

Preprint (post-review) of the following paper:

Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind : Evaluating neuropsychological evidence. *Trends in Cognitive Sciences*, 9(12), 572-577. <https://doi.org/10.1016/j.tics.2005.10.004>

Domain-specificity and theory of mind: Evaluating evidence from neuropsychology

Ian A. Apperly, Dana Samson and Glyn W. Humphreys

University of Birmingham, UK

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

Corresponding author: Apperly, I.A. (i.a.apperly@bham.ac.uk)

Abstract

Humans' unique aptitude for reasoning about mental states (Theory of Mind) may help explain the unique character of human communication and social interaction. ToM has been studied extensively in children, but there is no clear account of the cognitive basis of ToM in adults. Evidence from functional imaging and neuropsychology is beginning to address this surprising gap in our understanding, and is often thought to favour a domain-specific or modular architecture for ToM. We present a systematic approach to this issue for the paradigm case of belief reasoning, and argue that neuropsychological data provide no clear evidence for domain-specificity or modularity. Progress in understanding ToM requires new tasks that isolate potentially distinct components of this complex ability.

For adult humans it is second nature to think about mental states when explaining or predicting behaviour. For instance, when we are told something we know to be untrue we may make an inference about that person's mental state: we might consider whether the speaker is mistaken, whether he or she is lying to protect our feelings, or whether the person is trying to mislead us maliciously. Most research on Theory of Mind has studied this use of mental state concepts – beliefs, desires, intentions and the like - to frame explicit predictions and explanations of behaviour. Twenty years of research shows that these abilities may be unique to humans¹, they undergo protracted development (from infancy through middle childhood at least^{2,3}) and may be disproportionately impaired in clinical disorders such as autism and schizophrenia⁴⁻⁷ (see Box 1). More recently, brain imaging and neuropsychological studies have begun to explore the basis of adult ToM abilities. These studies suggest that ToM involves a network of regions within the medial prefrontal cortex, the temporal poles and the temporo-parietal junction^{3,8-16}. However, the roles of these brain regions are not well understood because there is no consensus about the cognitive requirements of ToM tasks. If the promise of these techniques is to be fulfilled we need tasks that allow the cognitive basis of ToM to be investigated more precisely.

A key debate in the existing literature concerns the degree to which ToM depends upon specialised processes, devoted to the purpose of ToM computations (domain-specific processes)^{3,4,9,17} or upon processes such as language and executive function that also serve other cognitive functions (domain-general processes)¹⁸⁻²⁰. Studying the patterns of association and dissociation of deficits in adults with acquired brain injury provides a unique source of evidence on this issue. If ToM is sustained by domain-specific processes we might expect a brain lesion to affect performance on ToM tasks while sparing performance on matched tasks that do not involve ToM. Also, we would expect that impaired domain-general processes are not necessarily associated with impaired ToM and

vice versa. In the current article we evaluate the neuropsychological evidence for the paradigm case of reasoning about beliefs, for which the strongest claims about domain specificity have been made^{3,4,9,17}. We argue that there is no clear evidence for domain-specificity because existing methods fail to separate belief reasoning itself from other processes associated with belief reasoning and with belief reasoning tasks.

False belief tasks

Tests of belief reasoning usually require acknowledgement that a belief is *false* because this guarantees that the participant cannot answer correctly from their own perspective^{21,22}. In one such task, participants are told a story where Sally puts her toy in a basket, then goes outside to play. In her absence, Anne moves the toy from the basket to the box. When Sally returns, participants are asked where she (Sally) will first look to find her toy⁷. Most adults successfully take account of Sally's false belief and predict that she will look in the basket. In contrast, children below four years of age tend to judge incorrectly that Sally will look in the box (i.e., where the child knows the toy is hidden)²³. Similar errors have been observed in adults with neurological damage^{10,12-15} and clinical disorders such as autism and schizophrenia^{4,5,7,24}.

The value of evidence from neuropsychological studies of belief reasoning depends upon being able to separate the contribution that domain-general processes such as language or executive function undoubtedly make to *performance* on false belief tasks, from any contribution they may also make to the *competence* to reason about beliefs at all. For example, working memory (a domain-general process) is necessary for maintaining and updating crucial information for completing the false belief task (e.g., where the toy began, where it ended up and who was present when it moved). Patients with impaired working memory could fail false belief tasks merely because they cannot meet these *performance* demands. However, it could also be that belief reasoning itself is carried out

in a generic working memory system. If so, then even if performance demands are adequately controlled, sufficiently impaired working memory would *necessarily* lead to a loss of belief reasoning *competence* (on the distinction between executive performance versus competence see²⁵). Finally, it is possible that components of belief reasoning are carried out in domain-specific or even modular processes devoted exclusively to this purpose. For example, it has been suggested that belief reasoning has unique representational requirements that are served by a specialised, domain-specific module⁴. By definition, modular processes would not rely upon domain-general working memory resources²⁶. It follows that if performance demands are adequately reduced then it should be possible to observe spared belief reasoning competence despite impaired working memory and vice versa. Thus, investigating the cognitive basis of belief reasoning requires 1) methods that allow clear separation of performance and competence demands 2) methods that allow competence to be investigated with sufficient precision.

Do existing false belief tasks allow a clear separation of performance and competence demands?

Demands on memory and language. Existing studies have reduced the likelihood of performance errors by reducing the pragmatic demands of linguistic test questions²⁷, by using pictures to support memory for story-based stimuli¹⁴, and by removing language entirely from the false belief tasks^{8,10} (see Figure 1). It is also common to test whether a patient can meet some of the performance demands that remain. To infer that Sally has a belief that is false it is necessary to remember where she put her toy and where the toy is now. Memory for these facts may be checked directly with explicit questions^{7,14}.

Alternatively, separate trials may be designed, where successful performance requires the participant to track similar object movements, but do not require belief reasoning^{8,10}. If a patient fails memory controls then the parsimonious assumption is that belief reasoning

errors are due to general performance problems^{10,14}. If a patient shows good performance on memory control trials then at least we can be sure that failure to remember basic facts about the story is not the cause of failure on false belief trials^{10-16,28-32}.

Interference from knowledge of reality. Standard false belief tasks confound the need to infer a false belief (e.g., that Sally thinks the toy is in the basket) with the need to resist interference from what is true from self-perspective (the toy is actually in the box). Since errors could occur in either process, separating these demands is essential for accurate assessment of belief reasoning problems in populations with impaired executive function. This is not only the case for people with brain damage, but also for people with schizophrenia or developmental disorders such as autism. Most existing studies *fail* to separate these processes.

In the only neuropsychological study to address this problem we compared performance on standard “reality-known” tasks with “reality-unknown” false belief tasks where the participant does not know the true state of reality (the location of an object) about which a target person has a false belief¹¹ (see Figure 1). WBA, who has a right fronto-temporal lesion, performed above chance on reality-unknown false belief trials when he lacked crucial knowledge of reality. However, he was not above chance on reality-known false belief trials, and he performed poorly on a series of other tasks where he had to set aside his own desire or a visual or emotional perspective in order to judge from someone else’s point of view.

The pattern shown by WBA is consistent with a problem in inhibiting ‘self-perspective’, rather than with belief reasoning per se. In WBA’s case, impairment exists across a range of social and non-social tasks demanding inhibitory control, so the simplest explanation is that WBA has impairment of a domain-general inhibitory process. However, it cannot be ruled out that WBA’s lesion is large enough to have affected separate social

and non-social inhibitory processes in adjacent brain regions. Future studies may provide more conclusive evidence by showing the degree to which dissociation is observed between self-perspective inhibition and inhibitory control on non-social tasks, and the degree to which overlap is observed in the lesions of patients who are impaired in one or both of these abilities.

The inhibition of self perspective is an interesting topic in its own right, but the important point for the current discussion is that most existing studies do not achieve a clear separation of belief reasoning competence from the performance demand of resisting interference from knowledge of reality (references 10 and 11 are exceptions). This is also true for imaging studies that report specific activation for false belief reasoning (see Box 2). Later, we describe how this confound restricts what these studies can tell us about the domain-specificity of belief reasoning.

Do existing methods allow precise investigation of belief reasoning competence?

Proceeding on the basis that existing studies are at least partially successful in controlling for the performance demands of belief reasoning tasks, we describe two approaches to the question of whether belief reasoning competence depends upon domain-general or domain-specific processes.

Standard measures of domain-general processes. The existence of spared belief reasoning in a patient with severe impairments of language or executive function is strong evidence that belief reasoning does not depend upon these processes. This approach has been taken in a small number of studies, which suggest that belief reasoning may remain intact despite impairment on several tests of grammar³³ and executive function³⁴.

On the other hand, if impaired belief reasoning competence is systematically associated with impaired performance on standard measures of language or executive function, this would be evidence that belief reasoning depended upon such processes.

Such associations have been observed, but the patterns are not consistent across different studies, patients or tasks (i.e., different standard measures and belief reasoning tasks)^{12,13,15,30,31}.

The balance of evidence here suggests that belief reasoning may not depend upon the domain-general processes that have been assessed in existing studies. However, this does not warrant the conclusion that belief reasoning is independent of domain-general processes in general. Although belief reasoning may survive impairment of grammar³³ we do not know how it is affected by semantic impairment, and although it may survive impairment of verbal fluency and some planning abilities³¹ other forms of executive impairment could lead to belief reasoning difficulties. Moreover, belief reasoning may depend upon multiple domain-general processes. Performance on such complex executive tasks may not have simple relationships with their component executive process. For example, Stuss et al³⁵ studied the effect of different brain lesions on the complex executive task of selecting a visually cued location following compatible or incompatible primes. They found that component executive processes (e.g., inhibiting automatic return to a previously correct location) were disrupted in different ways and required different neural systems depending upon the level of overall task difficulty. Such complex patterns suggest that standard measures designed to test aspects of language and executive function in isolation are likely to be blunt instruments for testing the dependence of belief reasoning on domain-general processes. A more powerful approach may be to look for dissociations between performance on false belief tasks and other reasoning tasks that demand a similar *combination* of domain-general processing demands but not belief reasoning.

Matched reasoning tasks. Existing evidence shows that patients who make belief reasoning errors need not be impaired for all types of inference; they may be able to make

simple non-mental inferences, such as appreciating that ink seen splashing from an ink-pot at one point in time will have landed on the table at a later point in time^{14,29}. However, theoretical accounts suggest that such simple physical inferences are not well matched to the demands of belief reasoning^{19,36}. Tasks that require reasoning about misrepresenting photographs are one example of a more closely related reasoning problem^{4,37} (see Figure 1). False belief and false photograph tasks follow very similar event sequences, and require a similar inference about a misrepresentation, which in one case is a belief and in the other a photograph. Closely matched tasks of this kind should be a powerful tool for investigating the domain-specificity of belief reasoning competence. However, no published neuropsychological study has compared false belief and false photograph tasks.

The success of this approach will depend upon careful analyses of the processing demands of the matching tasks. Like standard false belief tasks, existing false photograph methods contain the confounding requirement to resist interference from knowledge of reality (Figure 1a and 1b). Of course, this should make standard false photograph tasks a good comparison for standard false belief tasks (though see^{38,39}). But it means that we could not be sure how to interpret a pattern where patients performed better on false photograph tasks than on false belief tasks: their impairment could be with a domain-specific process for representing beliefs, or a domain-specific (i.e., belief-specific) process for resisting interference from reality, or both. Progress will depend upon developing new tasks that separate possible sub-processes of belief reasoning, and test the domain-specificity of these sub-processes independently. For example, we have recently created a “reality-unknown” false photograph task (Figure 1d) to match the “reality-unknown” false belief task (see Figure 1c). Preliminary data from eleven patients show no evidence for a dissociation: Performance was consistent (i.e., above chance or not) on both false belief and false photograph tasks⁴⁰. For these patients, whatever impairment led to errors on false belief trials was general enough in nature to lead to errors on false photograph trials.

Future investigations may reveal patients who show dissociations between performance on such tasks. If not, then this could be because false belief and false photograph reasoning have the same neural basis, or because it is very uncommon for a patient to have a sufficiently specific lesion to dissociate these processes. Data from neuroimaging will provide valuable converging evidence on this question (see Box 2).

Conclusion

The idea that reasoning about beliefs depends upon domain-specific cognitive processes, is a strong theoretical claim that requires strong empirical evidence. In view of the contributory factors that need to be ruled out to make the case for domain-specific processing, our conclusion is that current data from brain-damaged patients provide no compelling evidence in support of domain-specificity for belief reasoning. Of course selective impairments may exist, but they have not so far been demonstrated. Progress on this issue depends critically upon the development of comparison and control tasks that enable belief reasoning performance to be distinguished from competence and the basis of belief reasoning competence to be decomposed. If clear evidence of domain specificity does emerge, then an important challenge will be to explain how domain-specific or modular processes are integrated with the domain-general language and executive processes that undoubtedly play an important role in ToM.

Acknowledgements.

This work was supported by grants from the Leverhulme Trust, the Stroke Association and the Medical Research Council (UK) to the authors.

Box 1. Developmental studies and the domain-specificity of belief reasoning

Developmental and comparative psychology has led the way in research on ToM, and studies of typically and atypically developing children yield most evidence about the cognitive basis of belief reasoning. The finding that children's success or failure on false belief tasks is related to their performance on tests of inhibitory control¹⁸, working memory⁴¹ and language²⁰ suggests that these general cognitive processes may have a vital role in belief reasoning. However, a recent meta-analysis of 178 studies argued that such factors were insufficient to explain age-related change in performance on false belief tasks²³. Potentially compelling evidence for domain-specific ToM processes comes from the finding that children with autism may fail false belief tasks but pass closely matched reasoning tasks (false photograph tasks - see Figure 1)⁴. However, the adequacy of this matching is still debated and it remains to be seen whether these findings reflect a genuine dissociation^{38,39,42}.

Whatever consensus emerges from developmental research, such data require cautious interpretation because it is often unclear whether patterns of association or dissociation between false belief tasks and other measures arise because a process is involved in children's on-line belief reasoning, or in the *development* of belief reasoning abilities. There are likely to be continuities between the systems of children and adults. However, it is quite possible that a domain-general process, such as language, is a vital source of information²⁰ or representational structure⁴³ for ToM development, but has no necessary role in mature belief reasoning⁴⁴. Equally, domain-specific processes may be critical for development, but play no necessary role in the mature system. For example, it has been suggested that a mechanism for representing objects of shared visual attention provides infants with valuable data for the acquisition of concepts such as "belief"⁴⁵. But once acquired, these concepts might be employed in a domain-general reasoning process that does not require input from processes that were necessary for their acquisition⁴⁶.

A clearer account of the functional basis of ToM in the mature system will help disentangle the roles of domain-general and domain-specific processes in the development of these abilities. By studying adults it is possible to gain direct evidence about the cognitive basis of on-line belief reasoning without the confounding possibility that the data reflect a developmental process.

Box 2. Imaging studies and the domain-specificity of belief reasoning

The usefulness of neuroimaging methods for assessing the domain-specificity of belief reasoning depends upon being able to distinguish the neural activation due to domain-general processes (e.g., on working memory or language) from any activation due to domain-specific process involved exclusively in belief reasoning.

One approach is to compare the patterns of activation seen in studies using diverse materials (e.g., stories, videos and cartoons) and diverse ToM problems (e.g., reasoning about beliefs, intentions, knowledge and emotions). Assuming that ToM is all that these tasks have in common, then any brain areas that are activated consistently should reflect ToM processing. This meta-analytic strategy identifies a relatively circumscribed network of regions within the medial prefrontal cortex (mPFC), the temporal poles and the temporo-parietal junction (TPJ)^{3,9}. There is evidence that some of this activation reflects domain-specific ToM processes. For example, regions of peak activation for ToM tasks do not overlap with mPFC regions activated in studies of executive processes³ or with TPJ regions activated in studies of relevant language processes⁹. However, as with neuropsychological studies, we believe that a strong test of domain-specificity demands comparison tasks designed to match the *combination* of domain-general processing demands made by ToM reasoning.

Stronger tests of this kind have been attempted in the case of belief reasoning, where researchers have identified a similar network of regions, principally in mPFC and TPJ, which are activated more strongly when subjects reason about the content of a false belief compared with a closely matched false photograph condition⁴⁷⁻⁴⁹. However, the power of these methods depends critically upon the matching of false belief and false photograph tasks, and here we suggest two reasons for caution. First, it remains controversial whether existing false photograph tasks are a good match to the conceptual demands of false belief tasks^{38,39,42}. Second, standard false belief and false photograph

tasks both involve at least two processes: inferring and representing the content of the false representation (that the person's mind or photograph represents object in location A), and resisting interference from knowledge of reality (that the object is in location B). A brain region specifically activated in such false belief tasks but not in false photograph tasks might sustain one or both of these processes. A full understanding of the neural basis of belief reasoning will ultimately require methods that are capable of making these, and yet more fine-grained functional distinctions.

Box 3: Questions for future research

- To what extent can belief reasoning competence be fractionated? For instance could the possession of mental state concepts such as BELIEF dissociate from the inferential processes that enable such concepts to be attributed to a particular person (e.g., Sally) with a particular content (e.g., that the toy is in the basket).
- To what extent are the components of belief reasoning (or other aspects of ToM) automatic or controlled processes?
- To what extent does reasoning about different mental states (beliefs, desires intentions etc.) depend upon shared versus unique functional and neural systems?

Reference List

- 1 Tomasello, M. *et al.* (2003) Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends Cogn. Sci.* 7, 153-156
- 2 Robinson, E.J. (2003) Six-year-olds' contradictory judgments about knowledge and beliefs. *Trends Cogn. Sci.* 7, 235-237
- 3 Frith, U. and Frith, C.D. (2003) Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 358, 459-473
- 4 Leslie, A.M. and Thaiss, L. (1992) Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition* 43, 225-251
- 5 Corcoran, R. *et al.* (1995) Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophr. Res.* 17, 5-13
- 6 Langdon, R. *et al.* (2002) Disturbed communication in schizophrenia: the role of poor pragmatics and poor mind-reading. *Psychol. Med.* 32, 1273-1284
- 7 Baron-Cohen, S. *et al.* (1985) Does the autistic child have a "theory of mind"? *Cognition* 21, 37-46
- 8 Samson, D. *et al.* (2004) Left temporoparietal junction is necessary for representing someone else's belief. *Nat. Neurosci.* 7, 499-500
- 9 Saxe, R. *et al.* (2004) Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55, 87-124
- 10 Apperly, I.A. *et al.* (2004) Frontal and left temporo-parietal contributions to theory of mind: neuropsychological evidence from a false belief task with reduced language and executive demands. *J. Cogn. Neurosci.* 16, 1773-1784
- 11 Samson, D. *et al.* (2005) Seeing it my way: a case of a selective deficit in inhibiting self-perspective. *Brain* 128, 1102-1111
- 12 Lough, S. *et al.* (2001) Dissociation of social cognition and executive function in frontal variant frontotemporal dementia. *Neurocase* 7, 123-130
- 13 Rowe, A.D. *et al.* (2001) "Theory of mind" impairments and their relationship to executive function following frontal lobe excisions. *Brain* 124, 600-616
- 14 Stone, V.E. *et al.* (1998) Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* 10, 640-656
- 15 Gregory, C. *et al.* (2002) Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain* 125, 752-764

- 16 Happe,F. *et al.* (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
- 17 Friedman,O. and Leslie,A.M. (2005) Processing demands in belief-desire reasoning: inhibition or general difficulty? *Developmental Science* 8, 218-225
- 18 Carlson,S.M. *et al.* (2004) Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology* 87, 299-319
- 19 Perner,J. (1991) *Understanding the representational mind*, MIT Press: A Bradford Book
- 20 Astington,J.W. and Jenkins,J.M. (1999) A longitudinal study of the relation between language and theory-of-mind development. *Dev. Psychol.* 35, 1311-1320
- 21 Wimmer,H. and Perner,J. (1983) Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103-128
- 22 Dennett,D.C. (1978) Beliefs about beliefs. *Behav. Brain Sci.* 4, 568-570
- 23 Wellman,H.M. *et al.* (2001) Meta-analysis of theory-of-mind development: The truth about false belief. *Child Dev.* 72(3), 655-684
- 24 Langdon,R. and Coltheart,M. (1999) Mentalising, schizotypy, and schizophrenia. *Cognition* 71, 43-71
- 25 Russell,J. (1996) *Agency Its Role in Mental Development*, Erlbaum
- 26 Fodor,J.A. (2000) *The mind doesn't work that way. The scope and limits of computational psychology.*, MIT Press
- 27 Surian,L. and Siegal,M. (2001) Sources of performance on theory of mind tasks in right hemisphere-damaged patients. *Brain Lang.* 78, 224-232
- 28 Channon,S. *et al.* (2005) Social cognition after head injury: Sarcasm and theory of mind. *Brain Lang.* 93, 123-134
- 29 Happe,F. *et al.* (1999) Acquired "theory of mind" impairments following stroke. *Cognition* 70, 211-240
- 30 Milders,M. *et al.* (2003) Neuropsychological impairments and changes in emotional and social behaviour following severe traumatic brain injury. *J. Clin. Exp. Neuropsychol.* 25, 157-172
- 31 Stone,V.E. *et al.* (2003) Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia* 41, 209-220
- 32 Winner,E. *et al.* (1998) Distinguishing lies from jokes: theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain Lang.* 62, 89-106
- 33 Varley,R. and Siegal,M. (2000) Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Curr. Biol.* 10, 723-726

- 34 Bird, C.M. *et al.* (2004) The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain* 127, 914-928
- 35 Stuss, D.T. *et al.* (1999) Dissociation of attentional processes in patients with focal frontal and posterior lesions. *Neuropsychologia* 37, 1005-1027
- 36 Peterson, D.M. and Riggs, K.J. (1999) Adaptive modelling and mindreading. *Mind. Lang* 14(1), 80-112
- 37 Zaitchick, D. (1990) When representation conflicts with reality: The preschooler's problem with false belief and "false" photographs. *Cognition* 35, 41-68
- 38 Perner, J. (1995) The many faces of belief: reflections on Fodor's and the child's theory of mind. *Cognition* 57, 241-269
- 39 Russell, J. *et al.* (1999) What do executive factors contribute to the failure on false belief tasks by children with autism? *J. Child Psychol. Psychiatry* 40, 859-868
- 40 Apperly, I. *et al.* (2005) Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence from an improved false photograph task with reduced language and executive demands. *J. Cogn. Neurosci.*, 71
- 41 Gordon, A.C. and Olson, D.R. (1998) The relation between acquisition of a theory of mind and the capacity to hold in mind. *J. Exp. Child Psychol.* 68, 70-83
- 42 Slaughter, V. (1998) Children's understanding of pictorial and mental representations. *Child Dev.* 69, 321-332
- 43 de Villiers, J.G. and 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
- 44 Carruthers, P. (2002) The cognitive functions of language. *Behav. Brain Sci.* 25, 657-674
- 45 Baron-Cohen, S. and Ring, H. (1994) A model of the mindreading system: Neuropsychological and neurobiological perspectives. In *Children's early understanding of the mind: Origins and development* (Lewis, C., ed), pp. 183-202, Lawrence Erlbaum
- 46 Gopnik, A. and Meltzoff, A. (1997) *Words Thoughts and Theories*, MIT press
- 47 Saxe, R. and Kanwisher, N. (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *NeuroImage* 19, 1835-1842
- 48 Saxe, R. and Wexler, A. (2005) Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia* In Press
- 49 Sabbagh, M.A. and Taylor, M. (2000) Neural correlates of theory-of-mind reasoning: an event-related potential study. *Psychol. Sci.* 11, 46-50

Figure 1. The event sequences in reality-known and reality-unknown false belief and false photograph tasks. *A) Reality-Known False Belief task.* Woman places object in one box. In her absence the man swaps the boxes. Participant has to predict where woman will look when she returns, while resisting interference from their knowledge of the true location. *B) Reality-Known False Photograph task.* Woman places object in one box. Photograph is taken and placed face-down. The man swaps the boxes. Participant has to judge where the object appears in the photograph, while resisting interference from their knowledge of the true location. *C) Reality-Unknown False Belief Task.* The participant does now know which box contains a hidden object – their task is to work this out from the a clue given by the woman. The woman looks in the box. In her absence, the man swaps the boxes. The woman indicates where she thinks the object is. Since the participant has to infer that the woman has a false belief *in order to locate the object*, knowledge of the object’s location cannot interfere with this belief reasoning. *D) Reality-Unknown False Photograph Task.* The participant does not know which box contains a hidden object - their task is to work this out from the photographic clue. A photograph is taken and placed face-down. The man swaps the boxes. The photograph is shown. Since the participant has to infer that the photograph is false *in order to locate the object*, knowledge of the object’s location cannot interfere with this reasoning.

Note. Warm-up trials establish the purpose of each task (either locating the object or predicting where the woman will search). By using video stimuli it is possible to administer test trials without language.

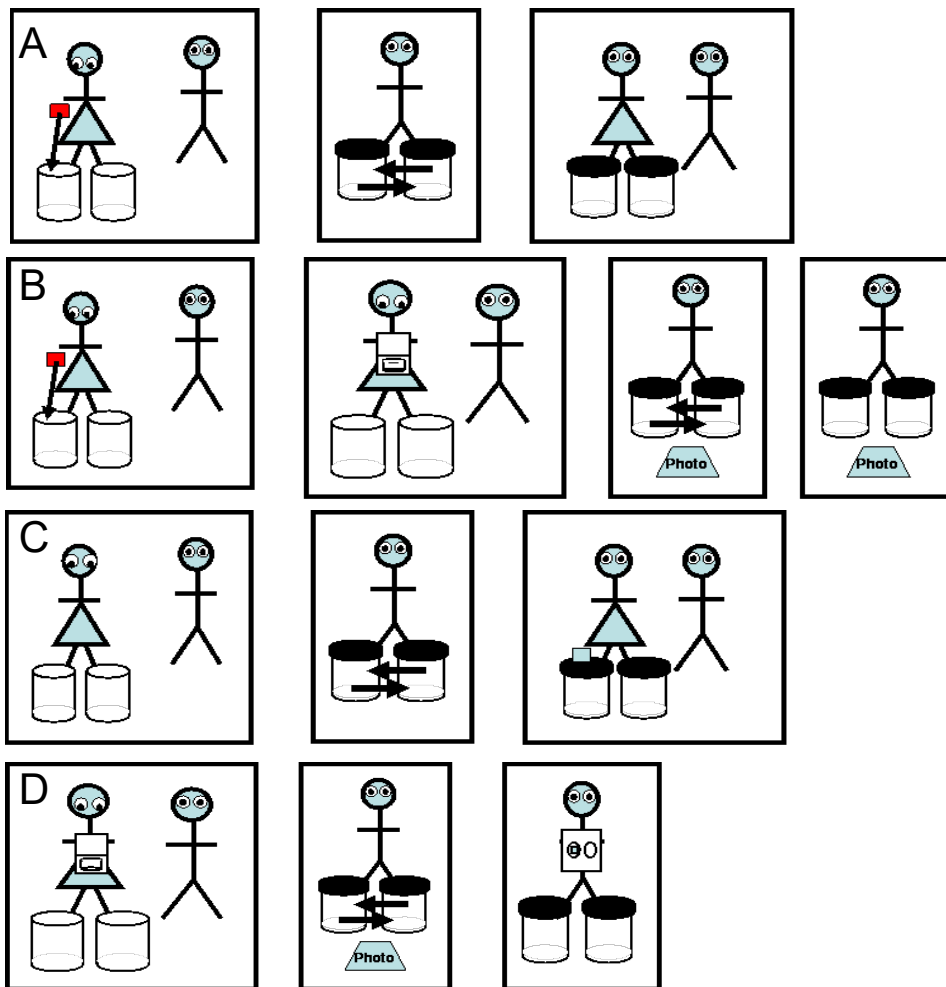


Figure 2. Schematic summary of our argument.

