Running head: REPRESENTATIONAL UNDERSTANDING IN CHILDREN

Non-Specificity and Theory of Mind: New Evidence From a Non-Verbal False Sign Task and Children With Autism Spectrum Disorders

Lai-Sang Iao* and Susan R. Leekam

Wales Autism Research Centre, School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, United Kingdom

*Corresponding author

E-mail: llsi@hku.hk

Telephone: +852 2859 0579

Present address: Centre for Communication Disorders, University of Hong Kong, 7/F Meng Wah Complex, University of Hong Kong, Pokfulam, Hong Kong

Acknowledgements

We thank the children, parents, head teachers and teaching staff who made this research possible. Thanks also go to Josef Perner, Ian Apperly, Helen McConachie and anonymous reviewers for helpful suggestions.

Please cite this paper as:

Iao, L.-S., & Leekam, S. R. (2014). Nonspecificity and theory of mind: New evidence from a nonverbal false-sign task and children with autism spectrum disorders. *Journal of Experimental Child Psychology, 122*, 1-20.

Doi: http://dx.doi.org/10.1016/j.jecp.2013.11.017

Non-Specificity and Theory of Mind: New Evidence From a Non-Verbal False Sign Task and Children With Autism Spectrum Disorders

Abstract

Understanding of false belief has long been considered to be a crucial aspect of theory of mind that can be explained by a domain-specific mechanism. We argue against this claim using new evidence from a non-verbal false representation task (false-sign task) with typically developing children and children with Autism Spectrum Disorders (ASD). Experiments 1 and 2 showed that typically developing children (mean age = 62.67 months) were equivalent in their performance across non-verbal and verbal forms of both the false-belief and false-sign tasks. Results for these two misrepresentation tasks differed from the results of an outdated representation task ("false" photograph task). Experiment 3 showed that children with ASD had difficulties with the false representation tasks and this could not be explained by executive functioning or language impairments. These findings support the view that children with ASD may not have a specific theory of mind deficit.

Keywords: Representational understanding; Theory of mind; False sign; Language; Executive function; Autism Spectrum Disorders

Non-Specificity and Theory of Mind: New Evidence From a Non-Verbal False Sign Task and Children With Autism Spectrum Disorders

It has long been claimed that there is a domain-specific cognitive mechanism which supports mental-related processing such as understanding of false beliefs (e.g., Baron-Cohen, 1995; Leslie, Friedman & German, 2004). In contrast to this claim, the domain-general hypothesis proposes a general conceptual development that accounts for both mental and non-mental processing as children come to understand false beliefs as representations of the world (e.g., Perner, 1991; Wellman, 1990). Others have argued that more general executive or language processing may explain children's false-belief performance (e.g., de Villiers, 2000; Russell, 1999). To date, the debate between domain-specificity and domain-generality of false-belief understanding is still ongoing (e.g., Cohen & German, 2010; Iao, Leekam, Perner, & McConachie, 2011). More importantly, false-belief performance continues to remain one of the key cognitive discriminators of Autism Spectrum Disorders (ASD; e.g., Begeer, Bernstein, van Wijhe, Scheeren, & Koot, 2012). While there is still a common assumption, even currently (e.g., Senju, Southgate, White, & Frith, 2009; Senju et al., 2010), that 'mentalising' or 'theory of mind' is the main problem in ASD, there are also suggestions that the problem for individuals with ASD may lie in executive and language processing (e.g., Pellicano, 2010; Paynter & Peterson, 2010). The current study critically tests these accounts regarding ASD and speaks to the long-standing debate concerning false-belief understanding.

Until recently, the debate has been tested empirically by comparing false-belief (FB) tasks (e.g., Wimmer & Perner, 1983) with non-mental tasks that share similar structural features, e.g., "false" photograph (FP) task (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007; Leekam & Perner, 1991; Leslie &

Thaiss, 1992; Zaitchik, 1990). The FB task itself involves a protagonist observing an object being placed in location A and subsequently moved to location B in the protagonist's absence. Children are then asked where the protagonist thinks the object is. In the FP task, the experimenter takes a photograph of an object in location A with a Polaroid camera, but then the object is moved to location B. Children are asked where the object is in the photograph.

While the FP and FB tasks appear to show structural similarity and equivalent performance in typically developing children, children with ASD failed the FB task but passed the FP task (Leekam & Perner, 1991; Leslie & Thaiss, 1992; see also Charman & Baron-Cohen, 1992). These results furthered Baron-Cohen, Leslie, and Frith's (1985) original proposal of a 'theory of mind' deficit in individuals with ASD, attributing specific difficulties to the processing of mental-related information, and offered evidence against the domain-general hypothesis. However, some researchers argued that these tasks present different conceptual demands. This is because a photograph is a *true* representation of the situation at the time the photograph was taken, whereas a false belief is a misrepresentation of whatever it is supposed to represent (e.g., Leekam & Perner, 1991). Thus, the FP task may not be an appropriate non-mental comparison to the FB task and whether individuals with ASD have a specific deficit in processing mental-related information needs further investigation.

In order to adequately test the domain-specificity debate, the false-sign (FS) task was subsequently devised to replace the FP task (Bowler, Briskman, Gurvidi, & Fornells-Ambrojo, 2005; see also Parkin, 1994, for an unpublished thesis). The logic behind this is that a sign, like a belief, represents what it is supposed to. In the FS task, children are shown a signpost that shows an object in location A but then the object is moved to location B. The signpost hence becomes a false sign. Children are asked where the signpost shows that the object is. It has been proposed that the FS task, relative to the FP task, is more comparable to the FB task in both behavioural studies with children (Leekam, Perner, Healey, & Sewell, 2008; Sabbagh, Moses, & Shiverick, 2006) and brain imaging studies with adults (Aichhorn et al., 2009; Perner, Aichhorn, Kronbichler, Staffen, & Ladurner, 2006). Training studies with children have also demonstrated that the FS and FB tasks are potentially transferable (Iao et al., 2011) whereas the FP and FB tasks are not (Slaughter, 1998). Further, children with ASD showed similar and associated performance on the FB and FS tasks (Bowler et al., 2005). These findings suggest that the FB and FS tasks share a developmental factor which is not shared by the FP task and that the suggestion that individuals with ASD are specifically impaired in mental-related processing may not be valid.

Yet while conceptual understanding of false representations might explain the correspondence between performance on the FB and FS tasks, it could be argued that the relation between them might be better explained by a third variable, notably other cognitive skills such as executive function or language demands. In terms of the inhibitory aspect of executive control, for example, these tasks involve several requirements. There is a requirement to disengage from knowledge about the salient reality and attend to the representation in question (cognitive inhibition) and to inhibit a prepotent response of pointing to the true location of the object (response inhibition). In terms of working memory there is also a requirement to maintain and process information in mind simultaneously. Indeed, based on the correlations found between performance on an executive function task (i.e., the bear/dragon task) and the FB/FS tasks but not the FP tasks, Sabbagh et al. (2006) suggested that the FP task may pose lesser executive demands than the FB and FS tasks. Furthermore, the FB and FS tasks are also linguistically demanding. Children have to comprehend a test question which involves syntactically complex structures (e.g., “Where does she think that X is?”; “Where does the sign show that X is?”). Such constructions within theory-of-mind tasks have been noted to create a challenge for young children (e.g., de Villiers & de Villiers, 2000). In contrast, the test question for the FP task, while also complex (“Where is X in the picture?”), is simpler. It is therefore possible that these skills in general cognitive processing explain the association between the FB and FS tasks.

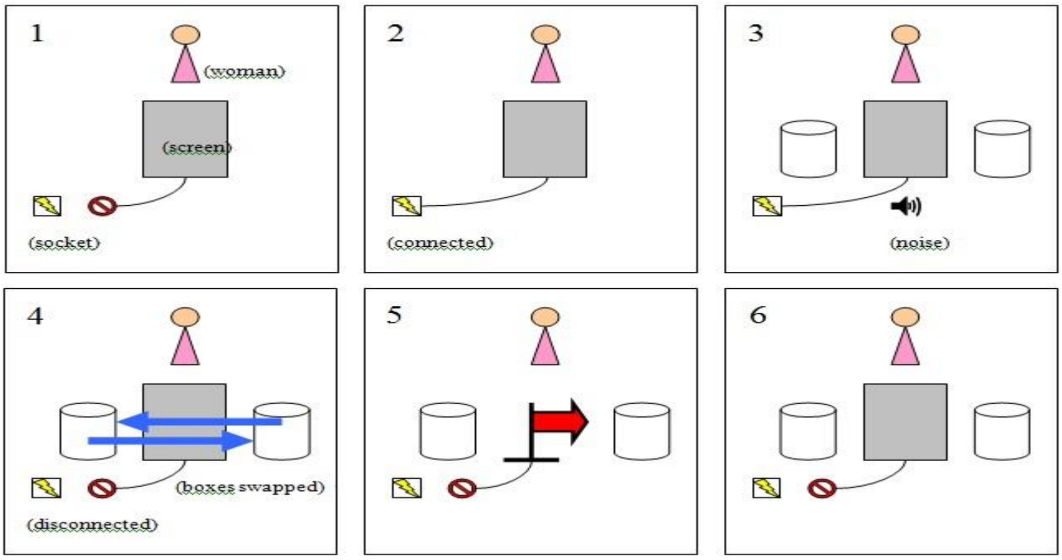
Individuals with ASD are widely documented to have abnormalities in executive function (e.g., Robinson, Goddard, & Dritschel, 2009) and language (e.g., Charman, Drew, Baird, & Baird, 2003). Recent research indicates that false-belief performance in ASD may be predicted by executive function (Pellicano, 2010) and syntactic language skills (Paynter & Peterson, 2010). It is thus possible that the consistently poor performance of children with ASD on the FB and FS tasks reflects their impairments in executive function and language. Although Bowler et al.’s (2005) finding was based on comparisons with two control groups which matched the ASD group on verbal mental age, it was measured by independent vocabulary tests which did not measure the exact incidental cognitive demands of the tasks. It is thus possible that the ASD and control groups were different in terms of their ability in meeting the incidental cognitive demands of the tasks. Robinson et al. (2009) also indicated that the impairments in executive function of children with ASD are independent of verbal and general intellectual abilities. Matching on verbal mental age and/or general intellectual ability does not control for the possible differences in executive function between the ASD and control groups. Therefore, it is crucial to investigate whether children with ASD have difficulties in understanding both mental and non-mental representations, independent of their deficits in language and executive function.

The main aim of the current study was to establish whether the association between understanding of mental and non-mental representations found in both typically developing children and children with ASD are best explained in terms of an underlying conceptual capacity or in terms of executive functioning and linguistic demands. We investigated this issue by employing established tasks and also devising a novel false representation task that disentangled the demands of executive function and language from representational understanding. The first task we employed was the FB task of Apperly, Samson, Chiavarino, and Humphreys (2004; adapted from a FB task originally devised by Call & Tomasello, 1999). This task was mainly *non-verbal* and *reality-unknown*, meaning that the real location of an object was not known by participants and therefore the requirement of cognitive inhibition was greatly reduced. In the task, a man hid an object in one of two boxes. A woman saw where he hid it while participants could not. For the *test* trial, the woman then left the room and the man swapped the boxes around, creating a false belief in the woman. When she returned, she briefly placed a marker on the box to indicate the location that she thought contained the object. Participants thus had to take her false belief into account to correctly find the object.

Another task we employed was the corresponding non-verbal reality-unknown FP task devised by Apperly et al. (2007). It followed the FB task above very closely except that the woman took a Polaroid photograph of the interior of the boxes (one contained an object), placed the photograph face-down in front of and mid-way between the boxes, and left without returning. The man swapped the boxes around and briefly showed participants the photograph which was a clue for them to find the object. According to Apperly et al., both of their FB and FP tasks made the same conceptual demand of representational understanding in the sense that both tasks required participants to consider a representation (belief or photograph) in relation to a current situation so as to figure out an object's current location. However, it is important to note that the photograph was "of" an outdated situation. In other words, the photograph acted as a clue, providing outdated information for participants to infer the current situation. In contrast, in Apperly et al.'s FB task, the woman's false belief was "about" the current situation. Her indication with the marker acted as a clue providing current but false information for participants to infer the current situation. Thus, Apperly et al.'s FP and FB tasks might still not be equivalent to each other.

The current study further devised a novel false non-mental representation task. This was a non-verbal reality-unknown FS task. In this task, a signpost was constructed with an electric plug attached. As a signpost, especially an electrically operated signpost, similar to traffic lights, is supposed to represent a current situation, it can become false when the electricity supply is disrupted and the situation changes. For example, an electrical signpost would represent an object's location as A by automatically turning its direction to point to A. However, if the electricity supply is disrupted, it cannot change its direction even though the object has been moved to B so it keeps representing the object's location as A. In this way, the false signpost matches with the woman's indication in Apperly et al.'s (2004) FB task, both of which act as clues providing current but false information for participants to infer the current situations. However, the signpost is non-mentally (electrically) operated whereas the woman's indication in Apperly et al.'s FB task is mentally generated.

The FS task involved initial phases in which participants learned how a signpost worked with its electric plug connected to an electricity supply to indicate an object's location. They also noticed that the signpost could be blocked from view by a screen and the object was hid in one of two boxes. The task then followed in steps (see Figure 1). A woman put the signpost's plug into a socket and then placed the two boxes on each side of the screen. A mechanical noise was heard (meaning the signpost was turning behind screen). It is important to note that the mechanical noise itself did not indicate direction, only the pointing of the signpost served this function. The signpost was then deprived of its electricity supply by the woman removing the plug, and it became a false sign when she swapped the boxes around. She helped participants to find the object by briefly taking away the screen to show the signpost. Participants thus had to take the false sign into account in order to find the object. Appendix A shows the key stages of the test trials of these non-verbal reality-unknown FB, FP, and FS tasks. Appendix B presents the analogy between the components involved in the FB and FS tasks.



False sign test trial. (1) Signpost was blocked from view by a screen but its electric plug was visible. (2) Woman connected signpost's plug to electricity supply. (3) Woman placed two boxes (one contained object) on each side of screen and mechanical noise was heard (meaning signpost was turning). (4) Woman disconnected signpost from electricity supply and swapped boxes so signpost became false. (5) Woman displayed signpost. (6) Signpost was blocked by screen again and participant was asked to identify object's location.

Figure 1. Event sequence of the test trial of the novel false sign task.

With this newly devised FS task, the current series of experiments re-examined the non-specificity claim of theory of mind. The overall question was whether the equivalence between tasks testing mental and non-mental representations might be explained in terms of equivalence in representational understanding or in terms of other incidental cognitive demands of language and executive function. Experiment 1 first tested the new FS task against Apperly et al.'s (2004, 2007) FB and FP tasks in a group of typically developing children. In this experiment, the tasks were non-verbal and reality-unknown so the demands of language and cognitive inhibition were low. If the association between the standard *verbal reality-known* false-representation tasks was replicated in this *non-verbal reality-unknown* context, it would suggest that language and cognitive inhibition did not play an important role in the association. Experiment 2 aimed to further replicate the association between the FB and FS tasks using both verbal and non-verbal tasks within one experiment; the standard *verbal reality-known* and *non-verbal reality-unknown* versions. Experiment 3 included children with ASD, in addition to typically developing children. Using the same tasks as Experiment 1, Experiment 3 aimed to specify whether children with ASD have difficulties in understanding representations, independent of other cognitive deficits. Ethical approval was obtained from Department of Psychology University ethics committees and parental informed consent was obtained before testing.

Experiment 1

Experiment 1 provided the very first test of the novel non-verbal reality-unknown FS task, investigating how it worked relative to Apperly et al.'s (2004, 2007) FB and FP tasks in typically developing children aged 3 – 7 years. As Apperly et al.'s tasks have never been used with young typically developing children in any published study, Experiment 1 also extended the applicability of their tasks.[1] If children's performance on these *non-verbal reality-unknown* tasks was similar to that of the standard *verbal reality-unknown* tasks, the associations previously found between the standard tasks would be replicated in this non-verbal reality-unknown context. This replication would suggest that the developmental factor that was shared between the FB, FS and FP tasks may not be the general cognitive skills of language and cognitive inhibition given that the demands of language and cognitive inhibition were greatly reduced in the non-verbal reality-unknown context.

As the *test* trial of each task involved both representational understanding and incidental executive demands, it was essential to ensure that children could meet those incidental executive demands which were tested with two *control* trials. This assurance was achieved if children's performance on these control trials was significantly better than that on the test trials. Each task also contained two *filler* trials which demanded even less executive skills than the control trials. A relatively good performance on these filler trials would reflect that children had paid attention and had not been guessing throughout the tasks. Given that the correct answers for four of the five trials in each task required children to indicate the opposite box from the one indicated by the woman, photograph or signpost, children may have adopted an incorrect strategy of always pointing to the opposite box. To exclude this alternative explanation, performance on the *true representation filler* trial whose correct answer was the box indicated by the woman, photograph or signpost was examined.

---

[1] The youngest mean age that Apperly et al.'s (2004) FB task has been established is 13 years 3 months with a mean verbal mental age of 6 years 10 months in a sample of atypically developing children (including children with fragile X syndrome and intellectual disability; Grant, Apperly, & Oliver, 2007). Thus, children aged below 7 years and over 3 years were recruited.

Method

Participants.

In total, 20 children aged 41 – 58 months (*M* = 49.90 months, *SD* = 5.95) and twenty children aged 60 – 83 months (*M* = 71.65 months, *SD* = 7.75) were recruited from two schools in North East England and South East Wales.[2] The sample had a mean verbal mental age (VMA) of 68.93 months (range = 39 – 100 months; *SD* = 16.30), which was calculated using the British Picture Vocabulary Scale - Second Edition (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997). All children were White British except five who were from minority ethnic backgrounds. The populations of the schools were generally of low socio-economic status.

Design.

This was a mixed design, testing younger versus older children with the non-verbal reality-unknown FB, FP and FS tasks. Each of the three tasks consisted of one test trial, two control trials and two filler trials. Children were tested in three sessions at one- to two-week intervals. In each session, children were tested on all three tasks. The presentation order of the three tasks was counterbalanced across children and sessions whilst the order of the five trials within each task was randomised in each of the three sessions.

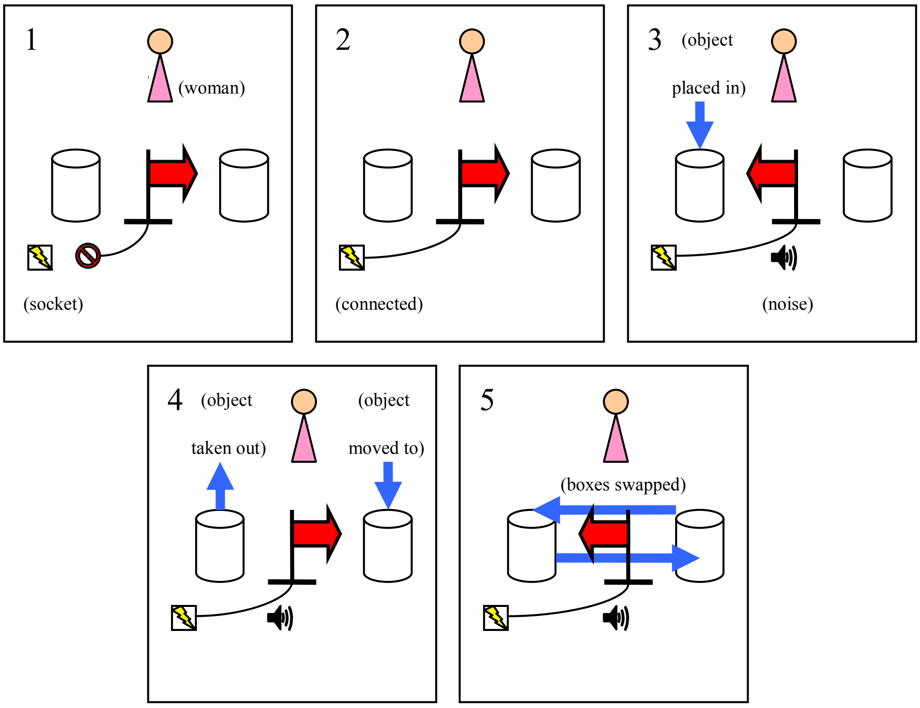
Materials and procedure.

Children were tested individually in a room within their own school by an experimenter. The tasks were video based and were presented on a laptop computer using PowerPoint software. The original FB and FP tasks were borrowed with kind permission from Apperly et al. (2004, 2007). As their FB and FP tasks were

---

[2] Half of the younger children and half of the older children were from each school and no significant difference in children's performance was found between the schools (*p*s > .05) although they were from different regions of the United Kingdom.

originally designed for adults, several adaptations have been made to accommodate these tasks and the novel FS task in a single experiment for children. First, children were given short training video clips, each of which illustrated how a Polaroid camera or a signpost worked, at the beginning of the FP and FS tasks respectively. The training video for the FP task showed a man who pressed the button of a Polaroid camera in front of some flowers. A picture then came out of the camera which eventually developed to show an accurate image of the flowers (procedure adapted from Zaitchik, 1990). Figure 2 presents the event sequence of the training video of the FS task. As children by the age of 3 years already understand the relationship between seeing and knowing (e.g., Pratt & Bryant, 1990), no training was required for the FB task. Second, calls for attention (e.g., calling children by their names) and encouraging words (e.g., “Well done!”) were used in each of the FB, FP and FS tasks. Third, a 2-minute break was given between each of the tasks.

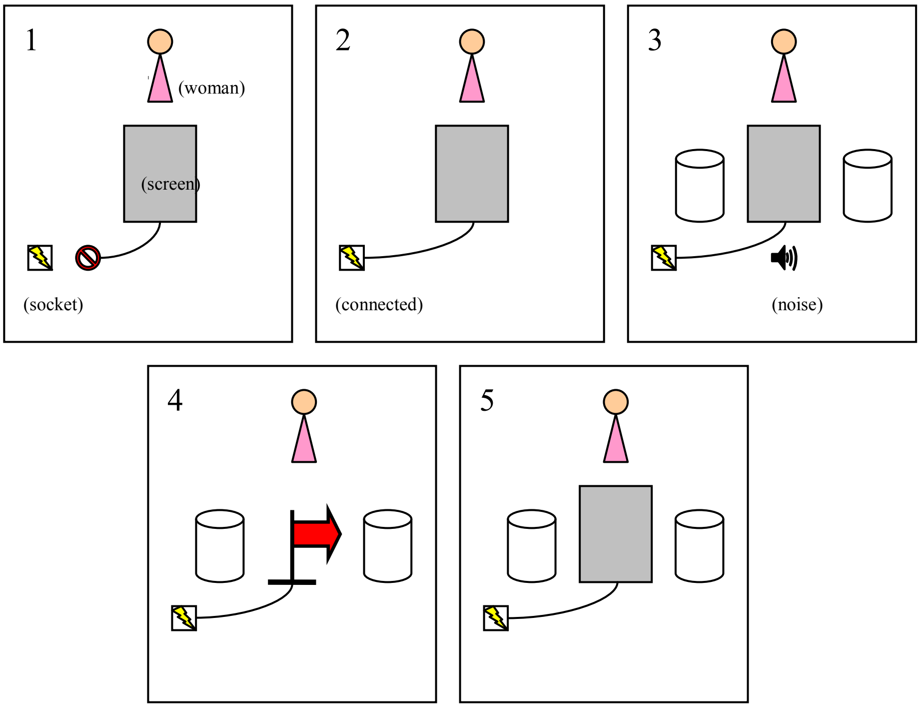


Training. (1) Signpost with an electric plug was not yet connected to electricity supply. (2) Signpost was connected to electricity supply by woman. (3) Object was put in one box and signpost turned, indicating object's location. Mechanical noise was always heard while signpost was turning. (4) Object was then moved to the other box and signpost turned accordingly (noise was heard). (5) Boxes were swapped with object inside one of them and signpost turned accordingly (noise was heard).

Figure 2. Event sequence in the training of the novel false sign task.

Before each task started, the principles of each task were explained verbally with still frames from the video, e.g., "We will play a hiding-and-finding game. Your job is to find a block. Look, the woman connects the sign. She puts its electric plug into the socket. Then she hides the block here. Now the sign turns to help you find the block." Warm-up trials were then given (e.g., see Figure 3 for the FS task). In each task, a woman provided a clue for the children to locate the block. For the FB task, a woman placed a marker on top of the box she thought contained the block. For the FP task, a woman took a Polaroid photograph showing the interior of the two boxes. For the FS task, a woman revealed a signpost which automatically turned to show the location of the block. Corrective feedback (e.g., "No, the block is in there.")

was provided as necessary. Two consecutive correct responses in the warm-up phase were required from children before they entered the test phase of the relevant task. None of the children required more than four warm-up trials to achieve two consecutive correct responses.



Warm-up trial. (1) Signpost was blocked from participant's view by screen but its electric plug was visible. (2) Signpost was connected to electricity supply. (3) Two boxes (one contained object) were placed on the sides of the screen and mechanical noise was heard. (4) Screen was removed and signpost was shown. (5) Screen was replaced and participant was asked to identifv obiect's location.

Figure 3. Event sequence of the warm-up trial of the novel false sign task.

The test phase involved one *false representation* test trial, two control trials (*working memory* and *response inhibition*), and two filler trials (*true representation* and *clue confirmation*) in each of the three tasks (see Table 1 for the FS task; Appendix C and D for the FB and FP tasks). Neither the control nor filler trials required an understanding of representation. However, the control trials involved the same incidental demands of (1) holding in mind the events that happened in the video while working out the object's location (i.e., working memory) and (2) inhibiting the tendency of pointing to the box indicated (i.e., response inhibition) as the test trials. Feedback which showed the interior of the two boxes was always presented at the end of each trial after children had responded. Two more testing sessions followed the same procedure but without training and warm-up phases. Each task started with an explanation of its principle, followed by the test phase of the relevant task.

Table 1

Key Stages of the Test, Control and Filler Trials of the Novel False Sign Task

<table>
<tr><th>False representation test trial</th><th>Working memory control trial</th><th>Response inhibition control trial</th><th>True representation filler trial</th><th>Clue confirmation filler trial</th></tr>
<tr><td colspan="5">A woman presented with an electrical signpost covered by a screen but its plug was visible. She put the plug into a socket.</td></tr>
<tr><td colspan="5">She placed two boxes on a table and the signpost turned with a mechanical noise behind a screen. The signpost indicating the location of an object was established.</td></tr>
<tr><td>She took the plug out.</td><td>The signpost was revealed and then covered.</td><td colspan="2">She took the plug out.</td><td>The signpost was revealed and then covered.</td></tr>
<tr><td>She swapped the boxes.</td><td>She took the plug out.</td><td>She took the object out of one box and put it into another box.</td><td>She held the boxes up and down vertically.</td><td>She took the plug out.</td></tr>
<tr><td>The signpost was revealed and then covered.</td><td>She swapped the boxes.</td><td colspan="2">The signpost was revealed and then covered.</td><td>She took the object out of the box indicated by the signpost and put the object into another box.</td></tr>
</table>

Results

Figure 4 shows the mean scores for the test and control trials of the three tasks. Older participants performed above chance on all the trials, *t*s(19) > 2.60, *p*s < .05, Cohen's *d*s > 1.19. However, younger participants performed below chance on the three test trials, *t*s(19) < -2.40, *p*s < .05, Cohen's *d*s > -1.10, but above chance on the control trials of all three tasks, *t*s(19) > 3.94, *p*s < .01, Cohen's *d*s > 1.81, except the working memory control trials of the FB and FS tasks on which they performed at chance level, *t*s(19) < .91, *p*s > .37.

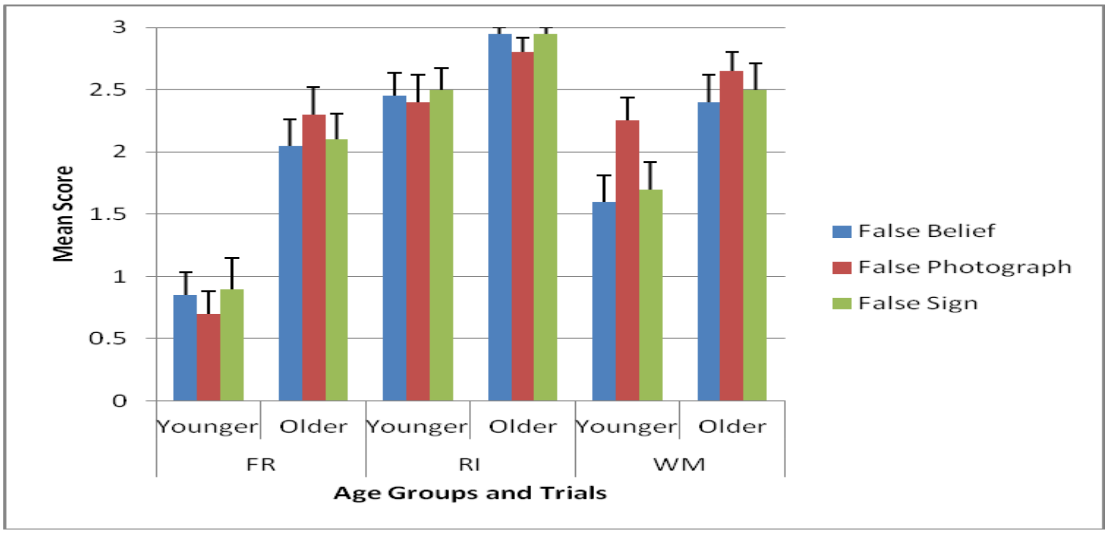


Figure 4. Mean Scores on Each Test and Control Trial Type of the False Belief, False Photograph and False Sign Tasks for Experiment 1. FR = False Representation Test Trial; RI = Response Inhibition Control Trial; WM = Working Memory Control Trial

For both test trials and response inhibition control trials, 2 (age) X 3 (task) ANOVAs revealed main effects for age, *F*s(1, 38) > 6.37, *p*s < .05, η_p^2s > .14, but no effect for task and no interaction, *F*s < 1.40, *p*s > .25. For the working memory control trials, there were significant main effects of age and task, *F*(1, 38) = 8.36, *p* < .01, η_p^2 = .18 and *F*(2, 76) = 5.23, *p* < .01, η_p^2 = .12 respectively, but no interaction, *F*(2, 76) = 1.25, *p* = .29. Performance on the working memory control trials of the FB and FS tasks was found to be worse than that of the FP task, *p* < .01 and *p* = .07

respectively, but no significant difference was found between the working memory control trials of the FB and FS tasks, $p = 1$. Table 2 showed significant differences between performance on the test trials, response inhibition control trials, and working memory control trials of the three tasks, suggesting that children could meet the incidental executive demands of the tasks.

Table 2

Repeated Measures ANOVAs and Post Hoc Tests Examining Performance Differences Between Test and Control Trials

	F	p	η_p^2	FR - RI	FR - WM	RI - WM
False Belief	27.87	< .001	.42	$p < .001$	$p < .05$	$p < .01$
False Sign	23.74	< .001	.38	$p < .001$	$p < .05$	$p < .01$
False Photo	19.08	< .001	.33	$p < .001$	$p < .001$	$p = 1$

Note. FR = False Representation Test Trial; RI = Response Inhibition Control Trial; WM = Working Memory Control Trial

The test trials of the three tasks were highly correlated, rs(40) > .67, ps < .001. Each of the three test trials was also correlated with VMA, rs(40) > .36, ps < .05. Given that the three test trials were correlated with each other and each of them were correlated with VMA, the third test trial (e.g., the FS test trial) and VMA had to be controlled for in order to have a purer measure of the correlations between any two of the test trials (e.g., the FB and FP test trials). Table 3 showed the correlations between any two of the test trials.

Table 3

Bivariate VMA-Controlled/Third-Task-Controlled Correlations and [VMA-and-Third-Task-Controlled Correlations] (With *N*) Between the False Belief, False Photograph and False Sign Tasks

	False Belief	False Photograph	False Sign
False Belief	-	.61***/.28 (37)	.71***/.55*** (37)
False Photograph	[.26] (35)	-	.66***/.42** (37)
False Sign	[.52**] (35)	[.34*] (35)	-

Note. ***$p < .001$, **$p < .01$, *$p < .05$

As feedback and encouraging words were provided, we also examined whether children improved performance on the test trials across the three sessions. Older participants did not improve on the FB and FS tasks, χ^2s(2, N = 20) < 3.82, *p*s > .24, but improved on the FP task, $\chi^2(2, N = 20) = 12.29$, $p < .01$. Further analyses showed that older participants' performance on the FP task improved from Session 1 to 2 and from Session 1 to 3, Wilcoxon *Z*s < -2.45, *p*s < .05, *r*s < -.39, but no improvement was shown from Session 2 to 3, Wilcoxon $Z = -1$, $p = 1$. Younger participants did not improve at all, χ^2s(2, N = 20) < 3.43, *p*s > .23.

Ceiling effects were found for both clue confirmation (100% correct) and true representation filler trials (88.33 – 95% correct) of all three tasks. Out of the 178 correct responses to the test trials of all three tasks, there were only 15 which were paired with an incorrect response to the true representation filler trial in the same task. These results suggested that participants were not likely to be using the strategy of always pointing to the opposite box.

Discussion

The main goal of Experiment 1 was to assess whether the novel FS task is a good comparison task to Apperly et al.'s (2004, 2007) non-verbal FB and FP task, providing the very first test of the novel FS task. The results showed that with respect to incidental working memory demands, the FS task was more equivalent to the FB task compared to the FP task. Comparisons between the working memory and response inhibition control trials also reflected that the working memory demand was higher than the response inhibition demand in both of the FB and FS tasks but not in the FP task. These findings extend Sabbagh et al.'s (2006) correlational findings. Instead of testing executive functioning with the bear/dragon task and correlating it with the FB, FS and FP tasks, the current experiment provided more solid evidence that executive functioning is *incidentally* required to a greater extent when reasoning about false beliefs and false signs than "false" photographs. Another difference between the tasks was that older participants significantly improved their performance on the false representation test trials of the FP task but not the FB and FS tasks across the three sessions, suggesting that the FP task was more subject to a practice effect for older participants. Despite these two differences, performance on the test trials of the tasks was not significantly different from but was correlated with each other. However, it is still worth noting that the correlation between the FB and FP tasks became non-significant when performance on the FS task and VMA were controlled. This finding replicated Leekam et al. (2008) who found similar results using standard verbal FB, FP and FS tasks.

Although both of Apperly et al.'s (2004, 2007) tasks have not been used in such a young age range in any published study, the current study showed that their tasks and the novel FS task were able to reveal a developmental change of false-belief, false-sign and pictorial understandings: Children before the age of 5 years performed below chance while those above 5 performed above chance. Call and Tomasello (1999) also found a similar developmental change in children below versus above 5 years old. Given that the usual age when children pass the standard *verbal reality-known* FB and FP tasks in which the real location of an object was known by participants (e.g., it was visibly moved from location A to B; Wellman, Cross, & Watson, 2001; Zaitchik, 1990) was 4 years old, the *non-verbal reality-unknown* tasks appeared to be more difficult than the standard *verbal reality-known* ones. However, Call and Tomasello demonstrated that the two versions of the FB tasks were in fact comparable. Following Call and Tomasello's procedure, Experiment 2 tested children with the two versions of the FB and FS tasks in order to replicate the current findings and validate the novel FS task.

Experiment 2

Experiment 2 compared a standard *verbal reality-known* version against the *non-verbal reality-unknown* version of the FB and FS tasks in a different group of typically developing children. The aim was twofold. First, to further replicate the associations found between the FB and FS tasks in both verbal and non-verbal versions within a single experiment. If a replication was illustrated, it would provide further evidence that language and cognitive inhibition do not play a role in the association between the two tasks. Second, the experiment aimed to validate the novel FS task against its standard version. If a validation was shown, it would suggest that the new FS task was virtually the same as the standard FS task but could further be employed to test populations who have language and cognitive inhibition problems such as children with ASD.

In order to accomplish these two aims, Experiment 2 followed Call and Tomasello's (1999) design of administering a standard version of the task in the context of the non-verbal reality-unknown version of the task series. Thus, the FB and FS tasks could be compared in both versions within a single experiment. Based on the findings of previous studies and of Experiment 1, no significant difference but significant correlations were expected between the FB and FS tasks in across both versions. Further, given the previously reported association between the non-verbal and verbal versions of the FB task (Call & Tomasello, 1999), a similar finding was expected in the current experiment. However, it was unknown whether an association would be shown between the two versions of the FS task, given that this was the first time that this association was tested.

Method

Participants.

In total, 13 children aged 39 – 62 months (*M* = 52.69 months, *SD* = 7.54) and thirteen older children aged 63 – 88 months (*M* = 78.38 months, *SD* = 8.93) were recruited from two schools in South East Wales. None of the children had taken part in Experiment 1. The sample had a mean VMA of 72.46 months (range = 37 – 100 months; *SD* = 17.62). All children were White British except four who were from minority ethnic backgrounds. The populations of the schools were generally of low socio-economic status.

Design.

This was a mixed design, testing younger versus older children with both versions of the FB and FS tasks. Children were tested with both versions of the tasks in each of three sessions at one-week intervals. Following the exact procedure of previous studies, the verbal version of the tasks consisted of a test question and a memory question; whereas the non-verbal version of the tasks consisted of the five trials as described in Experiment 1.

Materials and procedure.

The same non-verbal FB and FS tasks from Experiment 1 were used. The only difference between Experiment 1 and 2 was a verbal version of the tasks was added in the context of the non-verbal task series. For example, the verbal FB task was either presented consecutively after the warm-up trials or after the series of the test, control and filler trials of the non-verbal FB task. In this way, the presentation order of the verbal and non-verbal versions of the tasks was counterbalanced across children.

Following Call and Tomasello (1999), the verbal FB task was identical to the response inhibition control trial of the non-verbal FB task except that children were asked a test question, “Where does the woman think the block is?”, when the woman returned to the room. After children responded, they were asked a memory question, “Where is the block really?” For the verbal FS task, children saw the woman presented with the signpost covered by a screen but its plug was visible. She put the plug in the socket and placed two boxes on the table. The mechanical noise produced by the turning signpost was heard. She then revealed the signpost, took the plug out, and displaced the object visibly from one box to the other. Children were asked a test question, “Where does the sign show the block is?”, and the same memory question as above. Verbal feedback was provided, e.g., “No, the sign shows the block is in this box.”

Results

As previous studies using the standard verbal FB task considered participants as passers only if they passed both of the test and memory questions, the same criterion was employed for the current verbal FB and FS tasks. The mean scores for the tasks were shown in Figure 5. Performance on the control and filler trials of the non-verbal tasks was the same as that found in Experiment 1. We thus reported only the performance of the verbal tasks and the false representation test trials of the non-verbal tasks which was indeed the main focus of the current experiment.

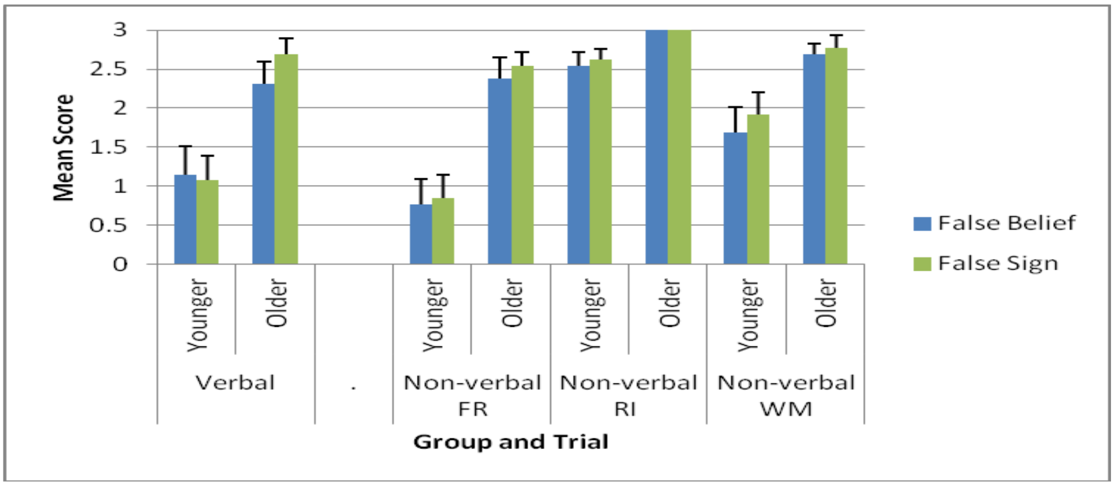


Figure 5. Mean Scores on Each Test and Control Trial Type of the False Belief and False Sign Tasks for Experiment 2. FR = False Representation Test Trial; RI = Response Inhibition Control Trial; WM = Working Memory Control Trial

Older children performed consistently above chance on both verbal and non-verbal FB and FS tasks, ts(12) > 2.82, ps < .05, Cohen's ds > 1.63. However, younger children performed below chance on the non-verbal ones, ts(12) < -2.21, ps < .05, Cohen's ds > -1.28, but at chance on the verbal ones, ts(12) > -1.37, ps > .20. A 2 (age) X 2 (version) X 2 (task) ANOVA revealed a main effect for age, $F(1, 24) = 22.81$, $p < .001$, $\eta_p^2 = .49$, but no effects for version and task, and no interactions, Fs(1, 24) < 2.13, ps > .16.

VMA correlated significantly with both versions of the tasks, rs(26) > .60, ps < .01. Performance on the FB task correlated with that of the FS task, even after controlling for VMA, $pr(23) = .74$, $p < .001$ for the verbal version and $pr(23) = .79$, $p < .001$ for the non-verbal version. Significant correlations were also found between the verbal and non-verbal FB tasks, $r(26) = .61$, $p < .001$, and between the verbal and non-verbal FS tasks, $r(26) = .58$, $p < .01$. However, these correlations did not remain significant after controlling for VMA, prs(23) < .34, ps > .10. This finding seemed to be driven by the younger children. When we investigated the older children only, the correlations between the verbal and non-verbal tasks remained significant after controlling for VMA, *pr*s (10) > .60, *p*s < .05. Across the three sessions, no significant improvement was shown by both older and younger children, χ^2s(2, *N* = 13) < 5.20, *p*s > .17 for both non-verbal FB and FS tasks, and χ^2s(2, *N* = 13) < 6, *p*s > .07 for both verbal FB and FS tasks. Thus, children's performance could not be explained by learning through feedback across sessions.

Discussion

Experiment 2 replicated the finding of Experiment 1 that the equivalence between the FB and FS tasks holds in both verbal reality-known and non-verbal reality-unknown versions. Moreover, the new non-verbal FS task was validated by the findings that performance on the non-verbal FS task did not differ from that on the verbal FS task and that the two tasks were correlated (*r* = .58). It was not clear why the correlations between the verbal and non-verbal tasks remained significant for the older children only after controlling for VMA. However, it was reasonable that children reached certain VMA and started passing the verbal tasks as well as the non-verbal tasks and thus VMA explained most of the variance between the tasks for the younger children. Further investigation is needed to clarify this speculation.

To our knowledge, none of the previous studies that have employed the same verbal and non-verbal FB tasks have measured participants' VMA (Call & Tomasello, 1999; Figueras-Costa & Harris, 2001) nor have they tested the correlations between the tasks after controlling for VMA. Given that previous studies have used the non-verbal FB task as a counterpart of the verbal FB task without measuring and controlling for VMA, we suggest that the new non-verbal FS task should still be appropriate for assessing false-sign understanding in children.

Experiment 3

Experiment 3 investigated whether children with ASD have difficulties in understanding representations which could not be accounted for by their deficits in language and executive function. The non-verbal reality-unknown FP task used in Experiment 1 was included in the current experiment for the following reasons. First, the inclusion of this FP task allowed an evaluation of whether children with ASD's performance on this FP task replicated that on the standard FP task. Second, if children with ASD performed relatively well on the non-verbal FP task, it would serve as a control task. Together with the control trials of the FB and FS tasks, a validation of equal executive abilities in meeting the incidental demands of the tasks between children with ASD and comparison children would be ensured. If children with and without ASD did not differ in their ability to meet the incidental demands of the tasks but significantly differed in their performance on the false representation test trials of the FB task only, it would clearly suggest that children with ASD suffer a specific impairment in understanding mental representations.

Typically developing children were recruited as comparison children to match with children with ASD for the following reason. If an association between the FB and FS tasks was replicated in typically developing children and also found in children with ASD, it would suggest that children with ASD were no different from typically developing children in that both groups process mental and non-mental representations by means of an underlying conceptual capacity for representational understanding. On the contrary, if the association was found in typically developing children but not children with ASD because they performed selectively worse on the FB task than on the FP and FS tasks, it would provide evidence for a specific deficit in understanding mental representations in ASD.

Method

Participants.

In total, 18 children with a diagnosis of ASD were recruited from two special schools and one resource unit of a primary school in South East Wales. All had a community multidisciplinary team assessment leading to a best estimate clinical diagnosis of an ASD (including autism and Asperger syndrome) according to DSM-IV-TR (American Psychiatric Association, 2000) and ICD-10 (World Health Organisation, 1993) criteria. To further assess participants' ASD symptoms, their parents were requested to complete the lifetime version of the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) which has a cut-off point of 15 for ASD. All participants were justified to be included in the ASD group of the experiment (range of score: 15 – 38; $M = 26.50$; $SD = 6.34$).

A comparison group of 18 typically developing children were recruited from two schools in South East Wales. The ASD and comparison groups were matched in terms of VMA and non-verbal intelligence quotient (NVIQ), as measured with the BPVS-II and the Brief Intelligence Quotient composite of the Leiter International Performance Scale-Revised (Roid & Miller, 1997) respectively.[3] Participant characteristics are presented in Table 4. The two groups did not differ in VMA, $t(34) = -.35$, $p = .73$, and NVIQ, $t(27.98) = -1.50$, $p = .15$, but differed in chronological age (CA), $t(27.40) = 5.59$, $p < .001$, Cohen's $d = 2.14$.[4]

---

[3] Using VMA, there were 12 children who were matched exactly (same VMA score), 10 children who were matched within 5 months of age, and 14 children who could not be matched individually. With NVIQ, there were 10 children who were matched exactly (same NVIQ score), 14 children who were matched within 6 points of NVIQ (less than 0.50 *SD*), and 12 children who could not be matched individually.

[4] The ASD group was older because 13 children with ASD above the age of the oldest typically developing children were included in order to match for VMA and NVIQ.

Table 4

Characteristics of the Experimental and Comparison Groups

Characteristic	ASD (n = 18, 16 males)			Comparison (n = 18, 8 males)		
	M	*SD*	Range	*M*	*SD*	Range
CA	104.89	19.95	70 - 133	74.44	11.66	52 - 88
VMA	75.56	20.70	51 - 119	77.61	13.84	53 - 98
NVIQ	97.89	19.93	65 - 133	106.11	12.06	91 - 135

Note. ASD = Autism Spectrum Disorder; CA = Chronological age in months; VMA = Verbal mental age in months; NVIQ = Non-verbal intelligence quotient

Design, materials and procedure.

This experiment used exactly the same design, materials and procedure as in Experiment 1, except that the between-participants factor was group (children with ASD versus typical development) rather than age.

Results

Performance on the control and filler trials of the three tasks was not significantly different between the two groups and was the same as that found in Experiment 1. This suggested that both groups were capable of meeting incidental cognitive demands of the tasks and that their capabilities were comparable even though they were not matched on CA. Moreover, both groups were not likely to be using the strategy of always pointing to the opposite box from the one indicated by the clues. To avoid redundancy, we report only children's performance on the false representation test trials of the three tasks which was indeed the main focus of the current experiment.

Participants' mean scores for the tasks are presented in Figure 6. Children with ASD's performance on the FB and FS tasks was not significantly different from chance, t(17) = .65, p = .52 and t(17) = 1.46, $p = .16$ respectively. However, their performance on the FP task was significantly above chance, $t(17) = 3.69$, $p < .01$, Cohen's d = 1.79. On the contrary, performance on all three tasks was significantly above chance for the comparison group, ts(17) > 6.76, ps < .001, Cohen's ds > 3.28. A 2 (group) X 3 (task) ANOVA revealed significant main effects of group and task, $F(1, 34) = 9.30$, $p < .01$, $\eta_p^2 = .21$ and $F(2, 68) = 3.12$, $p = .05$, $\eta_p^2 = .08$ respectively, but no significant interaction, $F(2, 68) = 2.43$, $p = .10$. Performance on the FP task was marginally better than that on the FB task, $p = .06$; whereas performance on the FS task did not differ from those on the FB and FP tasks, $p = .57$ and $p = .70$ respectively. Planned contrasts revealed that the groups were different on the FB and FS tasks, $t(34) = -3.11$, $p < .01$, Cohen's d = -1.07 and $t(34) = -2.79$, $p < .01$, Cohen's d = -0.96 respectively, but not on the FP task, $t(34) = -1.62$, $p = .11$.

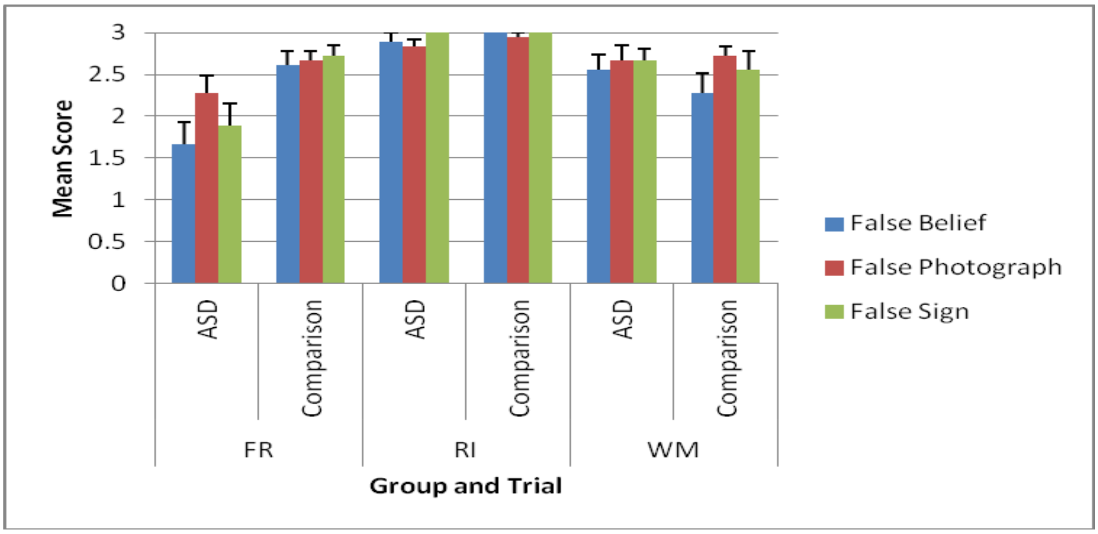


Figure 6. Mean Scores on Each Test and Control Trial Type of the False Belief, False Photograph and False Sign Tasks for Experiment 3. FR = False Representation Test Trial; RI = Response Inhibition Control Trial; WM = Working Memory Control Trial; ASD = Autism Spectrum Disorder

The test trials of the three tasks was correlated with VMA, rs(36) > .38, ps < .05, but not with NVIQ, rs(36) < .30, ps > .05. When VMA was controlled, the correlations between the test trials on the three tasks remained significant, prs(33) > .47, ps < .01. Looking at the groups separately, performance on the three tasks was also significantly correlated, rs $(18) > .50$, ps $< .05$ for the ASD group and rs $(18) > .46$, ps $< .05$ for the comparison group. However, only the correlation between the FB and FS tasks remained significant when performance on the third task (i.e., the FP task in this case) was controlled, $pr(15) = .58$, $p < .05$ for the ASD group and $pr(15) = .48$, $p < .05$ for the comparison group.

Children with ASD improved their performance across the three sessions on the FB task, $\chi^2(2, N = 18) = 10.89$, $p < .01$, the FP task, $\chi^2(2, N = 18) = 7.75$, $p < .05$, and the FS task, $\chi^2(2, N = 18) = 10.75$, $p < .01$. Improvement was significant only from Session 1 to 2 and 3 on the FB and FS tasks, Wilcoxon Zs < -2.45, ps $< .05$, rs $< -.41$, but marginal on the FP task, Wilcoxon Zs < -2.12, ps $= .07$ and $.06$, rs $< -.35$. On the contrary, typically developing children did not improve their performance on the FB and FS tasks across the three sessions, $\chi^2(2, N = 18) = 2.80$, $p = .40$ and $\chi^2(2, N = 18) = 6.50$, $p = .07$ respectively, but improved on the FP task, $\chi^2(2, N = 18) = 12$, $p < .01$. Improvement on the FP task was significant from Session 1 to 2 and 3 only, Wilcoxon Zs $= -2.45$, ps $= .05$, rs $< -.41$.

Discussion

Experiment 3 investigated whether children with ASD have difficulties in understanding representations, independent of other impairments in language and executive function. Results showed that children with ASD performed worse on the false representation test trials of the FB and FS tasks relative to typically developing children. When performance on the FP task was partialled out, the association between the FB and FS tasks remained. Together with their good performance on the control trials of the tasks, these findings suggest that children with ASD were neither selectively impaired in understanding mental representations, nor primarily impaired in language and executive function which mask their competence of understanding mental and non-mental representations.

Children with ASD performed as well as typically developing children on the FP task, supporting previous findings using the standard verbal FP task (e.g., Leekam & Perner, 1991). However, children with ASD have also been shown to fail a modified FP task (Russell, Saltmarsh, & Hill, 1999). If all the FP tasks involved only true representations, why did children with ASD succeed on some of them but not all of them? It is possible that children with ASD failed Russell et al.'s (1999) FP task because of its higher executive demands and its unnatural nature of taking a photograph of a screen rather than a person or an object, as Russell et al. has suggested. By having no object in the photograph, the photograph became less salient and thus harder to resist the interference from reality which was not the same as what was shown on the photograph.

General Discussion

The current study introduced a newly devised non-verbal FS task to test whether the non-specificity claim of theory of mind still holds in a non-verbal and reality-unknown context in both children with and without ASD. Experiment 1 and 2 showed that the *non-verbal reality-unknown* and *verbal reality-known* forms of the FS task showed equivalence across both forms of the FB task. Experiment 3 showed that children with ASD had consistent difficulties with the non-verbal reality-unknown forms of the false representation tasks. As a whole, the current findings suggested that the consistent performance by children with and without ASD may not be explained by the shared demands on language and cognitive inhibition between the tasks. It may better be explained in terms of an underlying conceptual capacity of understanding false representations, providing further evidence against the domain-specificity account of 'theory of mind'.

Testing Representational Understanding With Minimal Demands on Other Cognitive Skills

One advantage of the non-verbal reality-unknown tasks used in the current study was that these tasks do not require sophisticated language and cognitive inhibition skills in their assessment of children's understanding of representation. In this respect it is similar to the demands of other non-verbal tasks used in previous studies, which measured young children and infants' spontaneous looking behaviours while they were observing false-belief situations (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007). For example, Senju et al. (2009, 2010) employed Southgate et al.'s paradigm (2007) to test spontaneous anticipatory looking in individuals with ASD based on the protagonist's false belief. Results showed that individuals with ASD failed to show correct action anticipatory looking that would be in line with the protagonist's false belief. However, looking behaviours are ambiguous and alternative explanations are possible. For example, the lack of spontaneous anticipatory looking in individuals with ASD may be generally directed to social stimuli (Ruffman, Garnham, & Rideout, 2001) rather than specifically directed to false-belief attribution. Moreover, the correct anticipatory looking shown by typically developing individuals in the control groups could be explained by behaviour rules such as people look for objects where they last saw them (Perner & Ruffman, 2005; Ruffman & Perner, 2005). More importantly, Schneider, Lam, Bayliss and Dux (2012) recently highlighted that even implicit belief processing measured with looking behaviours requires executive resources and is affected by extra cognitive load. Therefore, these paradigms did not necessarily provide evidence that a false belief was attributed or genuinely understood with least demands on executive skills.

One might ask whether children's success on the FB and FS tasks in the current study could also be explained by behaviour rules (e.g., Perner & Ruffman, 2005). We suggested that this was not likely the case for the following reasons. First, the elements in the tasks of the current study were either rare (e.g., woman placed marker) or novel (e.g., electrical signpost) for children. Second, there was no evidence that children have learnt from the tasks and improved their performance across sessions in both Experiment 1 and 2. Although children with ASD (Experiment 3) did show improvement across sessions, they did not perform the tasks well overall. It was thus unlikely that children with and without ASD had picked up any rules from their everyday life or from the tasks per se that might have helped them pass the tasks.

Theory of Mind, False Belief and Domain Specificity

False-belief understanding has been considered as a crucial aspect of 'theory of mind' which is widely suggested to be a domain-specific mechanism dedicated to mental-related information processing only (e.g., Leslie et al., 2004). Moreover, individuals with ASD have been universally suggested to be specifically impaired in 'theory of mind' since Baron-Cohen et al. (1985; Senju et al., 2009, 2010). Cohen and German (2010) also recently showed that adults' reaction times to false-belief situations were faster than to false-map and false-sign situations, claiming that a domain-specific 'theory of mind' exists in human cognition. In the study, participants saw an actress hide an object and draw a map/arrow to indicate the object's location before leaving (but participants could not see the map/arrow). Participants then saw a man either move the object or leave it in place. Subsequently, an unpredictable test probe appeared, either saying "She thinks that the purse is in the right drawer" (false-belief condition) or "The map/arrow shows that the purse is in the right drawer" (false-map condition/false-sign condition). Participants had to make a "yes/no" response to the probe and their reaction times were measured. Given the map/arrow was drawn by the actress and was not shown to participants, we suggest that participants might have to first process what the actress believed in order to work out what she had drawn when they had to respond to the latter probe. This may lead to the longer reaction times found for the false-map/false-sign condition relative to the false-belief condition. If that was the case, then Cohen and German's results did not necessarily support the domain-specific claim of 'theory of mind'.

Similarly, one might argue that the non-verbal FS task in the current study may elicit mentalizing of the actress who intentionally manipulated the signpost's connection to electricity. However, this was not a necessary process for participants to figure out the object's location. An analogy is an electrical clock after a power-cut situation. One knows that the time shown on the clock falsely represents the current time as an earlier time without mentalizing of anyone who has possibly caused the power-cut. If participants with ASD failed the false representation trial of the task because they failed to mentalize, they would also have failed the control and filler trials of the task which involved the same actress performing the same sets of actions. Moreover, we made sure that every participant, with or without ASD, understood how the electrical signpost worked and that they could track the actions of the actress through the training phase, the warm-up trials, the control trials and the filler trials. It was therefore not possible that the participants with ASD failed the task because they failed to track the actress's actions.

Another theoretical concern about the claim against domain-specificity of 'theory of mind' is that a sign may arguably involve a mental component in its interpretation as it is intended by someone to be a felicitous communicative vehicle although it is physically non-mental. However, we argue that this concern cannot explain the asymmetries in associations between the FP versus FS and FB tasks found in the current study based on the nature of Apperly et al.'s FP task (2007). In this task, the protagonist showed children the photograph which served as a clue for them to find the object. Hence, the photograph here may also involve an intentional communicative component. This may be a reason why the raw correlations between the three tasks were significant in both Experiment 1 and 3. However, partial correlations between the FB and FP tasks were not significant but those between the FB and FS tasks were significant in both children with and without ASD. Thus, the falseness of representations which taps a genuine understanding of representation should be the core determinant of the equivalence between the tasks.

Indeed, ‘theory of mind’ is an umbrella concept which comprises processing of all types of inner, mental and emotional states. False-belief understanding involves not only mental but also representational characteristics so it should be regarded as a separate aspect of ‘theory of mind’ (e.g., Iao et al., 2011; Leekam et al., 2008). This separate aspect of ‘theory of mind’ can be specified as “*representational* theory of mind” which is distinguished from “*non-representational* theory of mind” by the requirement to interpret a mental state as a representation. Although individuals with ASD are widely found to show poor performance on the FB task (for meta-analyses, see Happé, 1995; Yirimiya et al., 1998), this difficulty in ASD seems not to apply to every task concerning mental states. There is evidence that mental states which do not require a “*representational* theory of mind” are not as difficult as false beliefs for individuals with ASD. For example, their level 1 visual perspective-taking is intact (e.g., Baron-Cohen, 1989), and their understanding of simple emotions is not worse than that of other mentally retarded populations (see Begeer, Koot, Rieffe, Meerum Terwogt, & Stegge, 2008, for a review). Moreover, children with ASD are capable of understanding goals and intentions (see Hamilton, 2009, for a review). Thus, individuals with ASD may not be as impaired in “*non-representational* theory of mind” as they are in “*representational* theory of mind”.

To conclude, the current study was the first to show that the association between tasks testing false beliefs and false signs was not due to their shared demands on language and executive function in both children with and without ASD. Rather, representational understanding is probably the key concept that underpins the processing of both mental and non-mental representations, providing further evidence for the non-specificity claim of ‘theory of mind’. Last but not least, the new non-verbal reality-unknown FS task opens further research opportunity for investigating atypical populations who have cognitive difficulties and the human brain functions.

References

Aichhorn, M., Perner, J., Weiss, B., Kronbichler, M., Staffen, W., & Ladurner, G. (2009). Temporo-parietal junction activity in theory-of-mind tasks: Falseness, beliefs, or attention. *Journal of Cognitive Neuroscience, 21*, 1179-1192.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revised). Washington, DC: American Psychiatric Association.

Apperly, I., Samson, D., Chiavarino, C., & Humphreys, G. (2004). Frontal and temporo-parietal lobe 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., Samson, D., Chiavarino, C., Bickerton, W., & Humphreys, G. (2007). 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. *Cognition, 103*, 300-321.

Baron-Cohen, S. (1989). Perceptual role-taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology, 7*, 113-127.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

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

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism, 16*, 357-366.

Begeer, S., Koot, H. M., Rieffe, C., Meerum Terwogt, M., & Stegge, H. (2008). Emotional competence in children with autism: Diagnostic criteria and empirical evidence. *Developmental Review, 28*, 342-369.

Bowler, D. M., Briskman, J., Gurvidi, N., & Fornells-Ambrojo, M. (2005). Understanding the mind or predicting signal-dependent action? Performance of children with and without autism on analogues of the false-belief task. *Journal of Cognition and Development, 6*, 259-283.

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

Charman, T., & Baron-Cohen, S. (1992). Understanding beliefs and drawings: A further test of the metarepresentation theory of autism. *Journal of Child Psychology and Psychiatry, 33*, 1105-1112.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the

MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language, 30*, 213-236.

Cohen, A. S., & German, T. C. (2010). A reaction time advantage for calculating beliefs over public representations signals domain specificity for ‘theory of mind’. *Cognition, 115*, 417-425.

de Villiers, J. G. (2000). Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg, D. Cohen (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (2nd ed., pp. 83-123). Oxford: Oxford University Press.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic Determinism and the Understanding of False Beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's Reasoning and the Mind* (pp. 191-228). Hove, UK: Psychology Press.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale, Second Edition*. London: NFER Nelson.

Figueras-Costa, B., & Harris, P. L. (2001). Theory of mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies in Education, 6*, 92-102.

Grant, C. M., Apperly, I., & Oliver, C. (2007). Is theory of mind understanding impaired in males with Fragile X Syndrome? *Journal of Abnormal Child Psychology, 35*, 17-28.

Hamilton, A. F. C. (2009). Research review: Goals, intentions and mental states: challenges for theories of autism. *Journal of Child Psychology and Psychiatry,* 50, 881-892.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66,* 843-855.

Iao, L.-S., Leekam, S., Perner, J., & McConachie, H. (2011). Further evidence for non-specificity of theory of mind in preschoolers: Training and transferability in the understanding of false beliefs and false signs. *Journal of Cognition and Development, 12*, 56-79.

Leekam, S., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203-218.

Leekam, S., Perner, J., Healey, L., & Sewell, C. (2008). False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology, 26*, 485-497.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in ‘theory of mind.’ *Trends in Cognitive Sciences, 8*, 528-533.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development:

Neuropsychological evidence from autism. *Cognition, 43*, 225-251.

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

Parkin, L. J. (1994). *Children's understanding of misrepresentation*. Unpublished D. Phil. Thesis, University of Sussex.

Paynter, J., & Peterson, C. (2010). Language and ToM development in autism versus Asperger syndrome: Contrasting influences of syntactic versus lexical/semantic maturity. *Research in Autism Spectrum Disorders, 4*, 377-385.

Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology, 46*, 530-544.

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

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Thinking of mental and other representations: The roles of left and right temporo-parietal junction. *Social Neuroscience, 1*, 245-258.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214-216.

Pratt, C., & Bryant, P. (1990).Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973-982.

Robinson, S., Goddard, L., Dritschel, B. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition, 71*, 362-368.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale, Revised*. Chicago: Stoelting Co.

Ruffman, T., Garnham,W., & Rideout, P. (2001). Social understanding in autism: Eye gaze as a measure of core insights. *Journal of Child Psychology and Psychiatry, 42*, 1083-1094.

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences, 9*, 462-463.

Russell, J. (1999). Cognitive development as a executive process – in part: a homeopathic dose of Piaget. *Developmental Science, 2*, 247-295.

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, 40*, 859-868.

Rutter, M., Bailey, A., & Lord, C. (2003). *SCQ: Social Communication Questionnaire*. Los Angeles: Western Psychological Services.

Sabbagh, M. A., Moses, L. J., & Shiverick, S. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs and false signs. *Child Development, 77*, 1034-1049.

Schneider, D., Lam, R., Bayliss, A., & Dux, P. E. (2012). Cognitive load disrupts implicit theory of mind processing. *Psychological Science, 23*, 842-847.

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., … Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology, 22*, 353-360.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science, 325*, 883-885.

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

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*, 587-592.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

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

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

World Health Organisation. (1993). *International classification of diseases* (10th ed.). Geneva: World Health Organisation.

Yirimiya, N., Erel, O., Shaked, M., & Solomonica-Levy, D. (1998). Meta-analyses comparing theory of mind abilities in individuals with autism, individuals with mental retardation and normally developing individuals. *Psychological Bulletin, 124*, 283-307.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and "false" photographs. *Cognition, 35*, 41-68.

Appendices

Appendix A

Key Stages of the Test Trials of False Belief, False Photograph and False Sign Tasks

False belief (Apperly et al., 2004)	False photograph (Apperly et al., 2007)	False sign (novel)
A man and a woman appeared.		A woman appeared with an electrical signpost covered by a screen but its plug was visible. She put the plug into a socket.
The man placed two boxes (one contained an object) on a table by the two sides of a marker and showed the interior of the boxes to the woman.	A man placed two boxes (one contained an object) on a table. A woman took a photograph of the boxes' interior and placed it face-down on the table.	The woman placed two boxes (one contained an object) on a table by the two sides of the screen. A mechanical noise was heard.
The woman left the room (so she was not able to update her belief about the location of the object).		The woman took the plug out (so the signpost was not able to update its indication of the object's location).
The man swapped the boxes.	The man swapped the boxes.	The woman swapped the boxes.
The woman's belief was manifested when she placed the marker on one of the boxes.	The photograph was shown when the man turned it over.	The signpost was revealed when the woman took the screen off.

Appendix B

Analogy Between the Components Involved in the False Belief and False Sign Tasks

False belief (Apperly et al., 2004)	False sign (novel)
Man (hiding object and showing to woman)	Woman (hiding object and enabling signpost)
Woman's presence	Signpost being plugged in
Marker placed between boxes (no indication of object's location)	Screen
Woman viewing inside the boxes	Signpost's turning (suggested by a mechanical noise)
Woman's absence	Signpost being unplugged
False belief being manifested by placing marker	False sign being revealed by removing screen

Appendix C

Key Stages of the Test, Control and Filler Trials of the False Belief Task (Apperly et al., 2004)

<table>
<tr><th>False representation test trial</th><th>Working memory control trial</th><th>Response inhibition control trial</th><th>True representation filler trial</th><th>Clue confirmation filler trial</th></tr>
<tr><td colspan="5">A man and a woman presented.</td></tr>
<tr><td colspan="5">The man placed two boxes (one contained an object) on a table by the two sides of a marker and showed the interior of the boxes to the woman.</td></tr>
<tr><td>The woman left the room.</td><td>The woman placed the marker on one box and put it back on table.</td><td colspan="2">The woman left the room.</td><td>The woman placed the marker on one box and put it back on table</td></tr>
<tr><td>The man swapped the boxes.</td><td>The woman left the room.</td><td>The man took the object out of one box and put it into another box.</td><td>The man held the boxes up and down vertically.</td><td>The woman left the room.</td></tr>
<tr><td>The woman returned, placed the marker on one box and put it back on table.</td><td>The man swapped the boxes.</td><td colspan="2">The woman returned, placed the marker on one box and put it back on table.</td><td>The man took the object out of the box indicated by the woman and put it into another box.</td></tr>
</table>

Appendix D

Key Stages of the Test, Control and Filler Trials of the False Photograph Task (Apperly et al., 2007)

False representation test trial	Working memory control trial	Response inhibition control trial	True representation filler trial	Clue confirmation filler trial
A man placed two boxes (one contained an object) on a table. A woman took a photograph of the boxes' interior and placed it face-down on the table. Then she left.				
The man swapped the boxes.	The man revealed the photograph and then covered it.	The man took the object out of one box and put it into another box.	The man held the boxes up and down vertically.	The man revealed the photograph and then covered it.
He revealed the photograph and then covered it.	He swapped the boxes.	He revealed the photograph and then covered it.		He took the object out of the box indicated by the photograph and put it into another box.