and executive functions in a patient with early left amygdala. *Brain*, *124*, 287–298.

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

Harris, M. E. (1988). *Wisconsin Card Sorting Test: Scoring program (Version 2.0)*. Odessa, FL: Psychological Assessment Resources.

Harris, P. L. (2005). Conversation, pretence and theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind*. Oxford, UK: Oxford University Press.

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test*. Bury St Edmunds, UK: Thames Valley Test Company.

Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic Assessment of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum Associates, Ltd.

Kumar, V., & Humphreys, G. W. (2006). *Semantic and non-semantic representations involved in spelling rehabilitation: Evidence from deep dysgraphia*. Manuscript submitted for publication.

Laiacona, M., Barbarotto, R., Capitani, C., & Trivelli, E. (1993). Dissociazioni semantische intercategoriali: Descrizione di una batteria standardizzata e data normativi. Archivio di Psicologia. *Neurologia e Psichiatria*, *2*, 209–248.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, *61*, 722–730.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.

Perner, J., & Wimmer, H. (1985). John thinks that Mary thinks that": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471.

Pickup, G. J., & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: Symptomatology, severity and specificity. *Psychological Medicine*, *31*(2), 207–220.

Riddoch, M. J., & Humphreys, G. W. (1993). *Birmingham Object Recognition Battery (BORB)*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, *18*, 139–158.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain*, *128*, 1102–1111.

Smith, M., Apperly, I. A., & White, V. (2003). False belief reasoning and the acquisition of relative clause sentences. *Child Development*, *74*, 1709–1719.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test*. Grand Rapids, OK: Psychological Assessment Resources, Inc.

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and ''theory of mind'' in an agrammatic aphasic patient. *Current Biology*, *10*, 723–726.

Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*, 489–493.

Violon, A., & Seyll, S. (1984). *Test d'apprentissage progressif de dessins sans signification*. Braine-le-Château, France: Application des techniques modernes.

Warrington, E. K. (1984). *The Recognition Memory Test*. Windsor, UK: NFER Nelson.

Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655–684.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition: International Journal of Cognitive Science*, *13*, 103–128.