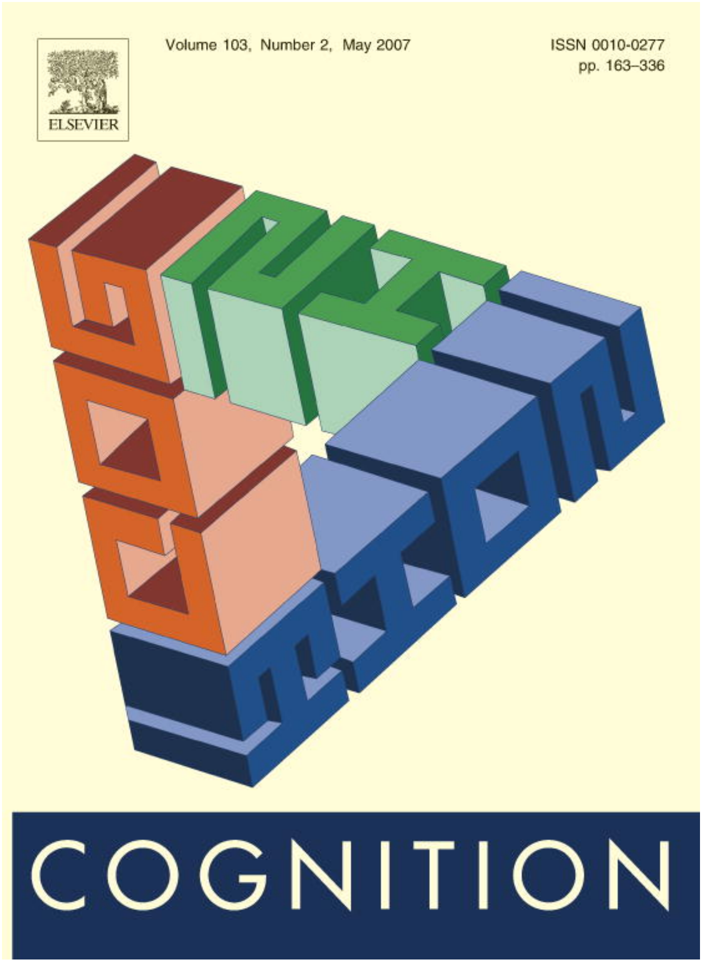
Volume 103, Number 2, May 2007
ISSN 0010-0277
pp. 163–336
ELSEVIER
COGNITION

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Cognition 103 (2007) 300–321
COGNITION
www.elsevier.com/locate/COGNIT

Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "reality-unknown" false belief and false photograph tasks ☆

Ian A. Apperly *, Dana Samson, Claudia Chiavarino, Wai-Ling Bickerton, Glyn W. Humphreys

School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 20 October 2004; revised 22 August 2005; accepted 28 April 2006

Abstract

To test the domain-specificity of "theory of mind" abilities we compared the performance of a case-series of 11 brain-lesioned patients on a recently developed test of false belief reasoning (Apperly, Samson, Chiavarino, & Humphreys, 2004) and on a matched false photograph task, which did not require belief reasoning and which addressed problems with existing false photograph methods. A strikingly similar pattern of performance was shown across the false belief and false photograph tests. Patients who were selectively impaired on false belief tasks were also impaired on false photograph tasks; patients spared on false belief tasks also showed preserved performance with false photographs. In some cases the impairment on false belief and false photograph tasks coincided with good performance on control tasks matched for executive demands. We discuss whether the patients have a domain-specific deficit in reasoning about representations common to both false belief and false photograph tasks.

☆ This manuscript was accepted under the editorship of Jacques Mehler.
* Corresponding author. Tel.: +44 121 414 3339.
E-mail address: i.a.apperly@bham.ac.uk (I.A. Apperly).

0010-0277/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2006.04.012

Keywords: False belief; False photograph; Theory of mind; Domain specificity; Metarepresentation; Social cognitive neuroscience; Temporo-parietal junction

1. Introduction

The ability to explain and predict the behaviour of agents in terms of folk-psychological concepts such as belief, desire and knowledge is widely regarded as central to our uniquely human capacities for social interaction and communication (e.g., Baron-Cohen, Tager-Flusberg, & Cohen, 2001; Malle, Moses, & Baldwin, 2001; Sperber, 2000). It is widely accepted that young children lack some of these "theory of mind" abilities (e.g., Astington, Harris, & Olson, 1998; Flavell, 1999; Lewis & Mitchell, 1994; Mitchell & Riggs, 2001). Specific difficulty on theory of mind tasks is also held to be distinctive of a number of psychological disorders such as autism (e.g., Baron-Cohen et al., 2001) and schizophrenia (Frith & Corcoran, 1996; Langdon & Coltheart, 1999; Lee, Farrow, Spence, & Woodruff, 2004) and to be a possible result of focal damage to a variety of brain structures (e.g., Happe, Malhi, & Checkley, 2001; Samson, Apperly, Chiavarino, & Humphreys, 2004; Stone, Baron-Cohen, & Knight, 1998). A number of authors have argued that theory of mind depends upon specialised functional and neuro-anatomical mechanisms (e.g., Frith & Frith, 2003; Leslie & Thaiss, 1992; Saxe, Carey, & Kanwisher, 2004). However, this conclusion remains highly controversial because of disagreement about what should count as evidence of specific difficulty with theory of mind (e.g., Apperly, Samson, & Humphreys, 2005; Perner, 1995; Russell, Saltmarsh, & Hill, 1999; Saxe et al., 2004). In the current paper, we focus on reasoning about false beliefs as one particular kind of theory of mind task. Some of the strongest evidence in favour of strong domain-specificity for theory of mind comes from functional dissociations between performance on false belief tasks and tasks that require reasoning about non-mental representations such as photographs (Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992), and from patterns of neural activation associated with reasoning about beliefs and not photographs (Sabbagh & Taylor, 2000; Saxe & Wexler, 2005; Saxe & Kanwisher, 2003). By examining for the first time the effects of brain damage on participants' ability to reason about non-mental representations we aim to advance understanding of the cognitive processes responsible for theory of mind. We describe a novel false photograph task, designed to solve a number of problems with the "false" photograph methods used in these existing studies. We then use our new task to test the domain-specificity of the belief reasoning problems of three patients with lesions to the left temporo-parietal-junction (TPJ), who show evidence of a relatively specific belief reasoning deficit (Samson et al., 2004), compared with eight further patients who show less specific patterns of success or failure (Apperly et al., 2004).

In a common form of false belief task (Baron-Cohen, Leslie, & Frith, 1985; Flavell, 1999; Wimmer & Perner, 1983), one character, Sally, puts her marble in a basket. She then leaves the scene, and in her absence, Anne moves the marble from the basket to the box. The key question for the participant is where Sally thinks the marble is, or alternatively where she will look for her marble when she returns. Like adults, typically developing 4-year-olds most commonly judge correctly that Sally will look in the basket. These correct judgements are widely regarded as good evidence that the participant considers Sally's behaviour to be determined by her (false) belief. In contrast, 3-year-olds commonly err by judging that Sally will look in the marble's new location. This response pattern is also common in children with autism (Baron-Cohen et al., 2001) and has been reported in adults with certain forms of mental illness or brain damage (Frith & Corcoran, 1996; Happe et al., 2001; Samson et al., 2004; Stone et al., 1998). However, the interpretation of incorrect responses is controversial for at least two reasons. First, although the false belief task is designed to test participants' ability to reason about beliefs, it clearly also makes demands on other cognitive processes such as language, working memory and inhibitory control. Second, these other cognitive processes are often immature in typically developing 3-year-olds, or impaired in children with autism, adults with schizophrenia and adults with brain damage (e.g., Carlson & Moses, 2001; Heinrichs & Zakzanis, 1998; Stuss & Benson, 1986). Therefore, errors on false belief tasks could either reflect a relatively specific problem with belief reasoning (a domain-specific effect) or difficulties with any one of these other more general cognitive processes. To test whether belief reasoning makes use of domain-specific processes these possible sources of error need to be separated.

1.1. The "false" photograph procedure

One way to control for the large incidental processing demands of false belief tasks is to devise structurally similar comparison tasks that do not require a false belief to be inferred. One such task is Zaitchick's (1990) "false" photograph procedure, which can be made to follow a very similar event sequence to the false belief task just described. For example, Sally puts her marble in the basket. A Polariod camera is used to photograph the scene with the basket containing the marble and the empty box. While the photograph is developing Anne moves the marble from the basket to the box. The key test question concerns the location of the marble in the photograph, and is designed to be directly analogous with asking for Sally's false belief about the location of the marble in the false belief task. Zaitchick's (1990) original findings suggested that typically developing children performed no better on the "false" photo task than on the false belief tasks; if anything their performance was somewhat worse, showing no evidence that children's problems with belief reasoning were domain-specific. In marked contrast, a number of studies have shown that children with autism perform significantly better on "false" photograph or equivalent "false" drawing tasks than on false belief tasks (Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992). Together, these findings suggest that children with autism may have a selective, domain-specific problem with reasoning about beliefs (Charman & Baron-Cohen, 1992; Leslie & Thaiss, 1992). Moreover, on the assumption that these "false" photograph tasks are indeed a closely matched control for false belief tasks, "false" photo versus false belief subtractions have been employed in three recent ERP and fMRI studies designed to identify the brain regions specifically activated by belief reasoning (Sabbagh & Taylor, 2000; Saxe & Wexler, 2005; Saxe & Kanwisher, 2003). These studies suggest that belief reasoning activates distinct areas of the brain in a highly selective way. These findings, and those from children with autism, have been used to argue for the modularity or domain-specificity of belief reasoning (e.g., Frith & Frith, 2003; Saxe et al., 2004).

However, the literature offers two important reasons for a cautious interpretation of the findings from these comparisons of false belief and "false" photograph tasks. The first limitation is that both false belief and "false" photograph methods confound the need for the participant to infer the content of a representation (the marble is in the basket), with the need to resist interference from the participant's own conflicting knowledge about the true state of affairs (the marble is in the box). There are theoretical and empirical reasons for thinking that these processes might have a separate functional and anatomical basis (Frith & Frith, 2003; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005). However, with existing methods it is unclear whether dissociations between false belief and "false" photograph methods would reflect domain-specificity in the process of inferring the content of a belief, the process of resisting interference from knowledge of reality, or both. This limitation needs to be addressed by further work using methods that enable the contribution of these component processes to be studied independently. This is one objective of the current paper.

The second, and more serious, reason for caution comes from a series of studies that suggest both conceptual and empirical reasons for doubting the similarity of "false" photograph and false belief tasks. Conceptually, Perner (1995) argues that the comparison is fundamentally misleading because, whereas false beliefs are genuinely false representations of the current situation, "false" photographs are actually true representations of an outdated situation. Thus, a proper understanding of false belief requires the participant to process a conflicting relationship between a current (mis)representation and the current situation that it misrepresents. In contrast, understanding a "false" photograph would merely require the participant to recall the past situation that was photographed; the conflicting current situation need not be considered. If this analysis is accepted, it undermines the utility of the "false" photograph task, where the photograph does not misrepresent and does not conflict with the current situation, as a closely matched control for the conceptual and processing demands of the false belief task, where the belief does misrepresent and does conflict with the current situation. Comparison of performance on false belief and "false" photograph tasks would not be a good test of the domain specificity of belief reasoning.

Russell et al. (1999) offer empirical reasons for thinking that the standard "false" photograph task makes lower demands on executive control than the false belief task. The false belief task requires the participant to evaluate the content of an intangible (and thus non-salient) belief state while resisting interference from their own knowledge of the true (visible and more salient) physical state of affairs. In contrast, the "false" photograph task requires evaluation of a two-dimensional physical state (more salient than a belief) while resisting interference from the changed three-dimensional physical state. If the photograph were more salient than the belief, then resisting interference from the current situation would be less demanding on executive control for the "false" photograph task than for the false belief task. Russell et al. (1999) argue that this difference could account for why children with autism perform better on false photograph tasks than false belief tasks, since children with autism are known to have executive function problems. Russell et al. (1999) presented children with autism with a modified "false" photograph condition where the relative salience of the photographed scene and the changed three-dimensional scene was designed to be more comparable to the false belief condition. When the photograph was taken there was no object in the target scene (so the content of the photograph had lower salience) but an object was added to the scene before the test questions were asked (so the physical situation had higher salience). On this modified "false" photograph task, children with autism performed no better than on the false belief task, whereas they performed significantly better on a standard "false" photograph task. These findings are consistent with Russell et al.'s (1999) contention that the standard "false" photograph task is a poor control for the executive demands of the false belief task.

Whereas the critiques just described suggest reasons why the false photograph might be easier than the false belief task, there are also reasons for thinking that the standard form of the "false" photograph task might be anomalously hard because the test question is easily misunderstood by children. Perner (1995) and Russell et al. (1999) point out that a question such as "In this photograph, where is the marble" could be misunderstood by children as "Where is the marble (that you could see in this photograph)". Slaughter (1998) tried to make sure that children understood the question to be about the item in the photograph by pointing to the face-down photograph when the question was asked. On this variation of the task, Slaughter (1998) found that 3- to 4-year-olds performed significantly better than on false belief tasks. It is difficult to predict the effect of this change to the task instructions on the performance of children with developmental disorders such as autism. However, the possibility of an artificially inflated error rate in previous forms of "false" photograph task, and the potential for the severity of this effect to vary across typically and atypically developing children should make us cautious about previous studies that have compared "false" photograph and false belief task performance in typically developing children and children with autism (e.g., Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992).

Finally, if it is true that false belief and "false" photograph tasks test a common underlying concept of representation (e.g., Perner, 1991; Zaitchick, 1990) then performance on the two types of task should be correlated, and training that improved performance on one type of task might generalise to the other type of task. In fact Slaughter (1998, Study 1) found no correlation between performance on false belief and false photograph tasks, and successful training on one type of task did not generalise to the other type of task (Study 2).

In sum, the available evidence casts serious doubt on the suitability of existing "false" photograph procedures as a closely matched comparison for false belief tasks. They may not make the same representational demands (e.g., Perner, 1995), or executive demands (Russell et al., 1999), some "false" photograph test questions may be misunderstood by children (Slaughter, 1998) and, despite the surface similarity of the photograph and belief tasks, there is no evidence of consistency in children's responses on the two types of task. This suggests that differences between performance (or neural activation in imaging studies) on existing "false" photograph and false belief tasks do not provide reliable evidence of domain-specific belief reasoning processes. These arguments notwithstanding, existing comparisons of false belief and "false" photograph tasks cannot distinguish domain-specificity in belief reasoning from domain-specificity in the process of resisting interference from knowledge of reality. In what follows, we describe a recently devised false belief task that eliminates key processing requirements of more typical methods (Apperly et al., 2004; Call and Tomasello, 1999). We then consider how a photograph task matched to this new false belief method answers the problems just described.

1.2. A revised false belief task

Apperly et al.'s (2004) objective was to improve on existing methods for assessing the belief reasoning of brain damaged adults. The commonly used story-based false belief tasks make high incidental demands on language and executive function. Only patients who can meet these incidental demands can be tested reliably, with the result that it is impossible to work with many potentially interesting patients, including most of those in the current study (Apperly et al., 2004). Based on Call and Tomasello's (1999) non-verbal false belief task, Apperly et al. (2004) created short videos designed to be suitable for adult participants. The participant's task in both test and control trials was to work out which of two identical boxes contained a hidden object. On each trial, a female character saw inside the two boxes, and offered a clue to the participant by placing a marker on top of one of the boxes. The participant did not see inside the boxes. False belief trials consisted of four key stages. (1) The woman's belief was fixed when she saw inside the boxes. (2) The woman's belief became false when a male character switched the locations of the two boxes while the woman was out of the room. (3) The content of the woman's belief was made manifest when she returned to the room and gave her clue by mistakenly placing the marker on the incorrect location. (4) The participant could then identify the object's true location, provided they took the falseness of the woman's belief into account (see Fig. 1).

This method has a number of advantages over story-based tasks. Language is completely eliminated from the test trials, making it possible to test participants with significant language impairment. Control trials, where finding the location of the object does not require the participant to take into account the woman's belief, allowed us to check that participants could follow the basic procedure and meet critical performance demands such as remembering that the boxes have swapped and pointing to a different box from the one indicated by the woman (see Apperly et al., 2004). Most useful of all in the current context is the fact that the true location of the object is unknown to the participant when they must work out that the woman has a false belief. In fact, working out that the woman has a false belief is a necessary step to locating the object. This difference from more standard false belief tasks means that the need to infer the false belief is de-confounded from the need to resist

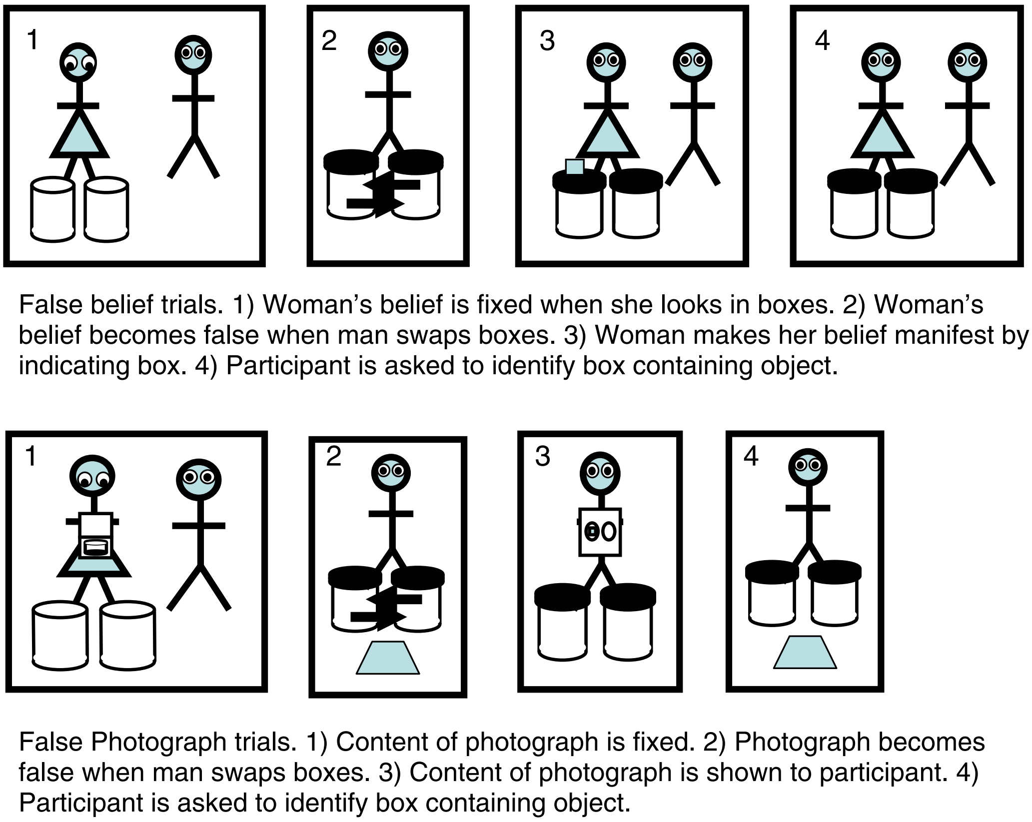

Fig. 1. Schematic diagram of participants' view of the event sequence for the False Belief task (upper panel) and the False Photograph task (lower panel).

interference from knowledge of the correct answer, and also helps address a number of problems with creating a photograph comparison task.

1.3. A revised false photograph task

We created a photograph analogue of Apperly et al.'s (2004) non-verbal false belief task (see Fig. 1). The participant's task was to identify the location of an object hidden in one of two identical boxes. On each trial, a female character took a polariod photograph showing the interior of the two open boxes, which was subsequently shown to the participant as a clue to the object's location. False photograph trials were closely matched to the four key stages of the false belief trials. (1) The content of the photograph was fixed when the woman took a photograph of the interior of the boxes. (2) The photograph became false when the male character switched the locations of the two boxes. (3) The content of the photograph was shown by turning it over to face the participant. (4) The participant could then identify the object's true location, provided they took the falseness of the photograph into account (see Fig. 1).

The first advantage of this new false photograph procedure is that the task is entirely non-verbal. Thus, unlike existing "false" photograph methods, there is no possibility that misinterpretation of the language of the test question could be an unwanted source of errors. The absence of language in the procedure is also a particular advantage for work with participants such as brain-damaged patients whose language problems may prevent them from undertaking existing “false” photograph tasks.

Second, Russell et al. (1999) raised the concern that the “false” photograph method employed in earlier studies was a poor match for the substantial executive demands of standard false belief tasks. In particular, these authors argued that the need to resist interference from knowledge of reality was not equivalent in “false” photograph and false belief tasks. Our procedure answers this particular concern directly: In both false belief and false photograph versions of the task, knowledge of the real location of the object *is unknown to the participant*. This eliminates any possible differences between the interfering effects of such knowledge in false belief and false photograph conditions. It also reduces the possibility that errors on either task are due to a “reality bias”, a “curse of knowledge” or “epistemic egocentrism” (Birch & Bloom, 2004; Mitchell & Lacohee, 1991; Royzman, Cassidy, & Baron, 2003; Saltmarsh, Mitchell, & Robinson, 1995).

Third, Perner (1995) argued that “false” photograph tasks do not make the same conceptual demands as false belief tasks. A false belief that the marble is in the basket is a (mis)representation of a current situation where the marble is in the box. In contrast, it is quite misleading to call the outdated photograph “false” since the photograph is not a (mis)representation of the current situation, but rather a true representation of an outdated situation. Perner’s (1995) main concern is with this claim that the false belief and “false” photograph tasks are not *conceptually* equivalent. However, his analysis also implies that the *processing* of the two tasks will be quite different. On Perner’s (1991, 1995) account, false beliefs are “about” the true situation. Thus, mental representation of a false belief entails representing the belief *in relation to* the conflicting reality that it (mis)represents. In contrast, “false” photographs are not “about” the true situation. Mental representation of a photograph of an outdated situation does not entail representation of the (conflicting) current situation. Because the standard “false” photograph task would not require the representation of conflicting information, Perner’s (1995) account suggests that it would make lower executive demands than the false belief task. Our false photograph task answers Perner’s (1995) conceptual objection. Recall that the photograph is taken with the object in its original location, the boxes are swapped and then the photograph is revealed *to help the participant find the object*. The photograph is therefore a clue “about” the new situation, and so meets Perner’s (1991, 1995) criteria for a misrepresentation. In meeting Perner’s (1995) conceptual objection our method also eliminates the related concern that false photographs do not need to be processed in the same way as false beliefs. To infer the true location of the object the false photograph must be considered *in relation to* the current situation. Exactly, the same process is necessary in the false belief condition. Moreover, once the true location of the object has been inferred, it will be in conflict with the information from the clue in the same way in both photograph and belief conditions.

A fourth advantage of our new false photograph procedure is that the remaining incidental processing demands are closely matched to those of the false belief task. Both false belief and false photograph tasks require the participant to track the switch of location of the two boxes. Both tasks also require the participant to infer the true location of the object, and to point to the opposite box from the one recently identified (either in the photograph or indicated by the woman). As is clear from Fig. 1, both false belief and false photograph trials present the participant with a similarly neutral visual array for responding: in the false belief task the woman's marker is in a neutral position between the boxes and in the false photograph task the photograph has been placed face-down in front of the boxes. Thus, the visual array provides no direct information that could interfere with inferring the object's true location or with pointing to the correct box. It is of course possible that the process of inferring the true location of the object suffers different levels of interference from the participant's memory of the woman's false belief and her placement of the marker, compared with the interference from their memory of the false photograph. However, we believe that any such differences are likely to be small in comparison with the differences that we have eliminated with our modified procedures. We return to this issue in the Discussion.

1.4. The current study

In the current study, we used our new false photograph task to test the domain-specificity of the belief reasoning deficits of a series of brain damaged patients previously reported by Apperly et al. (2004; Samson et al., 2004). Using the non-verbal false belief task described above, these studies provided two sources of evidence on the domain-specificity of patients' errors. First, Apperly et al. (2004) showed that there was no reliable relationship between performance on false belief trials and performance on independent tests of inhibitory control, working memory or language. All seven patients who failed the false belief task (by failing to be above chance over twelve false belief trials) showed significant deficits on one or more of these independent measures. However, performance on the independent measures was also significantly impaired in one or more of the five patients who *passed* the false belief task. Thus, impairment on the independent measures could not explain patients' belief reasoning errors. Second, Apperly et al. (2004; Samson et al., 2004) found that three of the seven patients who failed the false belief trials performed perfectly on comparison trials designed to control for the incidental demands that the false belief trials made on inhibition and working memory. These three patients all had lesions to the left temporo-parietal-junction (TPJ). This finding adds to functional imaging studies that show selective activation of the TPJ during theory of mind tasks (Allison, Puce, & McCarthy, 2000; Saxe & Wexler, 2005; Saxe & Kanwisher, 2003) by suggesting that the TPJ is actually necessary for belief reasoning. However, while it seems safe to conclude that these patients have a relatively pure belief reasoning deficit these findings do not constitute a strong test of domain-specificity or modularity. As has already been described, reasoning about false beliefs requires the participant to process a conflicting relationship between a (mis)representation and the current situation that it misrepresents. This processing could plausibly make demands on domain-general executive functions that go beyond the incidental task demands controlled for by Apperly et al. (2004). Since our non-verbal false photograph task is specifically designed to match these processing demands of our non-verbal false belief task, comparison of performance on these conditions allows a much stronger test of the domain-specificity of the belief reasoning deficit in these three patients with TPJ lesions. If these patients pass the false photograph task, this would suggest that their deficit was not with handling representational relationships in general but was more specific to reasoning about beliefs. If, on the other hand, these patients are also impaired on the false photograph task this would suggest either that their deficit was due to difficulty with domain-general processes (most likely specific executive processes) that were shared by the belief and photograph tasks, or that cognitive apparatus for belief reasoning was also being recruited to handle non-mental representations such as photographs. Successful performance in the control conditions, though, would discount that any deficits reflect even more general limitations, for example in working memory or attentional disengagement.

2. Method

2.1. Participants

Eleven of the twelve patients reported in Apperly et al. (2004) were tested (PF, RH, DB, DS, FK, GA, PW, WBA, CN, MH and PH). The patients had lesions to temporal, parietal and/or frontal regions of the brain. These patients are described in Table 1. Their functional profiles are reported in Apperly et al. (2004).

2.2. Materials

Video-based false photograph task. The conditions for this task were based very closely on the non-verbal false belief task of Apperly et al. (2004), including trials corresponding to the false belief, true belief, memory control, inhibition control and clue confirmation trials used in this study. Participants watched short video sequences in which a character gave a clue to the location of a hidden object by revealing a photograph taken of the open boxes (see Fig. 1). In the video it appeared that the photo revealed to participants was the actual photograph taken with the camera. In fact, we surreptitiously switched for a pre-printed photograph, enabling the trials to be substantially shorter than if it had been necessary to wait for the photograph to develop. The task principles were explained to the participant at the beginning of each testing session, and comprehension was checked with a number of warm-up trials on which corrective feedback was given, as necessary.

In *false photograph trials,* the participants saw a female character taking a Polaroid photograph of the two open boxes, but they did not see in which box the object was located (see Fig. 1). The photograph was placed face-down on the table. A male character swapped the locations of the two boxes. He then gave the participant a clue

Table 1
Patients' characteristics and lesion description

Patient	Sex/age/ handedness	Main lesion site	Major clinical symptoms	Aetiology	Years post-onset
CN	M/47/R	Bilateral medial temporal lobes (more pronounced on left)	Mild amnesia	Herpes simplex encephalitis	10
DB	M/68/R	Left parietal inferior (angular gyrus), superior and middle temporal gyri	Aphasia	Stroke	6
DS	M/70/R	Left inferior, middle and superior frontal gyri	Right hemiplegia, aphasia	Stroke	14
FK	M/35/R	Bilateral superior and medial frontal regions, bilateral superior and medial temporal gyri, bilateral lateral occipital gyri	Agnosia, aphasia, dysexecutive syndrome	Anoxia	14
GA	M/49/R	Bilateral medial and anterior temporal lobes, extending into left medial frontal region.	Aphasia, amnesia, dysexecutive syndrome	Herpes simplex encephalitis	13
WBA	M/58/R	Right inferior and middle frontal gyri, right superior temporal gyrus	Aphasia	Stroke	3
MH	M/50/R	Left angular and supramarginal gyri; lentiform nucleus	Right extinction, optic ataxia	Anoxia	10
PF	F/55/R	Left inferior parietal (angular gyrus, supramarginal gyrus), and superior temporal gyrus	Right extinction, dysgraphia	Stroke	8
PH	M/31/R	Left medial and superior temporal, left inferior and middle frontal gyri	Right hemiplegia, aphasia	Stroke	5
PW	M/72/R	Right inferior and middle frontal gyri, right superior temporal gyrus	Left hemiplegia, dysexecutive syndrome	Stroke	4
RH	M/70/L	Left inferior parietal (angular and supramarginal gyrus) and superior temporal gyrus	Right neglect, aphasia	Stroke	8

about the location of the object by revealing the photograph (see Fig. 1 for representation of the event sequence in false photograph trials, compared with the sequence in false belief trials). The photograph was replaced face-down on the table, the video was paused and the participant was prompted to point to the box containing the object. To locate the object, participants needed to realise that the photograph was a (mis)representation of the true location of the object. Participants judged where they thought the object was located, then received feedback by viewing the end section of video-clip where the man opened the boxes and showed them to the camera.

False photograph trials required the participant to process the order of the events in the video, in particular that the photograph clue is revealed after the boxes have been swapped. To control for this incidental processing demand, *working memory control trials* reversed the order of clue-giving and box-swapping events. The photograph was revealed *before* the boxes were swapped, so enabling the participant to infer the location of the object. The photograph was replaced face-down on the table. The man then swapped the locations of the two boxes without opening them to reveal the object's location to the participant. The participant was prompted to point to the box containing the object. Thus, the participant had to use the fact that the boxes had swapped to update his or her knowledge of the object's location, and maintain this information until a response was requested. As for all other trial types, participants received feedback once they had given their response.

False photograph trials also required the participant to point to the opposite box from the one that the photograph showed to contain the object. Participants who lacked the inhibitory control to disengage their attention from the incorrect location would fail the task, whether or not they could reason about false photographs. *On inhibition control trials* the photograph was taken, then, in full view of the participant, the man moved the object from one box to another. The photograph was then revealed, showing the object in the box that the participant knew to be empty. The photograph was replaced face-down on the table. The participant was then invited to point to the box containing the object. As in the false photograph trials, a correct answer required participants to disengage their attention from the box that the photograph showed to contain the object. However, unlike false photograph trials, no reasoning about misrepresentations was required.

True photograph filler trials were designed to guard against participants passing the false photograph trials by adopting the strategy of always pointing to the opposite box from the one containing the object in the photograph. The photograph was taken, but the man did not swap the boxes, meaning that the photograph remained an accurate representation of the object's location. When the photograph was revealed, the participant had to point to the box that the photograph showed to contain the object. Although it was possible that participants were reasoning about the representational character of the photograph on these trials, we did not regard this as a reliable index of reasoning about representations as it is also possible to make a correct response simply based on the visual similarity of the photograph and the scene. This rationale directly follows the literature on theory of mind, where success on true belief trials is not regarded as a reliable index of belief reasoning (Dennett, 1978; Wimmer & Perner, 1983). The key point in the current study is that correct answers to the true photograph trials required the participant to point to the location indicated in the photograph, while correct answers to false photograph trials required the participant to point to the opposite location. Thus, if participants performed well on true photograph trials we could be confident that good performance on false photograph trials reflected genuine reasoning about representations, not a superficial strategy of pointing to the opposite box from that indicated in the photograph.

On clue confirmation filler trials the photograph was revealed immediately, to show the object's true location. The man opened this box to reveal the object, providing a very salient reminder that the photograph clues were being given in good faith. The man moved the object to the second box, and the participant was then prompted to respond. Despite these clue confirmation trials, a possible construal of false photograph trials might have been that the man was being deliberately deceptive in showing the participant the false photograph. Although we thought this unlikely, we were also satisfied that this unintended construal would be apparent because it would lead to errors on control and filler trials by causing participants to doubt the veracity of the photograph clues. Interpolation of both types of filler trial with experimental trials meant that experimental trials did not appear in any regular pattern and were not repeated in long sequences.

There were a total of 12 video trials of each type. The videos were presented on a standard desktop computer using PowerPoint software. Video presentation was controlled manually by the experimenter, enabling the time allowed for responding and the rate of progress to the next video to be adapted to the needs of the participant. The participant responded non-verbally by pointing to one of the two boxes on the screen, and this response was recorded by the experimenter. Each testing session lasted approximately twenty minutes and sessions were typically held at 1–2 week intervals. Photograph videos were somewhat shorter than the belief videos. We ensured that testing sessions were of a similar duration in photograph and belief conditions. Hence whereas the belief task was presented in 4 sessions of 15 trials (3 trials of each type), the photograph task was presented in 3 sessions of 20 trials (4 trials of each type). Trials were presented in a pseudo-random order designed to avoid runs of more than 2 trials of the same type. For each trial type overall, and across trial types within each session, half of the correct responses were the box on the right and half were the box on the left. All participants were first presented with the belief task then with the photograph task. The elapsed time between presentation of the last block of false belief trials and the first block of false photograph trials ranged from 10 to 14 months.

3. Results

All false photograph and control trials consisted of a binary choice-response, so performance was evaluated against a 50% chance baseline. For an individual to score statistically above chance on a particular trial type they needed to give 10 or more out of a possible 12 correct responses (10/12 correct has a one-tailed probability of 0.019 by binomial test).

From Fig. 2 it is clear that performance on false belief and false photograph trials was highly consistent. Every patient who had performed above chance on false belief trials (Apperly et al., 2004) also performed above chance on false photograph trials, whereas patients who were not above chance on false belief trials were not above chance on false photograph trials. All patients were above chance on true photograph trials, with the only errors being made by FK (10/12 correct). This finding is of particular relevance for WBA, CN, MH and PH, because it demonstrates that the above-chance performance of these patients on false photograph trials was not due to a superficial strategy of pointing to the opposite box from that indicated in the photograph.

Apperly et al. (2004) divided patients into three categories of response pattern, which were largely preserved on the conditions of the photograph task. The first group consisted of three patients (DB, PF and RH) who, on the belief task, failed on

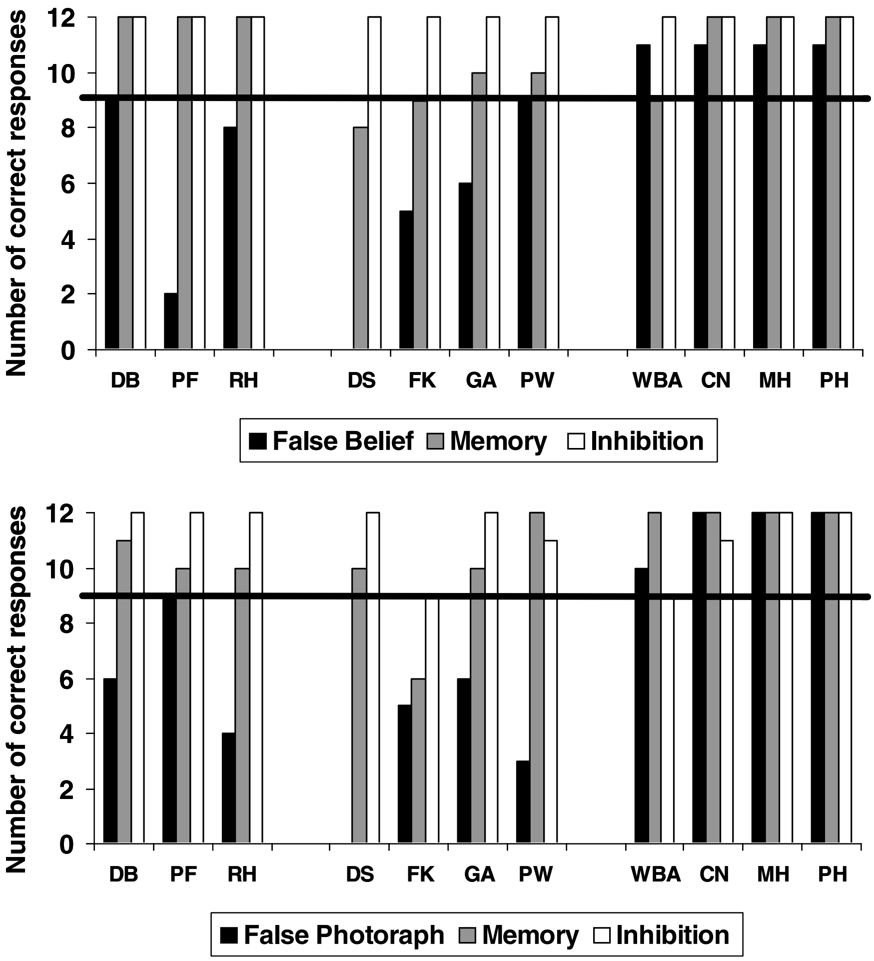

Fig. 2. Behavioural performance on the False Belief (upper graph) and False Photograph tasks (lower graph). The horizontal line indicates the level above which a participant is scoring above chance.

false belief trials but performed perfectly on memory and inhibition control trials. On the photograph task, these patients showed the same pattern for false photograph trials and inhibition control trials, but each made one (DB) or two (PF and RH) errors on memory control trials. Nonetheless, this performance on memory control trials was significantly above chance.

The second group (FK, DS, GA and PW) were patients who, on the belief task showed the following profile: They failed false belief trials, performed perfectly on inhibition control trials, and made errors on memory control trials (DS and FK were not above chance on memory control trials whereas GA and PW were above chance, but made two errors each). On the photograph task, all of these patients failed false photograph trials. On the inhibition control trials FK made three errors, so was not above chance. On memory control trials DS was above chance, making just 2/12 errors (compared with 4/12 errors on the corresponding condition of the belief task).

Finally, the third group of patients (WBA, CN, MH and PH), contained individuals who were above chance on false belief trials. The only patient in this group to make errors on control trials was WBA, who was not above chance on memory control trials. On the photograph task, all of these patients were above chance on false photograph trials. The only patient whose performance changed compared with chance was WBA, who performed perfectly on the memory control trials (better performance than equivalent trials in the false belief task) but was not above chance on inhibition control trials (worse performance than equivalent trials in the false belief task).

4. Discussion

The current study is the first to use a comparison of performance on false belief and false photograph tasks as a strong test of the domain-specificity of belief reasoning impairments in neurological patients. It is also the first study of any kind to use methods that de-confound reasoning about false beliefs and false photographs from the need to resist interference from knowledge of reality, and the first to employ a non-verbal false photograph task that enables the testing of participants with language impairment. The data showed overwhelming consistency in the performance of the eleven patients tested. On trials designed to control for incidental memory and inhibitory demands, performance was strikingly similar for belief and photograph tasks. This pattern suggests that the belief and photograph procedures were indeed well matched for incidental processing demands. Crucially, the four patients who passed the false belief task also passed the false photograph task, whereas the seven patients who failed the false belief task also failed the false photograph task. Thus, the current study provided no evidence that reasoning about false beliefs requires processes that are distinct from reasoning about false photographs.

Before discussing our findings in more detail we consider two potential methodological concerns with our study. First, the time-course of our investigations meant that all patients completed the false belief task before the false photograph task, with 10–14 months between testing on the two tasks. Could our findings be explained by spontaneous recovery of function over this time? It is noteworthy that each patient was a minimum of 3 years post-onset, making it unlikely that substantial changes in function would be observed. Perhaps more importantly, it seems highly unlikely that the observed pattern of consistency on false belief and false photograph tasks is a spurious effect resulting from spontaneous recovery of function or the learning of a response strategy in some or all of the patients. Had this been the case, we would have expected to observe better performance in the false photograph task compared with the false belief task, whereas in fact performance across the two tasks was consistent for all patients. A second methodological concern is that our novel false belief and false photograph tasks may not have been well-matched in their task demands. In defence of our methods it is important to note that the event sequences of false belief and false photograph tasks were closely matched, and there do not seem to be clear grounds for supposing that reasoning about a photograph would make different demands on working memory and inhibitory control compared with reasoning about a belief (that had been made manifest in the placement of a marker on a box). In addition it is very difficult to see how a pattern of consistent performance across our false belief and false photograph tasks could be the result of incidental or unexplained differences in task difficulty. Indeed, the striking consistency of the patients' performance on these tasks lends support to our contention that the tasks are well-matched.

The current findings suggest that the relatively specific belief reasoning deficit in the DB, PF and RH reported by Samson et al. (2004; Apperly et al., 2004) is not limited to false belief problems but extends to closely matched problems about false photographs. Thus, although the method we used had the potential to provide strong evidence in favour of domain-specific mechanisms for belief reasoning located in the left TPJ, none was found.

One interpretation of these findings is that PF, DB and RH have an intact cognitive system for belief reasoning, but the successful operation of this system is obscured by a deficit in domain-general executive processes that generates performance errors on both false belief and false photograph tasks. It is undoubtedly true that our video-based materials led to least two potential sources of performance error because participants needed to maintain and update information about the swapping of boxes or box contents and to point to the opposite box from the one indicated by the woman, or that the photograph showed to contain the object. It is possible to evaluate this interpretation by considering whether the patients were likely to have been able to meet these and other performance demands. First, DB, PF and RH performed perfectly on working memory and inhibition control trials that were specifically designed to control for performance demands of the false belief task. For the equivalent control trials of the false photograph task these patients made one or two errors, but were significantly above chance. Second, Apperly et al. (2004) report that these three patients did not show a distinctive pattern of impairment on a range of independent tests of working memory, inhibitory control or language compared with patients who passed the non-verbal false belief task. Third, by eliminating the need for participants to resist interference from knowledge of reality we eliminated a performance demand that is confounded with false belief and false photograph reasoning in standard methods. Altogether, these considerations provide strong reasons for thinking that DB, PF and RH possessed sufficient domain-general executive resources to meet the incidental performance demands of the false belief and false photograph tasks. This makes it unlikely that performance errors obscured the existence of an intact belief reasoning ability in these patients.

Our findings are clearly compatible with the possibility that domain-general processes have a necessary role in belief reasoning itself, or even that belief reasoning is achieved exclusively with domain-general resources. On this view, DB, PF and RH possessed sufficient domain-general resources to meet the performance demands of the tasks, but lacked the necessary domain-general resources for reasoning about beliefs and photographs per se. This could arise because reasoning about beliefs and photographs makes greater demands on domain-general resources, or demands that are different in nature from those assessed by the independent tests of executive function or the memory and inhibition control trials. This suggestion is consistent with claims that there are discrete components within "executive function" that may be independently spared or impaired by brain damage (e.g., Stuss et al., 2002). It is also consistent with the view that metarepresentational reasoning (about beliefs, photographs and other representations) makes distinctive cognitive demands on domain-general reasoning processes (e.g., Perner, 1991). Importantly, our methods help to locate these demands more precisely in the process of metarepresentational reasoning, rather than in the process of resisting interference from knowledge of reality (see also Apperly et al., 2005).

Another explanation of our findings is that PF, DB and RH have damage to a domain-specific reasoning mechanism, but that the mechanism's specialized domain extends beyond beliefs to include, at a minimum, photographs. This would be an important modification of the view that adult belief reasoning depends upon domain-specific mechanisms because it suggests that the specialization might be for metarepresentational reasoning, not for theory of mind alone. However, it is clearly compatible with a variety of developmental accounts, including the view that children are innately equipped with a theory-of-mind mechanism (e.g., Leslie, German, & Polizzi, 2005; Leslie & Thaiss, 1992). For example, a domain-specific mechanism developed phylo- or ontogenetically to handle beliefs might subsequently be recruited to enable older children and adults to solve formally similar reasoning problems about photographs and other external representations such as words. Alternatively, children might initially rely exclusively upon domain-general processes for reasoning about beliefs, photographs and the like, but these abilities could become progressively modularized and domain-specific in a developmental process. In the mature system of our adult participants, all such reasoning problems would make use of a mechanism that was relatively independent of domain-general processes (for a detailed exposition of the possibilities of modularisation in developmental processes see e.g., Elman et al., 1997; Karmiloff-Smith, 1992). Thus, the proposition that adults make use of the same specialized processing mechanisms for reasoning about beliefs and photographs allows for a variety of accounts of the nature of these processes in young children. However, it would nonetheless constrain developmental theories by characterizing the endpoint reached by the developmental process.

Finally, it is possible that the errors made by DB, PF and RH on false belief trials arise from damage to domain-specific processes for belief reasoning, whereas their errors on false photograph trials arise from damage to nearby brain regions that support domain-general executive processes. On grounds of parsimony we believe the proposition of two deficits is less plausible than the possibility that a single deficit accounts for errors on *both* false belief and false photograph tasks. However, the existence of two deficits, and therefore the existence in adults of a domain-specific process for belief reasoning, cannot be ruled out on the strength of the current data.

Evidence to distinguish between domain-general and domain-specific accounts for PF, DB and RH (and other patients or populations with a similar functional profile), will come from further tests of the scope of the impairment to reasoning abilities. For example, reasoning about desires is a "theory of mind" problem, and conflicting desire tasks may make similar processing demands to false belief tasks (e.g., Russell et al., 1999). However, many theorists hold that reasoning about desires is unlike reasoning about beliefs and photographs because desires are not *representational* mental states (e.g., Perner, 1991; Russell et al., 1999). Similarly, counterfactual reasoning may have some formal similarities to belief reasoning (Peterson & Riggs, 1999) and children's counterfactual reasoning correlates with their performance on false belief tasks (Riggs, Peterson, Robinson, & Mitchell, 1998). However, in common with desire reasoning, counterfactual reasoning does not entail reasoning about representations. By using such tasks it should be possible to chart the nature of the deficit in a systematic way. The general principle would be that the wider the range of tasks on which PF, DB and RH were impaired, the more difficult it would become to uphold the notion that their deficit was domain-specific.

Interestingly, Apperly et al. (2004; Samson et al., 2004) found that PF performed perfectly across 12 trials of a language-based counterfactual task, suggesting one limit on the scope of this patients' deficit. Language problems meant that DB and RH could not be tested reliably with this task. To exploit the full potential of a neuropsychological approach to the study of "theory of mind" it will be necessary to continue developing simplified tasks that can be used reliably with a larger number of potentially interesting patients with brain damage.

4.1. Contrasts between our data and other literature

The findings in the current study contrast with the dissociation observed in the developmental literature between children with autism (who tend to perform better on false photograph tasks than on false belief tasks) and typically developing children (who show no difference in performance or perform better on false belief than on false photograph tasks) (Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992). One potential explanation for this contrasting pattern follows from concerns about the validity of the comparison between false belief tasks and the "false" photograph tasks used in these studies of children (e.g., Perner, 1995; Russell et al., 1999). As summarized in the introduction, this raises the possibility that the dissociation reported between the performance of children with autism versus typically developing children is simply an artifact arising from poor matching of false belief and false photograph tasks. Another possibility is that the dissociation observed in children does indeed reflect the existence of a specialized mechanism for reasoning about beliefs, but that in adults the same mechanism has been co-opted for reasoning about photographs. It may be possible to address these questions in new studies of children with and without autism that use age-appropriate adaptations of our new false photograph and false belief methods. Such studies would also benefit from the fact that our methods de-confound the need to reason about beliefs or photographs from the need to resist interference from knowledge of reality, which may contribute the errors of both typically and atypically developing children on standard tasks.

Our findings also contrast with functional imaging studies that have shown neural activation associated specifically with false belief reasoning in comparison with "false" photograph reasoning (Sabbagh & Taylor, 2000; Saxe & Wexler, 2005; Saxe & Kanwisher, 2003). PF, DB and RH all have lesions to left temporo-parietal junction. It is notable that fMRI studies have shown bilateral activation of the TPJ when false belief reasoning is contrasted with "false" photograph reasoning (Saxe & Wexler, 2005; Saxe & Kanwisher, 2003). Moreover, the right-TPJ regions were activated more specifically than left TPJ regions by other tasks that required processing of the thoughts and feelings of story characters (Saxe & Wexler, 2005) indicating that right rather than left TPJ may be have a particularly specific role in theory of mind. In attempting to make sense of these findings it is important to recognize that all three imaging studies employed conventional false belief and "false" photograph tasks, and so inherit several concerns about the validity of this comparison (Perner, 1995; Russell et al., 1999). However, these problems notwithstanding, the observation of specific neural activation for false belief tasks (but not false photograph tasks) is perfectly compatible with the existence of patients with a relatively specific deficit on *both* false belief and false photograph tasks. It would merely be necessary to propose that the complex task of inferring and representing beliefs depends upon at least two kinds of process[1], one which is also used for inferring and representing non-mental representations, such as photographs, and one which is not.

A candidate for the *common* process in false belief and false photograph tasks is the need to perform a relatively specific kind of reasoning about mental and non-mental representations. It would be this ability that is affected in patients PF, DB and RH. For the reasons already discussed, it remains for future work to determine the degree to which such reasoning depends upon domain-general or domain-specific processes, and whether it is narrowly "metarepresentational" reasoning, or whether it includes a broader class of problems such as reasoning about desires and counterfactuals. This question may be addressed in further studies of patients. Additionally, the current analysis suggests that it may be fruitful for functional imaging studies to examine which, if any, brain regions are selectively involved in reasoning about beliefs *and* about photographs, alongside the strategy used in existing studies of identifying brain regions selectively involved in reasoning about beliefs and *not* photographs.

We might also speculate that the process that is crucially *different* across false belief and false photograph tasks is the use of different forms of semantic knowledge. Reasoning about beliefs may recruit semantic knowledge for "theory of mind", whereas reasoning about photographs recruits distinct semantic knowledge for artifacts (see also Frith & Frith, 2003; Stone, 2006). This suggestion seems compatible with Saxe and Wexler's (2005) finding that the left and right TPJ regions activated specifically for false belief reasoning (and not false photograph reasoning) were also

[1] There may well be more than two processes. For example, it has been suggested that belief reasoning depends upon the ability to resist interference from self-perspective (e.g., knowledge of reality in contrast to a false belief or one's own visual perspective as opposed to that of another). There is preliminary evidence to associate such processes with frontal brain regions (e.g., Frith & Frith, 2003; Samson et al., 2005). It remains to be seen whether such processes are domain-specific for theory of mind problems, or whether they are domain-general inhibitory processes.

activated when participants read sentences about the thoughts and feelings of story characters – i.e., a task that requires "theory of mind" semantics, though not necessarily theory of mind *reasoning*.

In conclusion, it is clear that solving false belief tasks makes demands on domain-general processes such as working memory and inhibitory control. Much more controversial is whether such domain-general processes exhaust the functional requirements of belief reasoning, or whether belief reasoning also requires devoted functional and anatomical mechanisms. We believe that the non-verbal false photograph and false belief tasks described in the current study offer a better instrument for investigating domain-specificity than existing methods. A clear dissociation between performance on these tasks in future studies of patients, children with developmental disorders, or in imaging studies of typical adult brains would count as strong evidence in favour of a domain-specific component to belief reasoning. Importantly, if dissociations were discovered, this would not exclude the possibility of future studies also confirming our observation of systematic association of impaired performance, or of demonstrating consistent patterns of neural activation for reasoning about false beliefs and false photographs. This follows from the likelihood that reasoning about beliefs (and by extension, other kinds of theory of mind task) is a relatively complex problem that is likely to depend upon multiple processes and multiple brain regions. An important task for future work is to identify these distinct components of belief reasoning, and the degree to which each one makes use of domain-general or domain-specific processes.

Acknowledgements

This paper was supported by grants from the Leverhulme Trust, the MRC and the Stroke Association. We are very grateful to all the participants for their kind participation.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.

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., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: evaluating evidence from neuropsychology. *Trends in Cognitive Sciences, 9*(12), 572–577.

Astington, J. W., Harris, P. L., & Olson, D. R. (1998). *Developing theories of mind* (First ed.). Cambridge: CUP.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2001). *Understanding other minds: Perspectives from developmental cognitive neuroscience* (Second ed.). New York: OUP.

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

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences, 8*, 255–260.

Call, J., & Tomasello, M. (1999). A nonverbal false belief task: the performance of children and great apes. *Child Development, 70*, 381–395.
Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053.
Charman, T., & Baron-Cohen, S. (1992). Understanding drawings and beliefs: a further test of the metarepresentation theory of autism: a research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*, 1105–1112.
Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 4*, 568–570.
Elman, J., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1997). *Rethinking innateness: A connectionist perspective on development*. Cambridge MA: MIT.
Flavell, J. H. (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology, 50*, 21–45.
Frith, C. D., & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychological Medicine, 26*, 521–530.
Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*, 459–473.
Happe, F., Malhi, G. S., & Checkley, S. (2001). Acquired mind-blindness following frontal lobe surgery? A single case study of impaired 'theory of mind' in a patient treated with stereotactic anterior capsulotomy. *Neuropsychologia, 39*, 83–90.
Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology, 12*, 426–445.
Karmiloff-Smith, A. (1992). *Beyond Modularity*. MIT.
Langdon, R., & Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. *Cognition: International Journal of Cognitive Science, 71*, 43–71.
Lee, K. H., Farrow, T. F. D., Spence, S. A., & Woodruff, P. W. R. (2004). Social cognition, brain networks and schizophrenia. *Psychological Medicine, 34*, 391–400.
Leekam, S. R., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203–218.
Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*, 45–85.
Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development – neuropsychological evidence from autism. *Cognition, 43*, 225–251.
Lewis, C., & Mitchell, P. (1994). *Children's early understanding of mind: Origins and development* (First ed.). Hove: Lawrence Erlbaum Associates.
Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.
Mitchell, P., & Riggs, K. J. (2001). Children's reasoning and the mind(vols. 1). Hove: Psychology Press.
Mitchell, P., & Lacohee, H. (1991). Children's early understanding of false belief. *Cognition: International Journal of Cognitive Science, 39*, 107–127.
Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press, A Bradford Book.
Perner, J. (1995). The many faces of belief: reflections on Fodor's and the child's theory of mind. *Cognition, 57*, 241–269.
Peterson, D. M., & Riggs, K. J. (1999). Adaptive modelling and mindreading. *Mind & Language, 14*(1), 80–112.
Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13*, 73–90.
Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). "I know, you know: epistemic egocentrism in children and adults. *Review of General Psychology, 7*, 38–65.
Russell, J., Saltmarsh, R., & Hill, E. (1999). What do executive factors contribute to the failure on false belief tasks by children with autism? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 859–868.
Sabbagh, M. A., & Taylor, M. (2000). Neural correlates of theory-of-mind reasoning: an event-related potential study. *Psychological Science, 11*, 46–50.

Saltmarsh, R., Mitchell, P., & Robinson, E. (1995). Realism and children's early grasp of mental representation: belief-based judgements in the state change task. *Cognition: International Journal of Cognitive Science, 57*, 297–325.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience, 7*, 499–500.

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.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia, 43*(10), 1391–1399.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55*, 87–124.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *NeuroImage, 19*, 1835–1842.

Slaughter, V. (1998). Children's understanding of pictorial and mental representations. *Child Development, 69*, 321–332.

Sperber, D. (2000). *Metarepresentations: A multidisciplinary perspective* (first ed.). Oxford: OUP.

Stone, V. E. (2006). An evolutionary perspective on domain-specificity in social intelligence. In P. Winkielman & J. C. Harris (Eds.), *Social cognitive neuroscience*. NY: Guilford Press.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10*, 640–656.

Stuss, D. T., Alexander, M. P., Floden, D., Binns, M. A., Levine, B., Mcintosh, A. R., et al. (2002). Fractionation and localization of distinct frontal lobe processes: evidence from focal lesions in humans. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 392–407). New York: OUP.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

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.

Zaitchick, D. (1990). When representation conflicts with reality: the preschooler's problem with false belief and "false" photographs. *Cognition: International Journal of Cognitive Science, 35*, 41–68.