Children's Understanding of False Representations

Lai-Sang Iao

Thesis submitted to Cardiff University
For the degree of
Doctor of Philosophy

December 2010

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ....LAI SANG IAO.................. (candidate)   Date ..12/03/2011.......

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ....PhD.............(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ....LAI SANG IAO.................. (candidate)   Date ..12/03/2011.......

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed ....LAI SANG IAO.................. (candidate)   Date ..12/03/2011.......

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....LAI SANG IAO.................. (candidate)   Date ..12/03/2011.......

To my dearest parents,

for giving me all they have

Acknowledgements

I would like to express my gratitude to my supervisors Susan Leekam and Helen McConachie for their help, encouragement, and continuous support of my work. I would also like to thank Josef Perner for his valuable advice. Thanks also go to my examiners Ian Apperly and Ulrike Hahn for an enjoyable viva, a helpful discussion, and their insightful comments.

I am very grateful to the children who were a great pleasure to work with and have taught me a lot about development and autism. Special thanks to the parents, head teachers and teaching staff for their enthusiastic participation and precious time.

I thank Alexandra Thompson and Jeremy Thompson who made the empirical work of Chapter 2 possible; and David Marshall for lending me the Leiter International Performance Scale. I am also indebted to Nathalie Walters for her help with the administrative issues which arose during my transference from Durham University to Cardiff University. Appreciation is also extended to the staff at Cardiff University, especially Lesley-Anne Strabel and Kevin Hotson, and Durham University.

Many thanks to Anastasia Kourkoulou for her enormous help and support all the way from Durham to Cardiff; Keith Iao for listening patiently to my doubts and worries; Lee Shepherd and Charlie Edge for their invaluable friendship and accompany; and all my colleagues at the Wales Autism Research Centre for their amusing discussions and encouragement.

Lastly, this work would not have been possible without the unfailing love and support of my dearest parents. I would like to offer my deepest thanks to my parents. They may not completely understand the content of this thesis but it is actually their success.

The following publication has resulted from the work in this thesis:

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.

Thesis Summary

This thesis investigates whether children's understanding of both mental and non-mental representations can be accounted for by the same underlying competence in representational understanding. This research question stems from a long-standing dispute between domain-specificity and domain-generality in understanding mental representations. In Chapter 1, I highlight the importance of using *false* representations to assess representational understanding, and discuss a fundamental problem inherent in previously devised false non-mental representation tasks in comparison to false mental representation tasks. I also outline the confounding of other cognitive skills such as language and executive function during the assessment of representational understanding. This motivates the subsequent empirical work for this thesis which includes (1) the development of novel measures for assessing children's understanding of non-mental representations; and (2) the investigation of the equivalence between children's understanding of mental and non-mental representations. Evidence for this equivalence was shown by a transfer of training between a new false non-mental representation task and an existing false mental representation task presented in Chapter 2. With the use of another novel false non-mental representation task which minimises and eliminates the confounding factors of language and cognitive inhibition, the two experiments in Chapter 3 further indicated that the equivalence between false non-mental and false mental representation tasks could not be explained by these confounding factors. Chapter 4 extended the research from typical to atypical development, namely Autism Spectrum Disorder (ASD) as individuals with ASD are known to be specifically impaired in processing social and mental information. Intriguingly, the findings in Chapter 3 were shown to be generalised to children with ASD. Finally, the consistent findings of the five experiments reported in this thesis are discussed in relation to the theoretical accounts and neurological basis of typical and atypical cognitive development.

Contents

Tables and Figures

Chapter 1 Introduction

"'Hallo!' said Piglet, 'what are you doing?' ...

'Tracking something,' said Winnie-the-Pooh mysteriously. ... 'Now, look there.' He pointed to the ground in front of him. 'What do you see there?'

'Tracks,' said Piglet. 'Paw-marks.' He gave a little squeak of excitement. 'Oh, Pooh! Do you think it's a – a – a Woozle?'

'It may be,' said Pooh. 'Sometimes it is, and sometimes it isn't. You never can tell with paw-marks.'

With these few words he went on tracking, and Piglet, after watching him for a minute or two, ran after him. Winnie-the-Pooh had come to a sudden stop, and was bending over the tracks in a puzzled sort of way. ... 'but there seem to be two animals now. This – whatever-it-is – has been joined by another – whatever-it-is – and the two of them are now proceeding in company ...'

... Christopher Robin came slowly down his tree. 'Silly old Bear,' he said, 'what were you doing? First you went round the spinney twice by yourself, and then Piglet ran after you and you went round again, together, and then you were just going round a fourth time -'"

— Milne (1926)
"Winnie-the-Pooh"

Winnie-the-Pooh, who was also known as the Bear of Very Little Brain, made a silly mistake of tracking his own paw-marks. But why did he do that? It was because he thought his own paw-marks were someone else's. In other words, he had a false belief about the paw-marks on the ground. Not only the Bear of Very Little Brain, also human beings, who are supposed to be the most intelligent creatures in the world, tend to make similar mistakes due to false beliefs.

Belief, whether true or false, is a type of mental representation that represents and sustains information. However, paw-marks also form a representation, which is non-mental. They are a sign or indicator of an animal and the direction in which it is travelling. Both mental and non-mental representations present themselves extensively in everyday life and understanding them has a vital importance. For example, paw-marks represent a track of some kind of animal, but they can also be a sign representing that danger is around if they are created by a fierce beast. What you believe the paw-marks to represent determines your reaction of tracking or running away. Based on how you react, I can understand what you think the paw-marks represent. It is also critical to realise that representation has a crucial characteristic in that it can be false. For example, Winnie-the-Pooh falsely believed his own paw-marks as someone else's; paw-marks might also falsely represent a track of someone who wears a giant animal feet costume as the track of a fierce beast. Misleading beliefs, appearances, and signs may lead to inefficient, redundant, or even hazardous behaviours. However, weakness can be turned into strength. By understanding false representations and being aware of them, negative behaviours can be avoided. Therefore, both interpreting what a representation *represents* (e.g., Winnie-the-Pooh's belief represented an animal which created the paw-marks) and understanding that it may not be what it is *represented as* (e.g., that animal was falsely represented as Whatever-it-is) can have adaptive significance and is essential to children's cognitive and social development.

However, whether children's understanding of mental representations can possibly be explained by their understanding of representation in general has continuously been debated. Using non-mental representations, such as pictures or photographs, questions have been raised about whether mental representations such as false beliefs are understood in the same way as non-mental representations (Zaitchik, 1990), and whether training on either type of representation will enhance understanding of the other type of representation (Slaughter, 1998). Inconsistent findings and controversial comments about the non-mental representation tasks used in the studies have led the dispute to remain unsolved. To make the dispute even more complicated, research shows that understanding false representations often requires other cognitive skills such as language and executive function.

This thesis attempts to address this dispute of whether mental and non-mental representations are understood on the basis of a single concept of representation. In this review chapter, I first define what "representation" is and the essential elements it constitutes. I then describe some research on the prerequisites for understanding representations, attempting to make clear what researchers mean by representational understanding. Then, I review traditional and recent research on representational understanding and clarify the distinction between true and false representations. I highlight the importance of using false representations to assess representational understanding and discuss the significance of addressing whether mental and non-mental representations are underpinned by the same competence of representational understanding. Following this discussion, an overview is given of subsequent empirical chapters, describing the work for these empirical chapters, including the development of new tasks for assessing children's understanding of non-mental representations and investigation of the equivalence between children's understanding of mental and non-mental representations in various aspects. The aim of Chapter 2 is to investigate the question of whether training on understanding non-mental representations will improve children's understanding of mental representations and vice versa. Chapter 3 examines whether other cognitive processes such as language and executive function can explain the equivalence between the understanding of mental and non-mental representations. Chapter 4 extends the research from typical development to atypical development. Results of the empirical work presented in Chapter 2 and 3 jointly suggest that both mental and non-mental representations are underpinned by the same competence of representational understanding and that this competence is independent of language and an aspect of executive function. Importantly, this finding can be generalised to a group of atypically developing children who have always been known to be specifically impaired in processing social and mental information. This result can be found in Chapter 4.

What is "Representation"?

Reviewing the literature in psychology, the term "representation" has been given different meanings. Piaget (1951) used "representation" to mean (1) something that stands for something else; and (2) general intelligence. More recently, Perner (1991) defined "representation" in more specific terms: representational *medium* which is an *entity* that represents (e.g., Winnie-the-Pooh's *belief*), or representational *process* which is an *activity* of representing (e.g., Winnie-the-Pooh was *entertaining a representation* of an animal which created the paw-marks). Another researcher, Denis (1994), also proposed a distinction between representation as *product* and representation as *process*. To avoid confusion, in this thesis the term "representation" is specified to refer to representational medium in Perner's sense.

A representational medium can be a mental entity (e.g., belief), but it can also be a physical object (e.g., paw-marks). It stands for, or represents, something else (e.g., Winnie-the-Pooh's belief represented an animal which created the paw-marks; paw-marks represent a track of an animal). Hence, representations could be considered according to two types: mental or non-mental. Some obvious examples of non-mental representations are pictures, signs, symbols, numerals, words, and sentences. Mental representations can be conceptualised as structures in the mind that represent and sustain information. Examples of mental representations include memory, knowledge, thoughts, and beliefs. As there are many kinds of things that can be representations, pictures, signs, and beliefs are the main focus of this thesis due to a heated debate about humans' understanding of these representations in the current literature (e.g., Apperly, Samson, Chiavarino, & Humphreys, 2004; Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007; Cohen & German, 2010; Leekam, Perner, Healey, & Sewell, 2008).

Having specified "representation" as a representational medium (mental or non-mental) which represents *something else*, it is also important to clarify the nature of this "something else". Goodman (1976) termed this "something else" as "referent". The referent of a representation can be almost anything, which varies from physical objects to non-exsisting entities. A picture or a thought can represent physical objects (e.g., sunflowers), non-physical objects (e.g., anxiety), and non-existing entities (e.g., dragons). However, Goodman furthered that a representation does not only represent a referent but also represents it as being a certain way (sense). For example, Winnie-the-Pooh's belief represented the animal that created the paw-marks (referent) as a Whatever-it-is (sense), whereas Piglet's belief represented it as a Woozle. Both Winnie-the-Pooh and Piglet's beliefs refer to the same animal (referent) but they differ in their senses.

Therefore, representations are entities which represent something (referent) as being a particular way (sense). The ability to understand or mentally represent that a representation is representing its referent as being a certain way is termed as metarepresentation (Pylyshyn, 1978). Without this ability, a representation is not understood as an entity which has a referent as well as a sense; rather, it is confused with its referent and its referent is not differentiated from its sense. Throughout this thesis, the terms "metarepresentation", "representational understanding" and "understanding of representations" are used synonymously.

Prerequisites for Metarepresentation

As early as 3 months of age, infants are able to perceive the similarity between a representation and its referent. For example, 3-month-olds recognised their mother's face in photographs (Barrera & Maurer, 1981) and 5-month-olds, after being habituated to a person's real face, showed dishabituation to a photograph of a novel face rather than a photograph of the familiar face (Dirks & Gibson, 1977). Later research further demonstrated that older infants do not only recognise the similarity between a representation and its referent, but also confuse the two. When presented with photographs of objects, 9-month-olds manually explore (e.g., feel, rub) the photographs and grasp at them as if attempting to pick up the depicted objects (DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998; Pierroutsakos & DeLoache, 2003). Infants act on photographs of objects as if the depicted objects were real, but this is not because they cannot discriminate between depicted and real objects. If they are presented with real objects and photographs of the objects, they prefer the former over the latter. However, their manual exploration of photographs greatly decreases from 9 to 15 months of age. By 18 months of age, manual exploration is replaced by pointing and talking (DeLoache et al., 1998; Pierroutsakos & Troseth, 2003), suggesting that infants gradually appreciate how representation differs from its referent with age and through experience.

It is also around the same age period that infants have been shown to be able to interpret representations. Evidence comes from studies which observed the reactions of infants, aged 9 to 24 months, with a red spot painted on their noses when they were in front of a mirror. If infants recognise their mirror images as themselves, they will show more self-directed behaviours, e.g., touch their noses. This nose-directed behaviour is first seen at 15 to 18 months of age and has been found to gradually increase with age; whereas sociable "playmate" or mirror-directed behaviours, e.g., smiling at and touching the "playmate" in the mirror, dominate in younger ages (Amsterdam, 1972; Lewis & Brooks-Gunn, 1979). Infants younger than 15 months act as if their mirror images were real persons. The finding of 18 months as the mean age of emergence for mirror self-recognition was replicated in a more recent study (Nielsen & Dissanayake, 2004).

More recent studies also indicate that 15- to 18-month-old infants are able to interpret words and pictures (Ganea, Bloom-Pickard, & DeLoache, 2008; Ganea, Preissler, Butler, Carey, & DeLoache, 2009; Preissler & Carey, 2004). The general methodology of these studies is to teach children a novel word (e.g., 'whisk') for a picture of a novel object (e.g., a whisk) and test whether they transfer the word from the picture to the object depicted (e.g., identifying the real whisk out of two real objects; indicating the real whisk, either alone or together with its picture, when being asked for a whisk). Results suggest that even 15-month-olds interpreted the word as corresponding to a certain object, not just to the picture which has been paired with the word. However, when being shown a real whisk that differed in colour from the one depicted in the picture, children as old as 24 months failed to extend the word "whisk" from the depicted whisk to the real whisk of a different colour (Ganea et al., 2009). This finding suggests that 24-month-olds interpreted the word and the picture as corresponding to a real whisk of the same colour, but not to a real whisk of a different colour because it did not correspond to the one depicted in the picture. Thus, the word and the picture were not yet understood as representations which *represent* a real whisk of any kind.

The ability to interpret representations as corresponding to something also allows children above 25 months old to match an outcome of a real or even pretend action to the correct picture that depicts this outcome (Harris, Kavanaugh & Dowson, 1997). Given that in this task the outcome of a pretend action was not visible when children made their choice of pictures, perceptual similarity between the outcome and the picture is not a necessary element for determining the correspondence between them. This explains why 24-month-olds were able to transfer the word "whisk" from a line drawing of a whisk to a real whisk in Preissler and Carey (2004) although the perceptual similarity between the line drawing and the real whisk was less than that between the highly realistic picture of a whisk and the real whisk of a different colour in Ganea et al. (2009). Thus, the lack of word transference from the picture of a whisk to the real whisk of a different colour in Ganea et al. (2009) is not simply due to the perceptual dissimilarity between them. Rather, it is children's interpretation of the picture/drawing determines the correspondence between the picture/drawing and the real whisk.

DeLoache and Burns (1994, Experiment 1 & 6) also reported that 24-month-olds could use a picture or a line drawing of a room to guide their placement of an object in the room, illustrating their ability of corresponding the picture/drawing and the room. However, they were not able to use the picture/drawing to guide their search of the object in the room (Experiment 1 - 5). These findings suggest that 24-month-olds can understand the correspondence between the picture/drawing and the room, but they may not be able to take advantage of this correspondence to infer the real location of the object so they failed the searching task. However, when they come to the age of 30 months, they can correctly infer the real location of the object from the picture/drawing. Salsa and de Mendoza (2007) further demonstrated that children aged 30 months succeeded in the searching task only when they were first asked to place the object in the corresponding real location as the depicted location the experimenter has pointed to. In other words, they were asked to *use* the picture to guide their mission towards the room. In a follow up experiment, the authors showed that children aged 36 months succeeded without any instructions about correspondence and functionality.

With the understanding of correspondence and the spontaneous ability of inference, children aged 36 months also understand what other people can and cannot see (i.e., level 1 visual perspective-taking). For example, Flavell, Everett, Croft, and Flavell (1981) showed children a drawing of a turtle which was bisected vertically by a piece of cardboard so that the turtle's back, but not its feet, was visible from one side, and its feet, but not its back, from the other side. Children were asked whether the experimenter, who sat opposite to them, saw the turtle's back or feet. Perfect performance was shown by children aged 3 years, suggesting that they understood the correspondence between other people's visual perceptions and the situation in front of their eyes, and correctly inferred what they saw. Research also shows that 3-year-olds consider mental representations, such as perceptions and beliefs, transparent in that they reflect or correspond to reality (e.g., Flavell, Flavell, & Green, 1983).

So far, none of the above studies showed a full understanding of representations in children aged 36 months and below because these children were not shown to understand a representation as an entity which (1) requires interpretation; (2) represents something (referent) as being a certain way (sense); and (3) may become false when its referent is distinct from its sense. Children by the age of 36 months interpret representations spontaneously but they do not realise that a representation is an object in itself and needs to be interpreted (e.g., the word "whisk" and the picture of a whisk need to be interpreted as representing a real whisk). They only spontaneously interpret representations as corresponding to something (e.g., the picture of a red whisk corresponds to a real red whisk). This understanding of correspondence does not differentiate *what* the representation is representing (referent) from how this referent is represented *as being* (sense). That is, the picture of a red whisk is interpreted as corresponding to a real red whisk, but is not treated as a representation which represents a real whisk (referent) as being red (sense). Hence, understanding correspondence does not require an understanding of representations. Further, an understanding of correspondence and an ability for inference allows children to *use* representations to solve a problem (e.g. taking advantage of the correspondence between a picture and a room to infer an object's location in the room), but not *understand* representations (e.g., the picture) as representing something (e.g., the room).

Metarepresentation I: Understanding of Representations

The understanding of representations develops after the age of 3 years. Children become aware that representations need to be interpreted and can be given different interpretations (e.g., paw-marks can be interpreted as the track of a Woozle or a Whatever-it-is). Masangkay et al. (1974), for example, investigated "level 2" visual perspective-taking and found that children understand that another person who views a picture from a different viewpoint will give the picture a different interpretation by the age of 4 years. In this study, children were shown a picture of a turtle and were required to distinguish whether the turtle was standing upright or lying on its back from their viewpoint or another person's viewpoint. A few 3-years-olds but many 4-year-olds understood that the other person who sat opposite to them would see the turtle as it was lying on its back when they themselves saw it as it was standing upright. In other words, they understood that the picture represented a turtle and it could be interpreted as the turtle standing or lying, depending on which viewpoint it was being viewed from.

Children from the age of 4 years also come to understand that an ambiguous figure has two different interpretations, e.g.,  can be interpreted as both a duck and a rabbit (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Rock, Gopnik, & Hall, 1994). Even though switching from one interpretation to the other, an ability termed as reversal, was found to be difficult until the age of 5 years, some 4-year-olds did report reversal and most 4-year-olds acknowledged both interpretations after being informed about the ambiguity of the figures. On the contrary, most informed 3-year-olds could neither reverse nor acknowledge the two different interpretations of each ambiguous figure.

In addition to the awareness of different possible interpretations of a single representation, metarepresentation involves the understanding that a representation is an entity which represents something (referent) as being a certain way (sense). If there is a reliable mechanism that ensures the correspondence between a representation and its referent, this representation is a true representation which faithfully represents its referent as the way the referent really is (sense). For example, photography, under normal circumstances, is a kind of reliable mechanism which produces a photograph that truly represents a scene at the time the photograph was taken. If Bert takes a photograph of Duckie while it is on a bed using a Polaroid camera, the scene of Duckie on the bed (referent) will be represented by the photograph as the way that scene exactly is (sense). Thus, a photograph is usually a true representation which has a sense and a referent that correspond with each other. If Duckie is then removed from the bed to a bathtub, the photograph will still truly represent the scene of Duckie on the bed as the way it *was*. This understanding of photograph was tested on children aged 3 to 5 years in Zaitchik (1990). Without being shown the photograph of Duckie, children were asked, "In the picture, where is Duckie?" Most 3-year-olds answered "bathtub" (incorrect answer) whereas 4- to 5-year-olds replied "bed" (correct answer). Zaitchik pointed out that this finding was not due to children's lack of knowledge of photography, and could neither be explained by their failure to infer the content of the photograph nor memory of Duckie's location at the time when the photograph was taken. However, whether the correct responses given by the 4- to 5-year-olds implied an understanding that a representation has a referent and a sense cannot yet be concluded for the following reason. As the photograph of Duckie was a true representation which corresponded to its referent and sense, children could possibly pass this photograph task using their understanding of *correspondence* rather than representation. Although 3-year-olds also understood correspondence, as mentioned in the previous section, they failed the task probably because their understanding of correspondence tended to be bound by *reality*. More evidence regarding this possibility is outlined and discussed below.

Another example of true representation is knowledge. It is a true *mental* representation which reflects information faithfully when one has had perceptual access to that information. In this case, perception is the reliable mechanism that ensures the truthfulness of the representation – knowledge. For example, if someone has looked inside a box and seen a red ball, he or she knows what is in the box. That is, he or she holds knowledge that represents the content of the box (referent) as what it really is (sense; i.e., as a red ball). Understanding the relationship between perception and knowledge is shown by children aged 3 to 4 years who attribute knowledge of the contents of an opaque box to people who have looked inside the box rather than to those who have not (e.g., Pratt & Bryant, 1990). However, whether knowledge was understood as a representation which has a referent and a sense was not demonstrated in this "see-know" task.

It is important to note that perception does not always reliably reflect reality and lead to knowledge. Perception itself is also a mental representation which differs from knowledge in that it may not be true. Children's ability to understand not only the distinction between perception and knowledge but importantly "seeing as" has been investigated by Flavell et al. (1983). In the study, children were shown a sponge which looked like a rock so they saw the sponge as a rock. In other words, their visual perception falsely represented the object in front of them (referent; i.e., the sponge) as a rock (sense) and hence they did not have the knowledge about the real identity of the object. After tactile exploration, their tactile perception led to knowledge, both of which truly represented the object as a sponge. Flavell et al. then asked children who aged 3 to 4 years two questions: (1) "What is this really, really? Is it really, really a rock or really, really a piece of sponge?" (2) "When you look at this with your eyes right now, does it look like a rock or does it look like a piece of sponge?" Results showed that most 4-year-olds could differentiate the real identity of the sponge from its appearance (i.e., answered both of the test questions correctly), whereas most 3-year-olds could not. Children aged 3 years tend to answer "sponge" to both questions. This finding implies that 3-year-olds, as mentioned earlier, tend to regard perception, whether it is visual or tactile, as corresponding to reality or knowledge. Some researchers termed this phenomenon as *reality bias* (e.g., Mitchell & Lacohee, 1991) whereas others called it *the curse of knowledge* (e.g., Birch & Bloom, 2004). As perception is not yet understood as a representation, 3-year-olds cannot appreciate the distinction between appearance (sense) and reality (referent).

In sum, a true representation confounds its referent with its sense that any success in tasks concerning true representations could be alternatively explained by children's understanding of correspondence. Only by evoking a striking difference between a representation's referent and its sense, can children's understanding of representation as having a referent and a sense be examined more unambiguously.

Metarepresentation II: Understanding of False Representations

When a representation's sense is distinct from its referent, this representation becomes a false representation. In the story of *Winnie-the-Pooh*, the animal that created the paw-marks was actually Winnie-the-Pooh himself. If Winnie-the-Pooh or Piglet's belief had represented that animal (referent) as Winnie-the-Pooh (sense), his belief would have been a true representation because its referent corresponded with its sense. However, that animal (referent) was either represented as a Whatever-it-is or a Woozle (sense). The sense was at odds with the referent so Winnie-the-Pooh and Piglet each had a false representation in mind.

Understanding of false representations requires a clear distinction between sense and reference. To determine whether children understand false representations, there is a well-established task in which a protagonist's belief becomes false when its referent no longer matches with its sense. This is the false belief (FB) task (e.g., Wimmer & Perner, 1983). It shows a protagonist who observes an object being placed in location A but is then absent when the object is moved to location B. Children are then asked where the protagonist thinks the object is or where he will go for the object when he returns. Both of these test questions ask about the protagonist's belief of the object's location. Before the move of the object, the protagonist held a belief which truly represented the object's location (referent) as location A (sense) because the object's location was really location A at that time. Then, the protagonist has not witnessed the change of the object's location so his belief still represents the object's location as location A. However, the object's location is now location B, causing a mismatch between the referent and the sense of the protagonist's belief. Thus, his belief is now false. In order to answer the test questions correctly, children have to understand belief as representation which has a referent and a sense, and be able to distinguish between reference and sense. Children aged 3 years usually fail the task by replying that the protagonist will search in location B for the object, providing evidence that they take belief as corresponding to reality; whereas children aged 4 years and above pass the task by replying location A.

Since the 1983 study, different versions of the FB task have been introduced in the literature. For example, there is a FB task which involves a change of contents rather than locations. In this *contents change* FB task (e.g., Hogrefe, Wimmer, & Perner, 1986; Perner, Leekam, & Wimmer, 1987), children are presented with a familiar box (e.g., a "Smarties" box) and asked what they think is inside. Having responded “Smarties”, children are shown that the box contains, not Smarties, but a pencil. The box is then closed with the pencil inside and children are asked what another person, who has never seen the real contents of the box, will think is in it. This test question, like the one in the previous *location change* FB task, assesses children’s comprehension of another person’s false belief so it is usually named as the ‘other’ question. Children are also asked what they thought was inside the box before they were shown its real contents. This is the ‘self’ question which tests children’s comprehension of their own prior false belief. Although there are a variety of FB tasks, which vary in materials and procedures, children across countries demonstrated a consistent developmental pattern on these tasks: they improved their performance from below-chance to above-chance levels between the ages of 2.5 and 5 years (for a meta-analysis, see Wellman, Cross, & Watson, 2001).

Belief, however, is not the only type of representation that can be used to investigate children’s understanding of false representation. Another type of representation that has been used is photographic or pictorial representation. Nevertheless, this research using “false” photographs and “false” drawings has been shown to have problems associated with it. In the following, I briefly outline this research and its findings. Its problems are subsequently discussed with conceptual analysis and empirical evidence.

Zaitchik (1990) was the first who attempted to make a photograph become “false” by changing the location of Duckie after photographing it, as described in the previous section. Without seeing the front of the photograph (Duckie on bed) which was in conflict with the current reality (Duckie in bathtub), children were asked where Duckie was in the photograph. This “false” photograph (FP) task was claimed to be a structurally similar comparison task to the FB task but tested children’s understanding of a non-mental representation. A similar developmental pattern to that found on the FB task was shown on this FP task: 3-year-olds failed but 4- to 5-year-olds passed. Other studies using similar photograph tasks (e.g., Leekam & Perner, 1991; Leslie & Thaiss, 1992) or drawing task (Charman & Baron-Cohen, 1992) also found comparable performance between these photograph/drawing tasks and the FB task in different samples of typically developing children. These findings were interpreted as demonstrating that typically developing children at the age period of 4 to 5 years are capable of understanding representations, regardless of whether these representations are mental or not.

However, it is important to note that the photographs and drawings used in these tasks represented the scenes at the time the photographs/drawings were taken/drawn. Moreover, these scenes were represented as the way they really were at the time the photographs/drawings were created. Therefore, these photographs and drawings were not false and no distinction between reference and sense was required. According to this conceptual analysis, Leekam and Perner (1991) argued that the difference between these photographs/drawings and false beliefs is not purely a difference in mental versus non-mental state. They are also different in the sense that these photographs/drawings are true representations of the situations at the time the photographs/drawings were created; in contrast, false beliefs are false representations of whatever they are supposed to represent. Hence, these photographs and drawings should not be considered as comparable to false beliefs.

Tasks using these photographs and drawings, collectively known as the FP task, can possibly be passed by employing an understanding of correspondence, as discussed in earlier sections; whereas the FB task requires an understanding of representation which clearly distinguishes reference from sense. Furthermore, if the FP and FB tasks were equivalent and reflected a common underlying ability of representational understanding, a positive correlation, in addition to similar performance, would have been illustrated between the two tasks. Five studies have been carried out and none of them found a correlation between the FP and FB tasks (Davis & Pratt, 1995; Lewis, Freeman, & Smith, 1992; Perner, Leekam, Myers, Davis, & Odgers, 1993; Peterson & Siegal, 1998; Slaughter, 1998). Although Leekam and Perner (1991) demonstrated a correlation, it was not reported whether this correlation remained when children's age or intellectual level were controlled. Two more recent studies found weak correlations but these correlations did not persist after age was controlled for (Leekam et al., 2008; Sabbagh, Moses, & Shiverick, 2006).

In addition to correlation, researchers have employed other experimental methodologies to test the equivalence between the FP and FB tasks. The general logic behind these methodologies is that any manipulation that enhances performance on one task should also improve performance on the other task if the two tasks are equivalent. However, Perner et al. (1993) demonstrated that drawing children's attention to the back of the photograph improved their performance on the FP task, but drawing their attention to the person holding a false belief did not improve their performance on the FB task. More specifically, training on one task should improve performance on that task as well as the other untrained task if the two tasks tap the same underlying concept. Nevertheless, training on the FP task has not been found to enhance children's performance on the FB task and vice versa (Slaughter, 1998, Study 2), indicating that the FP and FB tasks are not only conceptually distinctive but also empirically non-equivalent by being non-transferable. Hence, the FP task is not appropriate as a control for singling out the mental aspect of the FB task because the comparison confounds the mental nature of the tasks with their "falseness".

To adequately realise Zaitchik's (1990) original research goal of measuring children's understanding of *false* non-mental representation, a false sign (FS; Parkin, 1994) task was subsequently devised to replace the FP task. In this FS task, a signpost initially shows an object in location A but then the object is moved to location B without changing the signpost which thus becomes a false sign. Participants are asked where the signpost shows the object is. A sign, like a belief, represents whatever it is supposed to, which is usually up-to-date with respect to reality. For example, a sign indicates the way to an ice-cream van by pointing along a road that goes behind a house. Participants can thus use this sign to find the van behind the house. But then the van drives over to a church which is situated on a different road. The driver forgets to change the sign, which keeps pointing along the road that goes behind the house. When the van is stationed behind the church, participants are asked, "Where does the sign show the ice-cream van is?" In this case, the *referent* of the sign was the location of the ice-cream van (i.e., behind the church) which does not match with its *sense* of showing the ice-cream van behind the house. In sum, the FS task is non-mental, in the same way as the FP task; but it genuinely misrepresents a current state and requires a distinction between reference and sense, in the same way as the FB task. Consistent with this conceptual equivalence, Parkin obtained robust correlations between children's performance on the FS and FB tasks.

Sabbagh, Moses, et al. (2006, Experiment 2) adopted the FS task and found a marginally significant correlation between the FB and FS tasks but non-significant correlation between the FB and FP tasks when age was partialled out. Leekam et al. (2008) also observed no significant difference between the FB and FS tasks; moreover, the correlation between the two tasks was significantly greater than the correlation between the FB and FP tasks when controlling for the third task and age. Both studies also showed no significant difference between two different versions of the tasks: one in which the location changed and one in which the contents changed.

Furthermore, there is a study (Bowler, Briskman, Gurvidi, & Fornells-Ambrojo, 2005) which developed a false signal (non-mental, mechanical) task, in which participants had to predict the destination of an automatic train, to correspond with the FB task. The train in this false signal task had no driver. It automatically picked up cargo from planes based on a signal which either accurately or falsely indicated the planes' location. This false signal task was found to be highly correlated with the FB task in children with typical development, intellectual disabilities, and autism. These findings suggest that the FB and FS/false signal tasks are related, sharing a developmental factor which is not shared by the FP task.

This shared developmental factor which underlies both FB and FS tasks could be a general understanding of representations which clearly differentiate reference from sense. If so, false beliefs are difficult for young children not because they are mental but because they are representational and require distinctions between references and senses (Perner, 1991). Similar claims have been proposed by others in various forms (Astington, 1993; Fodor, 1992). However, this shared developmental factor could also be argued to be a developmental coincidence between two distinct concepts, one for understanding mental representations and the other for understanding non-mental representations. Or one might further argue that it could be a range of other general cognitive skills, e.g., executive function, which underlies both FB and FS tasks but not the FP task. The main aim of this thesis is therefore to critically investigate whether false mental and false non-mental representations are equivalently understood on the basis of a unitary concept of representations which enables the discrimination between reference and sense.

Main question: Can False Mental and False Non-Mental Representations be Understood Using the Same Underlying Concept of Representations?

Although there is evidence showing that children's understanding of false signs is equivalent to their understanding of false beliefs, the question of whether understanding beliefs should be treated as a capacity for representational understanding which covers both mental and non-mental representations is still under debate. This stems from the traditional view that the ability to understand beliefs is central to a mentalistic understanding of self and others which focuses on the mental rather than the representational aspect of beliefs. Performance on the FB task has long been used as the index of "theory of mind" which refers to the ability to read a broad range of inner, mental and emotional states (e.g., desire) that may not necessarily be treated as representations (Wellman et al., 2001). In particular, "theory of mind" has been regarded as a domain-specific mechanism which is limited to mental-related information processing since the 1980s (Leslie, 1987, 1994; Leslie, Friedman, & German, 2004; Leslie & Thaiss, 1992).

Recently, Cohen and German (2010) 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 were instructed to watch some videos and track an object's location in the videos. Each video showed the object being put in a drawer by an actress who then sat by a table with her back facing participants. At this point, participants *knew* that the actress was drawing a map/arrow to indicate the object's location because this had been communicated to them at the outset in a number of practice trials, so that they could make sense of the videos even when the maps/arrows were not shown in the main task (T. C. German, personal communication, August 11, 2010). Participants then saw the actress leave and a man coming in to either move the object or leave it in place. Subsequently, the video was interrupted with an unpredictable test probe, either saying "She thinks that the purse is in the right drawer" for the false-belief condition or "The map/arrow shows that the purse is in the right drawer" for the false-map/false-sign condition. Participants had to make a "yes/no" response to the probe and their reaction times were measured. Although Cohen and German claimed that false maps/signs were processed slower than false beliefs, I would like to propose a criticism of their claim. If the map/arrow was drawn by the actress and was not shown to participants, participants might have to first process what the actress believed in order to work out what she has drawn when they had to response to the latter probe, resulting in longer reaction times for the false-map and false-sign conditions relative to the false-belief condition. If that was the case, then their results did not necessarily support the domain-specific claim of "theory of mind".

As a result, even very recent experimental evidence has not provided a clear conclusion for the question of whether mental representations are understood with a general representational understanding system or a domain-specific mechanism which dedicates to mental-related processing only. For more than two decades, this question has been highly debated and it is still important because (1) it affects the explanation of *what* changes with typical development and *how* the change takes place, and (2) it relates to our understanding of the human brain. Each of these two issues is discussed next.

Developmental change: Competing theories of understanding mental representations.

There have been competing developmental theories of understanding mental representations. Theorists, taking the general system approach, emphasise that children's understanding of mental representations depends on a general theory of representation which develops with age (e.g., Perner, 1991; Wellman, 1990). Young children initially understand the world with a simplistic, intuitive theory of common sense and rules. For example, they understand "seeing" and "knowing" as synonymous with each other, without a representational understanding of "seeing" and "knowing". As children grow older, they are confronted with an accumulating amount of counterevidence (e.g., visual perception can be misrepresented by deceptive appearance) which then triggers a transition of "theory change". Hence, this conceptual change of theory, shifting from the theory of common sense to the theory of representation, is mainly determined by experience. Although this change of theory occurs around the age of 4 years, the theory of common sense remains as the default theory unless an understanding of representations is necessary in dealing with a certain situation.

On the other side of the debate, theorists supporting the domain-specific claim propose that the mechanism for understanding mental representations is inborn and takes the form of a specific module in the brain (e.g., Leslie, 1987, 1994). It is suggested to be separate and dissociable from mechanisms which are responsible for the understanding of non-mental representations. As a result, there are at least two types of mechanisms for the two different types of representations (mental and non-mental). The mechanism for understanding mental representations is assumed to emerge on its own maturational timetable which leads to an early conception of mental representations. In that sense, children's failure on the FB task is not due to their incompetence of understanding mental representations but because of a lack of sophisticated general cognitive skills required for successful performance of the task, such as language and executive function. Thus, the neurological maturation of the domain-specific module and general cognitive skills provide the mechanism for change occur through development. During this changing process, experience does not play a necessary role.

Role of other general cognitive skills: Language and executive function.

One of the problems of previous research on children's understanding of false belief is that the results might be due to other cognitive skills such as *language processing ability*. The FB task does not only require children to comprehend the story about a protagonist and an object's location, but also involves a linguistically complex structure such as a proposition embedded in another (e.g., Where does she think that X is?). Many studies have demonstrated a relationship between language and false-belief understanding (e.g., Astington & Jenkins, 1999; de Villiers & Pyers, 2002; Lohmann & Tomasello, 2003; for a meta-analysis, see Milligan, Astington, & Dack, 2007). Other confounding cognitive factors include: (1) the need to disengage from participants' own knowledge about the salient reality and attending to the protagonist's abstract representation (cognitive inhibition), (2) the need to inhibit a prepotent response of pointing to the true location of the object (response inhibition), and (3) the need to hold and process information in mind simultaneously (working memory). These three cognitive factors together with the others, such as planning and cognitive flexibility, are collectively referred to as *executive function*. Evidence has been provided for the relationship between false-belief understanding and executive function, such as inhibitory control (e.g., Carlson & Moses, 2001; Perner, Lang, & Kloo, 2002) and working memory (e.g., Davis & Pratt, 1995; Gordon & Olson, 1998).

Much research has focused on which specific aspects of language (e.g., grammar or pragmatics) and executive function (e.g., inhibitory control or working memory) are most strongly associated with false-belief understanding. The question of the extent to which language and executive function is involved in false-belief understanding is less investigated. There was a study by Call and Tomasello (1999) who first attempted to disentangle language and executive function from false-belief understanding by devising a non-verbal FB task which did not require cognitive inhibition. This study did not receive much attention although researchers have developed an interest in isolating other cognitive factors from false-belief understanding as a result of a greater interest in populations other than typically developing 3- to 5-year-olds over the last decade. In the following, I first outline Call and Tomasello's study and their non-verbal FB task. Then I move on to discuss the research on false-belief understanding in young children and infants, followed by other growing false-belief research developments in adults and brain-damaged patients.

Call and Tomasello (1999) tested children aged 4 to 5 years with a non-verbal version of the location change FB task which was embedded in a hiding-finding game context. In this context, children were asked to find an object, which was hidden by an experimenter, with the help of a communicator who knew the object's location. The experimenter hid the object in one of two identical containers without showing this hiding process to the children. However, the communicator saw the hiding process so she knew the object's location and helped the children to find it by placing a marker on the correct container. Having understood the game, children were then screened on their ability in mastering general task requirements: (1) to track the object being moved to a new location; (2) to track the container marked as containing the object being moved to a new location (i.e., swapping containers around); and (3) to ignore the communicator's marker when it was known to be incorrect (i.e., children had seen the object being moved while the communicator was away). After successful completion of the screening phase, children were given the testing phase. During the testing phase, the communicator watched the hiding process and left while the experimenter swapped the containers around. It is important to note that children until this point did not know where the object was, so no cognitive inhibition of knowledge about reality was required. The communicator returned and placed the marker on the container at the location where she saw the object being hidden. As the containers had been swapped while the communicator was away, she had a false belief about the object's location and hence placed the marker on the wrong container. Children had to reason about the communicator's false belief in order to correctly find the object. In addition to this non-verbal FB task, a verbal FB task which corresponded to the standard FB task was incorporated into the game context. It was found that children's performance on both non-verbal and verbal versions of the FB task were highly correlated. This result was consistent with the typical developmental change found with the standard FB task in previous studies. This finding suggests that language and cognitive inhibition did not have a critical effect on false-belief understanding otherwise children should have performed better on Call and Tomasello's non-verbal FB task.

Although Call and Tomasello's (1999) non-verbal FB task was later adapted to be used with deaf children who are raised by hearing parents (i.e., late exposure to sign language; Figueras-Costa & Harris, 2001), brain-damaged adult patients (Apperly et al., 2004), and low-functioning children with autism and children with specific language impairment (Colle, Baron-Cohen, & Hill, 2007), it has still not gained as much attention as the non-verbal paradigms designed for testing false-belief understanding in young children and infants (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007). All these infant studies used the violation-of-expectation and anticipatory-looking paradigms which measure spontaneous looking behaviours while infants are observing false-belief situations to infer their attribution of false beliefs. These studies have all concluded that young children are capable of attributing false beliefs to other people, further suggesting that false-belief understanding exists independently of the development of language and executive function in ages earlier than 3 years.

In the violation-of-expectation paradigm, young children are tested as to whether they look significantly longer when a protagonist's behaviour is incongruent, as opposed to congruent, with his/her false belief. The logic behind this is that infants show their expectation of people behaving according to what they believe by looking longer at behaviours that are incongruent with the actors' beliefs. For example, the well-known study of Onishi and Baillargeon (2005) first familiarised 15-months-olds to a protagonist hiding an object in one of two boxes (e.g., box A) and reaching inside box A. Then, in a false-belief condition, the object moved to another box (e.g., box B) in the protagonist's absence. When the protagonist reappeared, infants looked reliably longer if she reached box B (incongruent with her false belief) instead of box A (congruent with her false belief). Scott and Baillargeon (2009) extended this finding from false belief about location to false belief about identity. Together with other violation-of-expectation studies (e.g., Surian et al., 2007), it was concluded that infants as young as 13 months expected a protagonist to act on the basis of his/her false belief.

However, alternative explanations are available. For example, Perner and Ruffman (2005) suggested that infants may create associations between different elements (e.g., protagonist-object-box A); and that a similar combination of elements will need less processing and thus a shorter looking time than a different combination of elements (e.g., protagonist-object-box B). In other words, it is possible that infants look longer at behaviours or situations based on their novelty-familiarity rather than unpredictability-predictability. Similar criticisms towards this violation-of-expectation paradigm have been proposed by other researchers (e.g., Cohen, 2004; Haith, 1998). Wellman (2010) also questioned the assumption that a longer looking time at behaviour A reflects infants' expectations of behaviour B. Although a longer looking time at behaviour A could be interpreted as suggesting that infants do not expect behaviour A, it does not imply that infants expect behaviour B. Hence, the violation-of-expectation paradigm may not be good enough to provide clear evidence for infants' attribution of false beliefs.

To have more confidence about infants' expectation of someone behaving in a particular way, the anticipatory-looking paradigm measures whether infants visually anticipate where a protagonist with a false belief about an object's location will search for the object. Southgate et al. (2007) familiarised 25-month-olds with a bear hiding a ball in one of two boxes while a protagonist was watching; then the protagonist reached towards the box which contained the ball. In the test trial, the protagonist saw the bear hide the ball in one box and then looked away. Meanwhile, the bear moved the ball from its location and took it away. Thus, the ball was no longer in either one of the boxes and infants' looking at either one of the boxes was not due to the ball's location but their expectation of the protagonist's searching location. Results showed that 25-month-olds correctly anticipated the protagonist's searching location. However, this finding could be explained by behaviour rules such as people look for objects where they last saw them (Perner & Ruffman, 2005; Ruffman & Perner, 2005). If this was the case, 25-month-olds did not necessarily attribute to the protagonist a false belief about the ball's location in order to correctly anticipate her searching location.

To conclude, all these infant studies appeared to support the domain-specific account that the concept of mental representations exists early in infancy and may even be innate. However, looking behaviours are so ambiguous that explanations other than an attribution of false belief are possible. Moreover, there is no evidence that young children understand belief as representation and are able to distinguish between reference and sense. As a result, none of the above non-verbal FB tasks which involve least cognitive demands provided compelling evidence for an understanding of false belief in children below age 3.

Given the difficulty in demonstrating a real understanding of false beliefs at the younger end of the developmental span, other researchers have sought to extend the testing population from children to adults. Interestingly, adults who are more skilful in language, executive function and any other cognitive processing than children are not necessarily performing at ceiling on different versions of the FB task, when parametric measures such as probability estimates and error rates are used to assess performance (e.g., Birch & Bloom, 2007; Keysar, Lin, & Barr, 2003). It was found that the need to resist interference from one's knowledge about reality (termed as *cognitive inhibition* hereafter) in the FB tasks is a problem for adults, just as it is for young children.

Other research with brain-damaged adults indicated that an impairment in cognitive inhibition hindered the expression of false-belief understanding (Samson, Apperly, & Humphreys, 2007; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005). Samson et al. (2005) used a non-verbal FB task which followed closely to the traditional location change FB task for 3- to 5-year-olds and showed that a particular patient (known as WBA) who had a brain lesion to the right frontal lobe performed below chance with a strong tendency of committing reality bias, but performed well on memory control trials. WBA was also tested with another non-verbal FB task (Apperly et al., 2004) which was adapted from the one originally devised by Call and Tomasello (1999). In Apperly et al.'s (2004) non-verbal FB task, following Call and Tomasello's, participants are not shown the location of the object. Instead, they are asked to find the object with the help of a clue given by a protagonist who has a false belief about the object's location in the test trials. The protagonist puts a marker on the box where she thinks the object is. Participants thus need to understand that the protagonist's false belief represents the object's location (referent) as the box she puts the marker on (sense) and reason the object's location to be the alterative box as they have witnessed the switch of boxes. As participants do not know where the object really is, they will not be biased. Therefore, no cognitive inhibition is required while they are inferring the protagonist's false belief. The task also involves control trials, similar to those in the screening phase of Call and Tomasello's task, in which participants are tested on incidental executive demands of the task (i.e., response inhibition and working memory). Results showed that the patient, WBA, referred to above, performed above chance on this non-verbal reality-unknown FB task. Taken together, these findings show that WBA was able to understand false beliefs but may not have been able to express this ability due to his impairment in cognitive inhibition.

Another brain-damaged patient, PH, who has severe grammatical impairment after a left hemisphere lesion, performed almost perfectly on the same non-verbal FB tasks used in Apperly et al. (2004) and Samson et al. (2005) and a more complex non-verbal FB task. This finding suggested that general language and grammatical impairments did not diminish performance on false-belief understanding (Apperly, Samson, Carroll, Hussain, & Humphreys, 2006). Having reviewed these studies and other related research on children and adults, Apperly, Samson, and Humphreys (2009) proposed that executive function plays essential roles in the *competence* and *expression* of false-belief understanding, lasting along the developmental span; whereas language, especially grammar, may only be necessary for the *development* of false-belief understanding, showing its effect on children but not adults.

Returning to the competing accounts of mental-representation understanding, would Apperly et al.'s (2009) suggestion speak to either the domain-specific account or the general system account? As the domain-specific account suggests that children's failure on the standard FB task is due to their lack of language and executive skills, it appears that their suggestion supports this account. However, they also suggest that the development of false-belief understanding requires the contribution of language. In other words, a delay in language development may hinder the acquisition of false-belief understanding. For example, Figueras-Costa and Harris (2001) found that deaf children who are born of hearing parents demonstrated a developmental delay in passing a non-verbal FB task which was adapted from Call and Tomasello (1999). As a result, false-belief understanding may not be innate and mature on its own timetable. Despite that, this suggestion does not exclusively support the general system account of mental-representation understanding either. To provide evidence for the general system account, it is essential to illustrate that false-belief understanding relies on a general understanding of representations which develops through experience. This illustration is attempted in the following chapters of this thesis which report and discuss a series of behavioural experiments with typically and atypically developing children.

The human brain: A general-purpose device or a set of specialised components?

In addition to the competing theoretical explanations for the development of mental-representation understanding, the importance of the debate in question also reflects in the long-standing investigation of brain function. For almost two centuries, researchers have been arguing whether the human brain functions as a general-purpose device, in which each region engages in a broad range of cognitive processes, or as a set of specialised components, each of which has a specific mission. One of the first to propose ideas about localisation of brain function is Franz Joseph Gall (1758-1828). Pierre Flourens (1794-1867) and others, on the other hand, argued for distributed cognitive processing throughout the brain. An empirical example for the argument about brain function is the posterior part of the left inferior frontal gyrus described by Broca in 1861 and many later studies as a specialised region for language (e.g., Dronkers, 1996; Penfield & Roberts, 1959; Schäffler, Lueders, & Beck, 1996). However, this region has been recently found to be also recruited in non-linguistic processing, e.g., human action (e.g., Baumgaertner, Buccino, Lange, McNamara, & Binkofski, 2007; Fazio et al., 2009; for a review, see Fadiga, Craighero, & Roy, 2006).

Is there a specific neural mechanism which selectively responds to mental rather than non-mental representations and is independent of language and executive demands? It has been suggested that the medial prefrontal cortex is specifically responsible for distinguishing mental representations from physical representations (for a review, see Frith & Frith, 2003). However, more recently, the importance of the temporo-parietal junction (TPJ) in understanding both mental and non-mental representations has been increasingly stressed (Aichhorn et al., 2009; Perner, Aichhorn, Kronbichler, Staffen, & Ladurner, 2006; Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005). Evidence from brain-damage patients further suggests that the TPJ is more essential for processing mental representations than the medial frontal areas. For example, patients with damage to the medial frontal regions demonstrated impairments in both false-belief reasoning and executive functioning whereas patients with damage to the left TPJ showed deficit in false-belief reasoning only when all of them were tested with non-verbal FB tasks (Apperly et al., 2004; Samson, Apperly, Chiavarino, & Humphreys, 2004). Sommer et al. (2007) also suggested that the activity in the medial prefrontal cortex might reflect the executive demands inherent in the FB task, rather than the representation of mental representation.

A brain imaging study, which tested children aged 8 years and above with both verbal and non-verbal FB tasks, also found more robust brain activity in the TPJ during false-belief reasoning than physical causal reasoning, whereas the medial prefrontal cortex activated more during a baseline condition which did not require reasoning (Kobayashi, Glover, & Temple, 2007). Although the TPJ has consistently been demonstrated to be actively involved in mental-representation processing in both children and adults, it is suggested that the right TPJ is more important for the understanding of false beliefs whereas the left TPJ is associated with both false beliefs and false signs in adults (Aichhorn et al., 2009; Perner et al., 2006). Saxe, Whitfield-Gabrieli, Scholz, and Pelphrey (2009) who examined children aged 6 to 11 years also found a change in response selectivity for mental-related information with age in the right TPJ. Intriguingly, while the right TPJ is proposed to be a specific brain region for adults' mental-representation understanding, other studies have repeatedly shown that the right TPJ is also engaged in a general process of attentional reorientation (e.g., Mitchell, 2008; Serences et al., 2005). As a result, whether there is a specific neural mechanism which is responsible for mental-representation understanding in the brain is still controversial. Although this thesis does not include any brain imaging or neurological studies with brain-injury patients, I further discuss these issues of neural specificity and development in the last chapter of this thesis in relation to the behavioural findings from children with typical and atypical development.

Overview of This Thesis

This thesis aims to address the debate of the domain-specific versus general system account of understanding mental representations because this debate has crucial implications to the theoretical and neurological bases of cognitive development. Reviewing related literature, the understanding of mental representations appears to be a complex cognitive process which involves representational understanding, mentalising, language and executive function. These four factors often confound with each other, making the interpretation of results in this area of research difficult and leaving the debate in question unresolved.

The empirical work for this thesis involved the design and execution of a series of experiments which attempted to disentangle the above four factors in order to investigate whether representational understanding can exclusively explain children's understanding of both mental and non-mental representations. This was achieved by (1) introducing new false non-mental representation tasks, one of which was non-verbal and to which the usual demands of executive function were eliminated or controlled for, and (2) exploring the equivalence between these new false non-mental representation tasks and previously devised false mental representation tasks. The following sections briefly describe the series of experiments that are reported in the following empirical chapters.

In order to determine whether the understanding of mental representations is domain-specific or relies on a general concept of representations, it is not sufficient to find an association between tasks that assess understanding of mental versus non-mental representation. As previously mentioned, this association could be explained by a developmental coincidence between two independent concepts (mental versus non-mental representations) or other shared cognitive skills of language and executive function. Given that correlational studies leave open different interpretations, experimental studies which (1) manipulate children's understanding of mental versus non-mental representations, (2) control the confounding factors of language and executive function, or (3) compare typical with atypical development in understanding the two types of representations may provide a clearer picture.

Training and transferability: Manipulating children's understanding of mental versus non-mental representations.

The first step taken in the empirical work of this thesis is the development and administration of a training study to investigate transfer of understanding based on the following logic. If tasks that assess understanding of mental versus non-mental representations are underpinned by the same concept of representation, it is expected that the tasks are correlated and the effects of training on either one of the tasks are transferable to the other (Slaughter, 1998). If the tasks are supported by two independent concepts which happen to follow a similar developmental time course, an association but no transferability between the tasks are expected. There are several studies that have shown transfer effects between tasks that tested the same concept such as "theory of mind" (Slaughter & Gopnik, 1996), symbolic representation (DeLoache, 1991; DeLoache, Simcock, & Marzolf, 2004; Marzolf & DeLoache, 1994) and inductive reasoning (Klauer, Willmes, & Phye, 2002; see also Klauer &

Phye, 2008, for a meta-analysis). All of these studies involve either training or initial experience on one task and then better performance shown on an untrained but related task or more difficult tasks with similar (near transfer) and different (far transfer) contexts. In this way, training studies have an advantage for examining whether two tasks tap a single concept because they demonstrate a causal relation between training and later measures.

Chapter 2 deals with the equivalence between the FB and FS task by critically examining the transfer of training across tasks. In that chapter, I briefly review several successful training studies on false-belief understanding, and outline a design used by Slaughter (1998) which trained children on the FB or FP task and examined the transferability between the two tasks. I then introduce a new FS task that is matched closely with Slaughter's FB task. By substituting the FP task used in Slaughter's study with the new FS task, a training study is carried out to determine the transferability between the FB and FS tasks. If the FB and FS tasks are transferable, it would suggest that they are supported by the same underlying concept of understanding representations which develops through experience gained from training. On the contrary, it would imply that the FB and FS tasks are not equivalent and support the domain-specific account if the two tasks are not transferable.

Controlling for language and executive function using non-verbal reality-unknown false representation tasks.

The second step is to exclude the possibility that the equivalence between mental and non-mental representation tasks might be explained by other cognitive processes of language and executive function. As mentioned earlier, the FB task confounds the requirement of processing the content of a mental representation with the cognitive demands on language and executive function. The same problem also exists in the FS task as it involves the same language and executive demands as the FB task while tapping the understanding of a non-mental representation. It is thus possible that the association found between the two tasks was due to language and executive function rather than representational understanding. If the FB and FS tasks are shown to be transferable in Chapter 2, it might also be attributed to an increased ability of language and executive function rather than representational understanding after training.

There is only one study that has suggested that language and executive function may contribute to the association between the FB and FS tasks. That is Experiment 2 of Sabbagh, Moses, et al. (2006). They found a marginally significant correlation between the FB and FS tasks when age was partialled out but no significant relations between the two tasks when both age and verbal ability were controlled for, suggesting that language ability might play a role in the relation between the tasks. They also demonstrated that performance on an executive function task, the bear/dragon task (Reed, Pien, & Rothbart, 1984; Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996) which imposes both working memory (remembering a rule) and inhibitory (suppressing an action) demands, predicted performance on the FB and FS tasks. However, the language ability test (i.e., the Peabody Picture Vocabulary Test; Dunn & Dunn, 1981) and the executive function task (i.e., the bear/dragon task) used in this experiment were independent from the FB and FS tasks. The Peabody Picture Vocabulary Test measures receptive vocabulary and the bear/dragon task tests certain aspects of executive function but they do not assess the exact incidental cognitive demands of the FB and FS tasks (e.g., narrative comprehension which involved in both FB and FS tasks is not adequately tested by the Peabody Picture Vocabulary Test). Thus, whether the incidental cognitive demands required in both FB and FS tasks can explain the association between the two tasks is still unclear.

Chapter 3 investigates this issue by employing tasks that disentangle language and executive function from representational understanding. Based on the above review of studies that attempted to disentangle language and executive function from false-belief understanding, Apperly et al.'s (2004) non-verbal reality-unknown FB task appears to be the best task to employ for the following reasons. First, whether the non-verbal FB tasks measuring infants' looking behaviours are tapping a conceptual understanding of false belief which differentiates reference from sense is still controversial. Second, Apperly et al.'s FB task exceeded its ancestor, the original FB task devised by Call and Tomasello (1999), in its standardised and carefully controlled procedure. This involved presenting the task in videos rather than personal manipulation, and having test, control as well as *filler* trials intervene each other rather than screening with control trials before testing. Recall that from previous descriptions of Call and Tomasello's and Apperly et al.'s tasks, only the test trials tap false-belief understanding whereas the control trials were either for screening or controlling participants' ability in meeting incidental demands of the tasks. The intervening filler trials in Apperly et al.'s task provided additional controls on undesirable factors, such as inattention and adoption of incorrect strategies (more details can be found in Chapter 3). Moreover, Call and Tomasello's procedure of having the control trials in the screening phase followed by the test trials in the test phases had a disadvantage: increasing the likelihood of participants learning how to succeed in the test trials through successive exposure in the test phase. The third reason for employing Apperly et al.'s task is that it has been successfully used with older children with fragile X syndrome and intellectual disability (mean age = 13 years 3 months; Grant, Apperly, & Oliver, 2007) although it has not been used with typically developing children in any published study yet. Therefore, Apperly et al.'s FB task appears to be the only task that taps a conceptual understanding of false belief while minimising, eliminating and controlling the language and executive demands of the task.

However, to clearly determine whether children's performance on the FB task is mainly due to their understanding of *representations* rather than their language and executive skills, a comparable *non-mental* representation task is needed. Based on Apperly et al.'s (2004) non-verbal reality-unknown FB task, Apperly et al. (2007) devised a non-mental comparison task: the non-verbal reality-unknown FP task. As mentioned earlier, false photograph is not a good enough comparison to false belief and that false sign might be a better option. Hence, Chapter 3 introduces a novel non-verbal reality-unknown FS task which was designed based on Apperly et al.'s FB and FP tasks. Then, an experiment testing whether this new FS task, relative to Apperly et al.'s FP task, was closely matched to their FB task is reported. A second experiment follows to compare the non-verbal reality-unknown FS and FB tasks with a verbal reality-known version, which is essentially the standard version, of the two tasks. This comparison has two goals: (1) replicating the associations found between the FB and FS tasks in both standard and non-verbal reality-unknown versions within a single experiment, and (2) validating the novel non-verbal reality-unknown FS task with the standard version of the task. As the non-verbal reality-unknown version of the FB and FS tasks minimised and eliminated the demands of language and cognitive inhibition, if the results of these two experiments show that the association between the FB and FS tasks remains in both versions of the tasks, it would suggest that language and cognitive inhibition cannot explain the equivalence between the tasks. On the contrary, if the association only holds in the standard version of the tasks, it would imply that language and cognitive inhibition play a significant role in the association between the tasks.

Autism Spectrum Disorder (ASD): A special case?

The final part of the empirical work for this thesis is to extend the research from typical development to atypical development. So far, the empirical work described has been focused on typical development. It is however insufficient to tell whether understanding mental representations is underlined by a general concept of representational understanding or a domain-specific mechanism which specialises in mental-related processes by looking at typical development only. For example, both mental and non-mental representations may be understood with the same unitary general concept in typical cases but the development of understanding the two types of representations may take two independent pathways in atypical cases. Moreover, cognitive components which are difficult to be factored apart in typical cases may be dissociated in atypical cases. An example of such 'factoring out' in atypical populations was the dissociation of tasks testing the understanding of false beliefs and "false" photographs/drawings shown in individuals with a neurodevelopmental disorder namely Autism Spectrum Disorder (ASD; Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992). As discussed earlier, typically developing children demonstrated similar performance on the FB and "false" photograph/drawing (FP) tasks, but whether they passed the FP task by recruiting an understanding of representation, just as they did on the FB task, or an understanding of correspondence is unclear. By investigating individuals with ASD, the understanding of representation and the understanding of correspondence were found to be dissociated based on the evidence that individuals with ASD failed the FB task but passed the FP task.

Autism was first introduced by Kanner (1943) in his term "early infantile autism" to describe eleven children who were impaired in emotional relationships with others. In 1944, Asperger also reported four children with what he termed "autistic psychopathy" (Frith, 1991). These children had similar behavioural symptoms to those described by Kanner, but with higher functioning (e.g., better language and intellectual abilities). Later, Wing and Gould (1979) suggested that autism could be classified by a triad of impairments in behaviours within three domains: reciprocal social interaction, communication, and restricted, repetitive interests and behaviours. To this day, these behavioural domains are still in use for diagnosis (American Psychiatric Association, 1994; World Health Organisation, 1993). The primary difference between a diagnosis of autism and a diagnosis of Asperger syndrome is that individuals with Asperger syndrome have no history of language and cognitive delay (i.e., general intelligence within the normal range; Bennett et al., 2008; Eisenmajer et al., 1996). Therefore, autism and Asperger syndrome are viewed as situated on a continuum (e.g., Baron-Cohen, 1995; Wing, 1981, 1988), usually known as the *autism spectrum* which ranges from normal to severely autistic behaviours (Wing, 1996). More recently, it has also been suggested that the distinction between autism and Asperger syndrome should be abandoned (e.g., Leekam, Libby, Wing, Gould, & Gillberg, 2000; Mayes & Calhoun, 2003). In this thesis, I thus refer to individuals with diagnoses of autism and Asperger syndrome collectively as ASD.

There have been different theoretical explanations for the behavioural characteristics demonstrated in ASD. The most common one was proposed by Baron-Cohen, Leslie, and Frith (1985) who assessed individuals with ASD with the FB task, suggesting that individuals with ASD have a "theory of mind" deficit which contributes to their difficulties in social functioning. Since then, it has been proposed that individuals with ASD have a specific deficit in their capacity to process mental-related information. In particular, research has indicated that children with ASD performed poorly on tasks that assess false-belief understanding (e.g., Baron-Cohen et al., 1985; Leslie & Frith, 1988; for meta-analyses, see Happé, 1995; Yirimiya, Erel, Shaked, & Solomonica-Levi, 1998) but were at or near ceiling on the "false" photograph/drawing task (Charman & Baron-Cohen, 1992; Leekam & Perner, 1991; Leslie & Thaiss, 1992). However, in a more recent study, children with ASD showed similar and associated performance on the FB task and the false signal task devised by Bowler et al. (2005). Hence, Bowler et al.'s findings place a doubt on the suggestion that children with ASD are specifically impaired in mental-related processing.

In addition, individuals with ASD also show abnormalities in language (e.g., Charman, Drew, Baird, & Baird, 2003; Lord & Paul, 1997; Wilkinson, 1998) and executive function (e.g., Luna, Doll, Hegedus, Minshew, & Sweeney, 2007; Robinson, Goddard, & Dritschel, 2009; for earlier reviews, see Baron-Cohen, Tager-Flusberg, & Cohen, 2000; Hill, 2004). As language and executive function play a role in the FB task, it is possible that the poor performance of children with ASD on the FB task reflects their impairments in language and executive function. Although children with ASD's problem in understanding false beliefs was usually concluded from a comparison with a control group which matched the ASD group on verbal mental age and/or general intellectual ability, these verbal mental age and general intellectual ability were measured by independent tests which did not measure the exact incidental cognitive demands of the FB task as mentioned in earlier section. Given that the false signal task involves the same language and executive demands that matched with the FB task, children with ASD's performance on the false signal task may also be affected by their impairments in language and executive function. As in the case of typically developing children, the association found between the FB task and the false signal task in children with ASD may also be due to language and executive function rather than representational understanding.

Chapter 4 investigates whether ASD is a special case in which individuals are specifically impaired in understanding mental representations but are intact in understanding non-mental representations, independent of their deficits in language and executive function. This is done by testing a sample of children with ASD with the non-verbal reality-unknown FB, FP and FS tasks used in Chapter 3. If children with ASD perform worse in the non-verbal reality-unknown FB task than the non-verbal reality-unknown FP and FS tasks, it would support claims for a specific module for mental-representation understanding in the human brain (e.g., Leslie, 1987, 1994). If children with ASD do not show a specific difficulty on the non-verbal reality-unknown FB task, it would suggest that mental and non-mental representations are understood in a unitary system of representational understanding even in atypically developing children.

In sum, this thesis aims to address the debate of the domain-specific versus general system account of understanding mental representations. Chapters 2 and 3 explore the equivalence between the FB and FS tasks in typically developing children. In chapter 2, I investigate whether training on understanding false beliefs will improve children's understanding of false signs and vice versa. If the understanding of false beliefs and false signs are both supported by a general concept of representation, then training effect on the FB and FS tasks would be expected to transfer back and forth based on other successful training studies. If the understanding of false beliefs and false signs only coincidentally develop in a similar time course, then training effect would not be expected to transfer from one task to the other. In Chapter 3, I examine whether other cognitive processes, such as language and executive function, can explain the equivalence found between the FB and FS tasks. To do this, I design and test a novel non-verbal reality-unknown FS task. If the FB and FS tasks are based on the same underlying understanding of representation, it would be expected that the association between the two tasks cannot be explained by other shared cognitive functions. In Chapter 4, I continue the theme of exploring the equivalence between the FB and FS tasks and test whether the findings with typically developing children in previous chapters can be generalised to children with ASD. If mental and non-mental representations are understood by the same concept of representation, it would be expected to be universal across typical and atypical development. Based on the findings of the series of experiments reported in Chapter 2 to 4, implications to the debate of the domain-specific versus general system account in terms of typical versus atypical development and brain function are further discussed in Chapter 5.

Chapter 2 Training and Transferability in Children's Understanding of False Mental and False Non-Mental Representations

In Chapter 1, I argued that finding an association between a mental representation task and a non-mental representation task is not sufficient to determine that mental and non-mental representations are underpinned by a general concept of representation. The association found between the two types of tasks may be accounted for by a developmental coincidence between two independent concepts, each of which supports one of the tasks. This alternative explanation may be excluded if transferability is found between the two types of tasks in an experimental study which manipulates children's understanding of mental versus non-mental representations through training. As Slaughter (1998) has pointed out, equivalence between two tasks may only be properly inferred if the effect of training on one task transfers to the other. The central goal of Chapter 2 is thus to address the generality claim of mental-representation understanding by critically evaluating the equivalence between the FB and FS tasks using a different approach – that of a training study.

There are a number of training studies which successfully improved children's performance on the FB task with relevant training and controlled procedures (e.g., Appleton & Reddy, 1996; Clements, Rustin, & McCallum, 2000; Lohmann & Tomasello, 2003; Melot & Angeard, 2003; Slaughter, 1998, Study 2; Slaughter & Gopnik, 1996). Children in these studies were first pre-tested and were then randomly assigned to experimental and control training groups. They were thus trained one-to-one on different tasks according to their group assignment, e.g., Slaughter and Gopnik (1996) trained their experimental training groups on tasks testing concepts of belief, desire and perception, whereas their control training group was trained on number conservation tasks. Training usually took place over a certain period of time with at least two sessions. Moreover, appropriate feedback or explanations about the correctness of children's responses to the training tasks were provided in each session. Finally, training effects were indicated by children in the experimental training group who improved their performance on a FB post-test task compared to the control training group. However, enhanced performance on post-test tasks that are identical in structure to training tasks might just be due to task-specific acquired strategies, not an increased understanding of the relevant conceptual skill in question (Knoll & Charman, 2000). Hence, it is important for a training study to demonstrate concept generalisation, meaning that concepts trained with training tasks are generalised to structurally different post-test tasks, e.g., Appleton and Reddy (1996) found that training on location change FB tasks generalised to contents change FB post-test tasks.

Not only can children's performance on the FB task be improved by training, but also this training effect can be transferred to other untrained "theory of mind" tasks (Slaughter & Gopnik, 1996, Study 2). Transfer effects on tasks tapping the same concept, other than "theory of mind", have also been successfully demonstrated in several different studies (e.g., Marzolf & DeLoache, 1994; Klauer et al., 2002). Therefore, Slaughter (1998) suggested that if the FB and FP tasks were supported by a general concept of representation, then a transfer effect should have resulted between the two tasks after training. In fact, her results showed that training on either task did not enhance performance on the other task. This finding supports the argument presented in Chapter 1 that the FB and FP tasks are not conceptually equivalent.

If the FS task rather than the FP task is a better non-mental comparison to the FB task because both FS and FB tasks rely on the same representational capacity of distinguishing between reference and sense, then the FS and FB tasks should be equivalent by being transferable after training. That is, training on the FS task would be expected to enhance performance on the FB task and vice versa. In order to critically test whether the FS and FB tasks are tapping the same underlying concept of representation, I adapted Slaughter's (1998, Study 2) design and replaced her FP task with the FS task. In the following, I briefly describe the design of Slaughter's Study 2 and then use this design as a basis from which to propose a new training study for the investigation of the equivalence between the FB and FS tasks (see Figure 2.1). The main differences between the designs of Slaughter's Study 2 and my new training study are indicated in italics, where the FP task in Slaughter's study is replaced by the FS task, in the figure.

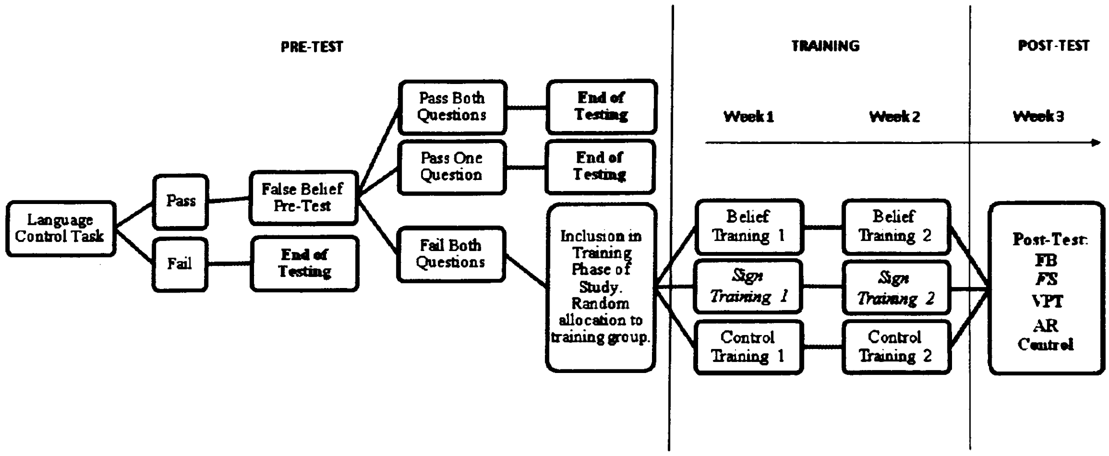


Figure 2.1. Design of a new training study for the false belief and false sign tasks. FB = false belief; FS = false sign; VPT = visual perspective-taking; AR = appearance-reality; Control = number conservation. Italics indicate where new FS task replaces Slaughter's (1998) original FP task.

Slaughter (1998, Study 2) first pre-tested children on a language control task, which was used to control for both children's memory ability and linguistic competence in answering test questions. Then she pre-tested them with a contents change FB task (e.g., a Band-aid box would lead someone to believe that there were plasters inside but it was then shown to contain a book, so the original belief became a false belief). Children who passed the language control task but failed the FB task were included in training. Slaughter had originally intended to pre-test children on the FP task as well and to include those who failed both FB and FP tasks in training, but then abandoned this design because her Study 1 found insufficient children who failed both tasks. Thus, children included in training were not pre-tested on the FP task. They were randomly assigned to three training groups: FB (trained on FB tasks), FP (trained on false photo and false drawing tasks), or Control (trained on number conservation tasks). During training, tasks were presented and simple feedback was given based on children's performance. After two training sessions, one per week, children were post-tested on theory of mind (false belief, appearance-reality, and visual perspective-taking) tasks, FP (false photo and false drawing) tasks, and a number conservation task. Results showed that training on the FB tasks did not lead to an enhancement of performance on the FP tasks and vice versa in comparison to the Control training group, indicating that there was no transfer between the two tasks.

Using a design identical to Slaughter's (1998) Study 2, and replacing her FP set of training and post-test with a FS set of training and post-test (see the italics in Figure

2.1), the new training study reported here investigated the equivalence between the FB and FS tasks. Exactly the same design as Slaughter's was used. This included a pre-test of the language control and FB tasks, a training scheme of two training sessions with explicit corrective feedback, and a post-test of a set of tasks (false belief, visual perspective-taking, appearance-reality and number conservation). There were two reasons for replicating Slaughter's design exactly. First, any difference in results between Slaughter's Study 2 and the new training study would not be attributed to the difference in design, pre-test, training scheme and post-test. If a transfer effect was found in the new training study but not in Slaughter's, this would be because the FS task shared the same underlying skills of representational understanding with the FB task whereas the FP task which Slaughter used did not. Second, another study which used the same design and training scheme also demonstrated a successful training effect in children's false-belief understanding (Slaughter & Gopnik, 1996). With the same design and training scheme, a higher possibility of getting a training effect would be likely in the new training study.

However, before carrying out the new training study, a suitable FS task which matches with Slaughter's (1998) FB task was needed. It was essential for the two tasks to be of the same difficulty and capture the same representational and cognitive demands, with the only exception that the FB task is mental whereas the FS task is non-mental. As any difference in difficulty between the two tasks would confound with the mental nature of the tasks, it would then be impossible to make a clear conclusion regarding the understanding of mental versus non-mental representations if no transfer was found between the tasks. Given that Slaughter used a contents change FB task, a contents change FS task was required.

Previous studies have devised different versions of the contents change FS task including different materials. Sabbagh, Moses, et al.'s (2006) contents change FS task involved two objects and one container (i.e., a cat jumped out of a box with a sign showing a picture of a cat, and then a frog jumped into that box) whilst Leekam et al.'s (2008) task contained two objects and two containers (e.g., a tin with a label of 'cakes' contained cakes and another tin with a label of 'biscuits' contained biscuits, then the cakes and biscuits were switched over). However, Slaughter's (1998) contents change FB task involved only one object and one container (e.g., a book contained in a Band-aid box). Although this difference in numbers of component materials should not make much difference to the cognitive demands of the tasks, it was reasoned that a FS task which is equivalent in every structural aspect to Slaughter's FB task would be more appropriate. Experiment 1 therefore aimed to prepare optimal materials for a new contents change FS task which would match Slaughter's contents change FB task closely. If the two tasks matched well, the newly developed FS task would then be employed in Experiment 2 – the new training study, which investigated whether training on the FB task would improve performance on the FS task and vice versa.

To summarise, this chapter tests the proposal that the developmental difficulty overcome by false-belief passers lies in the ability to represent false representations. Experiment 1 first aimed at establishing a newly developed contents change FS task to match with Slaughter's (1998, Study 2) contents change FB task. Then the newly developed FS task was employed in Experiment 2 to address the critical question of whether the FB and FS tasks are transferable. If the two tasks are transferable, it would suggest that they are supported by the same underlying concept of understanding representations. On the contrary, if they are not transferable, it would imply that they are not equivalent and further support the domain-specific account. Ethics approval which applied to both Experiment 1 and 2 was obtained from the Ethics Advisory Sub-Committee in the Department of Psychology in Durham University.

Experiment 1

A newly developed contents change FS task with different sets of materials was compared with Slaughter's contents change FB task to test whether the two tasks were matched and to find out the best sets of materials for the new FS task to match closely to those used in Slaughter's FB task. Following Slaughter's study, children had to pass a language control task in order to be included in the analyses. Given that the FB and FS tasks are the same in every aspect, except the test questions, and previous studies (Bowler et al., 2005; Leekam et al., 2008; Sabbagh, Moses, et al., 2006, Experiment 2) had found no significant difference but instead significant correlations between the two types (FB and FS/false signal) of tasks, it was expected that children would perform similarly and their performance would be correlated between the FB and FS tasks. Children aged 3 to 5 years were recruited due to the universal developmental pattern found between this period of age on the FB task (e.g., Wellman et al., 2001) and a similar pattern demonstrated on the FS task (e.g., Leekam et al., 2008). Based on these previous studies, older children were expected to perform better than younger children on both FB and FS tasks.

Method.

Participants.

Twenty-eight typically developing children (aged 3;4 – 5;2; M = 52.25 months; SD = 7.43) were recruited from a school in North East England. Fourteen were aged 3;4 – 4;5 (M = 45.79 months; SD = 4.34) and another 14 were aged 4;6 – 5;2 (M = 58.71 months; SD = 2.43). Three additional children, aged 3;3, 3;4, and 3;7, were tested but they failed the language control task so they were excluded from the analyses. All children were White British (representative of the geographical area) and the population of the school was of low socio-economic status. Parental informed consent was obtained before testing.

Design.

This was a mixed design, testing younger versus older children with both FB and FS tasks. Each child was firstly presented with the language control task used by Slaughter (1998); then he/she was tested with three trials of the FB task and three trials of the FS task. The purpose of giving three trials was to gain more reliable data for each child as each trial involves a single pass/fail response. The presentation orders of the tasks and the trials within each task were counterbalanced.

Materials and procedure.

Children were individually tested in a session which lasted approximately 10 minutes by an experimenter. Slaughter's (1998, Study 2) contents change FB task which used two sets of materials (a "Band-aid" box with a book inside and a "crayon" box with candles inside) were employed in this experiment. In order to have three trials for each of the FB and FS tasks, one more set of materials (a "Smarties" box with a pencil inside) was added to the FB task and three other sets of materials (a "Coca-Cola" bottle with a toothbrush inside, a "Fruit Shoot" bottle with a paintbrush inside and a "Lego" box with a bag of crisps inside) were used for the new FS task. A summary of materials employed from Slaughter's Study 2 and new materials added for the current experiment is listed in Table 2.1. Two out of the three sets of materials of the FS task that were the most comparable to Slaughter's "Band-aid" and "Crayon" contents change FB task would then be employed in Experiment 2. All of the materials used were familiar to the preschool children in England.

Table 2.1

Materials Employed From Slaughter's (1998) Study 2 and new Materials (in Italics) Added for Experiment 1

Slaughter's (1998) Study 2	Experiment 1
Language control task	
Dolls house, orange and toy car	
Contents change FB task	
Set 1: "Band-aid" box and book	
Set 2: "Crayon" box and candles	
	Set 3: "Smarties" box and pencil
	Contents change FS task
	Set 1: "Coca-Cola" bottle and toothbrush
	Set 2: "Fruit Shoot" bottle and paintbrush
	Set 3: "Lego" box and a bag of crisps

Following Slaughter (1998), the experiment started with an introduction of a doll called Charlie who was described as a friend who liked to help but was sleepy so was put to sleep under the table and would be woken up for help later. The language control task was the same as that in Slaughter's study. It contained a dolls house with an orange, which was then replaced by a toy car. Children were asked, "What was inside this house when you first saw it?" The criterion for passing was replying "Orange" to the question. Then the FB task and the FS task were presented. The presentation order of the tasks was counterbalanced. The procedure for the FB task was also the same as Slaughter's but with two exceptions: (1) Before asking the test question, the experimenter tried and failed to wake up Charlie so he stayed asleep under the table; and (2) One test question, instead of two ('self' and 'other' question), was asked, "What will Charlie think is inside this box when he first sees it?" (i.e., the 'other' question). The purpose of the first exception was to ensure that children did not think that Charlie overheard their conversation with the experimenter, whereas the second exception was made to match the FB task with the FS task which could only have one test question. The reason for choosing the 'other' question, asking about Charlie's belief, over the 'self' question is that the initial contents change FB task

(Hogrefe et al., 1986) described in Chapter 1 only asked a question about another person's belief and the procedure of introducing a doll at the beginning could be kept the same as that in Slaughter's for a rational purpose. The FS task was made similar to the FB task except that the test question was "What does the label on the bottle show that is inside?" Verbatim instructions used for the language control, FB and FS tasks are listed in Table 2.2.

Table 2.2

Verbatim Instructions Used for the Language Control, False Belief and False Sign Tasks

Language control task
Participants were shown a dolls house with an orange inside. "What is inside this house?" *The orange was removed and a toy car was put in the dolls house.* "Now what is inside this house?" "What was inside this house when you first saw it?"

False sign	False belief
Participant was shown a container. "What do you think is inside this box?"	*Participant was shown a container.* "What do you think is inside this box?"
The box was opened to reveal an unexpected content.	*The box was opened to reveal an unexpected content.*
"Look there is ......" *Requested that the participant say what's inside. If s/he didn't, told him/her and asked him/her to repeat.*	"Look there is ......" *Requested that the participant say what's inside. If s/he didn't, told him/her and asked him/her to repeat.*
The box was returned to its original state. "Now what do you think is inside this box?"	*The box was returned to its original state.* "Now what do you think is inside this box?"
Test Question: "What does the label on the box show that is inside?"	*Test Question:* *Charlie was not able to be awakened.* "What will Charlie think is inside this box when he first sees it?"

Results.

Children received a score of 1 for each of the trials if they correctly reported Charlie's false belief for the FB task and the container's false label for the FS task. As the scores were either 0s or 1s, non-parametric tests were used to examine whether there were differences between the sets of materials used within each task. There was no significant difference in children's performance across the three themes within both FB and FS tasks, Friedman test: $\chi^2(2, N = 28) = 3.43, p = .23$ and $\chi^2(2, N = 28) = 1.20, p = .85$ respectively.

Children's scores were thus collapsed across the three themes for each of the FB and FS tasks (range of scores for each task: 0 – 3). The means and standard deviations for the two tasks are presented in Table 2.3. The data were not normally distributed (the values of skewness were positive whereas the values of kurtosis were negative). Data transformations, including square root, log and inverse transformations, did not improve normality. However, the outcomes of the following parametric analyses were corroborated by running non-parametric analyses. First, children's performance across tasks was examined with a between-participants factor of age. It was expected that performance on the two tasks would not differ but that older children's performance would be better than younger children's. It was also expected that the two tasks would be correlated so correlation tests were then performed to test this prediction. Finally, which sets of materials of the FS task that matched best with the "Band-aid" and "Crayon" sets of materials of the FB task used in Slaughter (1998) were determined by Pearson's chi-square tests as the range of scores children got for each set of materials of each task was 0 – 1.

Table 2.3

Means and Standard Deviations for the False Belief and False Sign Tasks (Range of Each: 0 – 2)

		False belief		False sign	
Age group	*n*	*M*	*SD*	*M*	*SD*
Younger	14	0.57	1.09	0.64	1.08
Older	14	2.21	1.05	2.36	1.15

Older children performed above chance, $ts(14) > 2.54, ps < .05$ with effect sizes of $rs > .58$, whereas younger children performed below chance, $ts(14) < -2.96, ps < .05, rs > .40$, on both of the FB and FS tasks. A 2 (age) x 2 (task) mixed-design analysis of variance (ANOVA) with task as a within-participants factor revealed a significant main effect of age, $F(1, 26) = 22.85, p < .001$ with an effect size of $r = .68$, but no effect of task and no interaction, $Fs(1, 26) < .24, ps > .63$.

Equivalent non-parametric analyses showed identical results. Wilcoxon Signed Ranks tests showed that older children performed above chance, Wilcoxon *Zs* < -2.18, *ps* < .05 with effect sizes of *rs* < -.41, whereas younger children performed below chance, Wilcoxon *Zs* < -2.30, *ps* < .05, *rs* < -.43. Mann-Whitney U tests indicated similar age difference on the FB and FS tasks, $U = 30$, $p < .001$ with an effect size of $r = -.63$ and $U = 34.50$, $p < .01$, $r = -.60$ respectively. No difference between the FB and FS tasks was also found by Wilcoxon Signed Ranks test, Wilcoxon $Z = -.58$, $p = .58$. Both Pearson and Spearman's correlations showed that performance on the FB task was correlated with performance on the FS task, $r(28) = .66$, $p < .001$ and $r_s(28) = .63$, $p < .001$ respectively.

Pearson's chi-square tests were used to determine which sets of materials of the FS task matched better with the "Band-aid" and "Crayon" sets of materials of the FB task. The "Band-aid" set of materials of the FB task was associated with both "Coca-Cola" and "Fruit Shoot" sets of materials of the FS task, $\chi^2(1, N = 28) = 5.07$, $p = .06$ with a phi-coefficient of .43 and $\chi^2(1, N = 28) = 5.32$, $p < .05$ with a phi-coefficient of .44 respectively, but not with the "Lego" set of materials of the FS task, $\chi^2(1, N = 28) = 1.29$, $p = .45$ with a phi-coefficient of .22. The "Crayon" set of materials of the FB task was associated with the "Coca-Cola", "Fruit Shoot" and "Lego" sets of materials of the FS task, χ^2s$(1, N = 28) > 7.04$, *ps* < .05 with phi-coefficients > .50. As the "Lego" set of materials of the FS task was not associated with the "Band-aid" set of materials of the FB task, it was not employed in Experiment 2.

Discussion.

Experiment 1 aimed at establishing a newly developed contents change FS task to match with Slaughter's (1998, Study 2) contents change FB task so that the new FS task could be employed in Experiment 2. Results showed that the new FS task matched well with Slaughter's FB task, *rs* = .63 – .66. These correlation coefficients matched with those in previous studies (*rs* = .50 – .88; see Perner & Leekam, 2008, for a review). Moreover, both of the tasks were sensitive enough to capture the developmental change of understanding false belief and false sign demonstrated in previous studies. Older children were found to perform above chance whilst younger children performed below chance on both of the tasks.

This experiment also found out the best sets of materials for the new FS task to match closely to those used in Slaughter's FB task. As the "Lego" set of materials of the FS task was associated with the "Crayon" but not the "Band-aid" set of materials of the FB task, the "Coca-Cola" and "Fruit Shoot" sets of materials of the FS task were the best matches to the "Band-aid" and "Crayon" sets of materials of the FB task and were thus employed in Experiment 2. Previous studies which have successfully demonstrated the correlation between the FB and FS or false signal tasks (Bowler et al., 2005; Sabbagh, Moses, et al., 2006, Experiment 2) also used completely different sets of materials for the FB versus FS or false signal tasks. Thus, no critical influence to the results was expected even though the FS task used "Coca-Cola" and "Fruit Shoot" whereas the FB task used "Band-aid" and "Crayon" in Experiment 2.

Experiment 2[1]

The main aim of Experiment 2 was to critically address the issue of equivalence between the FB and FS tasks by investigating whether the two tasks were transferable after training. Children, who were different from those in Experiment 1 but of the same age range of 3 to 5 years, were screened for inclusion in training, following Slaughter's (1998, Study 2) procedure. This involved giving children pre-tests, including the language control task (see Table 2.2) and the contents change FB task (see Table 2.4), as described in Slaughter's study. The FB pre-test task involved two test questions: the 'self' and 'other' questions. Employing the same inclusion criterion as that used in Slaughter's study, the current experiment included only children who passed the language control task but failed both of the test questions of the FB pre-test task into the training phase.

For the training part of the experiment, children who met the inclusion criterion were randomly assigned to three training groups: False belief, False sign and Control. The FB training group received training on a FB task, the FS training group received training on a FS task, and the Control training group received the same training as that used in Slaughter (1998, Study 2), namely a number conservation task. With the Control training group, baseline performance could be obtained for comparing the training effect against development across time on understanding representations. During the training session, children were presented with one of the relevant tasks and given positive or negative feedback based on their performance. Finally, children were tested on the four post-test tasks used in Slaughter's study (false belief, visual

[1] Experiment 2 has been accepted to be published in the *Journal of Cognition and Development*.

perspective-taking, appearance-reality and number conservation) plus the new FS task developed in Experiment 1. As the FB task involved two test questions whereas the FS task involved only one test question, performance on the 'self' and 'other' questions of the FB task were examined separately in order to match with the FS task. If transferability was found, this experimental design would also enable the investigation of the potential mechanism underlying the training effect of representational understanding in children.

Table 2.4

Verbatim Instructions Used for the False Belief Pre-Test Task

False belief pre-test task
Participant was shown a container. "What do you think is inside this box?" *The box was opened to reveal an unexpected content.* "Look there is ......" *Requested that the participant say what's inside. If s/he didn't, told him/her and asked him/her to repeat.* *The box was then returned to its original state.* "Now what do you think is inside this box?" *Test questions:* Self: "What did you think was inside this box when you first saw it?" *Charlie was not able to be awakened.* Other: "What will Charlie think is inside this box when he first sees it?"

In addition to the main aim, there was a subsidiary goal which was to demonstrate that concept generalization was another outcome of training. This issue was addressed using different tasks in the training phase from those used in Slaughter's (1998) Study 2. Location change tasks were used in training whereas contents change tasks were used in the pre- and post-tests. A previous study (Appleton & Reddy, 1996) found an improved performance on the contents change FB post-test tasks after training children on the location change FB tasks, hence similar concept generalisation was expected in the current experiment. However, whether concept generalisation would also be found for the FS training group was not known. Children in this training group were trained on a location change FS task which used an arrow as the representational medium (e.g., an arrow pointing to a box which contained sweets; see also Table 2.5), whereas the contents change FS post-test task used a label as the representational medium (e.g., the "Coca-Cola" label on a "Coca-Cola" bottle; see also Table 2.6). Whether training on arrows would be generalised to labels was not yet established in literature. According to Leekam et al. (2008) and Sabbagh, Moses, et al. (2006, Experiment 2), no significant differences were found between their location change and content change versions for both FB and FS tasks; thus it was predicted that concept generalisation would also occur for the FS training group in the current experiment. Finally, it was also important to examine whether age would have an effect on training and subsequent post-test performance based on the findings in previous studies and Experiment 1 that older children performed better on both FB and FS tasks than younger children. The assumption is that older children perform better on false representation tasks because they possess a conceptual understanding of representation whereas younger children do not. However, there are also individual differences in performance across age and one would expect experience through training to have a similar effect on children without a concept of false representation regardless of age. In the case of the current experiment, children were screened for false-belief understanding. Therefore, none of them would have possessed the conceptual understanding of representation before training and age was unlikely to have an effect on the training and post-test performance.

To conclude, the current experiment intended to examine the proposal that understanding of false beliefs and false signs depend on a general underlying ability of representing false representations. Under this proposal the following predictions could be made: (1) No significant difference but an association would be found between performance on the FB and FS post-test tasks across training groups, replicating the results of Experiment 1; and (2) A transfer effect would be found after training. That is, when compared to the Control training group, the FB training group would show better performance on the FS post-test task and the FS training group would show better performance on the FB post-test task. In addition, the experiment also tested for concept generalisation between training and post-test by using different versions of the FB and FS tasks involving location and contents change at each phase. If the trained concept could be generalised, training on the location change task should enhance performance on the contents change post-test task for the FB and FS training group in comparison to the Control training group. Finally, the age effect on training and post-test performance was also investigated with the expectation that age would not be associated with how well children responded to the training and performed in the post-test.

Method.

Screening phase.

166 typically developing children (aged 3;3 – 5;11) were recruited from three schools in the northeast of England. None of them had participated in Experiment 1. The current experiment followed directly Slaughter's (1998) design and procedure, by first screening children on two tasks to assess inclusion in training. Inclusion depended on passing the language control task and failing the contents change FB pre-test task. The language control task, which was exactly the same as that in Experiment 1, was always administered first and then followed by the FB pre-test task. The FB pre-test task involved two deceptive boxes (a Band-aid box and a crayon box) with unexpected contents (a book and some candles respectively) inside. Half of the children received the "Band-aid" set of materials whereas the other half received the "Crayon" set of materials. There were two test questions. For the 'self' question, children were asked, "What did you think was inside this box when you first saw it?" For the 'other' question, the experimenter asked, "What will Charlie think is inside this box when he first sees it?" Charlie, same as in Experiment 1, was earlier identified as a doll asleep under the table. The experimenter tried but failed to wake Charlie up before she asked the 'other' question. The presentation order of the two test questions was counterbalanced. Verbatim instructions for the FB pre-test task are listed in Table 2.4. If children passed the language control task but failed both of the FB pre-test questions (i.e., both 'self' and 'other' questions were incorrect), they were included in training.

Although the ideal design would be to include children who failed both FB and FS tasks in training, it was reasoned that this goal would be difficult to achieve as Slaughter's (1998) herself found insufficient children who failed both mental (FB) and non-mental (FP) tasks. As a result, the design of Slaughter's study was followed exactly by pre-testing children on the language control and FB tasks only, and including those who passed the language control task but failed both questions of the

FB task. One hundred and three (62%) of the children recruited did not qualify for inclusion in training[2]. As a result, 63 children were included. The percentage of children that were included in training (38%) was similar to that of Slaughter's Study 2 which was 35%. However, there were 6 children who were absent from school either on the day of one of the training sessions or post-test so they were not included in the following analysis, resulting in 19 children in each training group (which was almost twice the number of the children in Slaughter's Study 2).

Main experiment: Training and post-test.

Participants. Fifty-seven children were randomly assigned to one of the three training groups: FB, FS or Control. The age range of the 19 children in the FS training group was 3;3 – 5;3 (M = 52.89 months; SD = 7.61), the FB training group's age range was 3;7 – 5;3 (M = 52.58 months; SD = 6.41), and the Control training group's age range was 3;11 – 5;4 (M = 55.58 months; SD = 5.80). A one-way ANOVA showed no significant difference in terms of age, $F(2, 54) = 1.22$, $p = .31$, between the three training groups. The sample was White British and the populations of the three schools were generally of low socio-economic status. Parental informed consent was obtained before the pre-test.

Design. This was a mixed-design with training group as a between-participants factor and post-test task as a within-participants factor. Following Slaughter's (1998) Study 2, the training phase consisted of two sessions, one session per week. During training, children were given the relevant tasks and positive or negative feedback according to their responses. A post-test of five tasks was then administered a week after the last training session. The design of the experiment is shown in Figure 2.1.

Materials and procedure. Children were trained and tested individually in a room in their own school by the same experimenter who also administered the pre-test. However, the experimenter was unaware of the child's training condition when the post-test session was conducted. She did not have any information in front of her

[2] Thirty-two children failed the language control task; 22 children passed one of the two questions in the FB task, 8 of them passed the 'self' question whereas 14 passed the 'other' question; 28 children passed both questions in the FB task; and 21 children failed to answer the prompt question "Now what do you think is inside this box?" correctly but passed one or both questions in the FB task. Although there were another 9 children who also answered the prompt question incorrectly and should normally be excluded, they failed both questions in the FB task as well so they were considered as not possessing the understanding of false belief yet, which was in fact the most important criterion for inclusion, and were therefore included in the training phase.

at time of post-test that would identify the child's training condition. Moreover, she was not able to recall this information due to the complexity of the design and number of children participated. An assistant was present to help and record children's responses.

There were two training sessions: one was scheduled within 1 month after children's completion of the pre-test and the other was 6 – 9 days after the first training session. During each training session, children were presented with one of the relevant tasks according to their group assignments (FB, FS or Control) and positive or negative feedback was given based on their answers to test questions (verbatim in Table 2.5). The FB and FS training group was trained on the location change tasks and the Control training group was trained on the number conservation task. The experimenter's feedback usually marked the end of each training session. However, there were a few exceptions when children disagreed with the feedback by opening the box to show the experimenter that the objects were actually there. The experimenter responded with a smile and said, e.g., "The arrow shows us that the sweets are in this box (initial location)". Each training group of children was shown two sets of materials (i.e., 'shop with sweets in one of two boxes' and 'kitchen with bananas in one of two boxes' for the FS and FB training group; and 'blocks' and 'candles' for the Control training group); one set of materials in each training session, and the order of presentation was counterbalanced across children. Only the FB training group involved a protagonist (doll) called Maxi. After two sessions of training, children were post-tested.

Table 2.5

Verbatim Instructions Used in one of the two Sessions for Each of the Three Training Groups: False Sign, False Belief and Control

False sign	False belief	Control
This is a shop. The shop sells sweets. The sweets are in this box (*point*). And, look, there is an arrow showing us where the sweets are. *Prompt:* Where are the sweets?	This is a shop. This is Maxi. The shop sells sweets. Maxi shows you that the sweets are in this box (*point*). *Prompt:* Where are the sweets?	*Participant was presented with two rows of 10 small blocks (row A and B). They were manipulated so that the rows were first equal lengths.*
Now, we put the sweets in the other box. *Prompt:* Where are the sweets now?	Maxi goes home. While Maxi is at home, we put the sweets in the other box. *Prompt:* Where are the sweets now? Now, Maxi wants some sweets. He returns to the shop in order to buy some sweets.	*One row was made longer (half of the participants had row A longer, whereas the other half of the participants had row B longer), then the rows were re-equalized, and then the other row was made longer.*
Test question: Where does the arrow show us that the sweets are?	*Test question*: Where does Maxi think the sweets are?	*Test question:* Does one of these rows have more or are they the same?
Positive feedback: "Yes, that's right. The arrow shows us that the sweets are in this box (*initial location*)."	*Positive feedback:* "Yes, that's right. Maxi thinks the sweets are in this box (*initial location*)."	*Positive feedback:* "Yes, that's right. They are the same."
Negative feedback: "No, the arrow shows us that the sweets are in this box (*initial location*)."	*Negative feedback:* "No, Maxi thinks the sweets are in this box (*initial location*)."	*Negative feedback:* "No, they are the same."

The post-test was presented 5 – 8 days after the second training session. The doll, Charlie, was introduced and put to sleep under the table again. Children were then given five tasks: false sign, false belief, appearance-reality, visual perspective-taking and number conservation (verbatim in Table 2.6; both FS and FB tasks were of the contents change version, same as those in Experiment 1). The presentation order of the tasks was counterbalanced, with the FB and appearance-reality (AR) tasks paired because they employed the same materials (i.e., the "Band-aid" and "Crayon" sets of materials used in Experiment 1 and the pre-test of the current experiment; see Table 2.6). Half of the children had the FB test questions asked first whilst the other half of the children had the AR test questions asked first after the real contents of the deceptive boxes were revealed. Both FB and AR tasks had two test questions: the 'self' and 'other' questions in the FB task (possible range of scores was 0 – 2); the 'appearance' (What does it look like is inside this box?) and 'reality' (What is really and truly inside this box?) questions in the AR task. Children had to answer both 'appearance' and 'reality' questions correctly in order to be scored as 1 (range: 0 – 1) for the AR task. The presentation order of the two test questions in each of the tasks was also counterbalanced. The FB task was exactly the same as the FB pre-test task except that the set of materials used was different. Children who received the "Band-aid" set of materials in pre-test would receive the "Crayon" set of materials in post-test, whereas the other children received the reverse ordering. Although the FB and AR tasks were administrated in pairs, children's performance on the two tasks was analysed separately.

Table 2.6

Verbatim Instructions Used in Each of the Post-Test Tasks: False Sign, False Belief, Appearance-Reality, Visual Perspective-Taking and Number Conservation

False sign	False belief & appearance-reality	Visual perspective-taking	Control
Participant was shown a bottle of Fruit Shoot / Coca-Cola. "What do you think is inside this bottle?"	*Participant was shown a Bandaid / Crayon box.* "What do you think is inside this box?"	*Participant was shown a straight stick.*	*Participant was presented with two rows of 5 pennies in equal lengths.*
The bottle was opened to reveal a paintbrush / toothbrush. "Look there is ......" *Requested that the participant say what's inside. If s/he didn't, told him/her and asked him/her to repeat.*	*The box was opened to reveal a book / some candles.* "Look there is ......" *Requested that the participant say what's inside. If s/he didn't, told him/her and asked him/her to repeat.*	*Then the stick was put into a glass of water, making it appear to be bent.*	*One row was made longer, then the rows were re-equalized, and then the other row was made longer.*
The bottle was then returned to its original state. "Now what do you think is inside this bottle?"	*The box was then returned to its original state.* "Now what do you think is inside this box?"		
Test question: "What does the label on the bottle show that is inside?"	*Test questions:* "What did you think was inside this box when you first saw it?" *Charlie was not able to be awakened.* "What will Charlie think is inside this box when he first sees it?" "What does it look like is inside this box?" "What is really and truly inside this box?"	*Test questions:* "What kind of stick did you see before we put it in water, a bent stick or a straight stick?" *Charlie was not able to be awakened.* "What kind of stick will Charlie see if we take it out of the water, a bent stick or a straight stick?"	*Test question:* "Does one of these rows have more or are they the same?"

The FS post-test task was the same as that in Experiment 1 but it involved only two sets of materials: a bottle of Coca-Cola with a toothbrush inside and a bottle of Fruit Shoot with a paintbrush inside. These two sets of materials for the FS task have been demonstrated to work appropriately compared to the "Band-aid" and "Crayon" sets of materials of the FB task in Experiment 1 (phi-coefficients > .43). Half of the children received "Coca-Cola" and the other half received "Fruit Shoot". The test question was "What does the label on the bottle show that is inside?" so the possible range of scores for this task was 0 – 1.

For the visual perspective-taking (VPT) post-test task, children were presented with a straight stick and a glass of water. Two test questions ('self' and 'other' questions) were asked after the stick was put into the glass of water (range of scores: 0 – 2). The presentation order of the two test questions was again counterbalanced. For the number conservation (NC) post-test task, two rows of five pennies were presented in front of the children and the transformation of the two rows was performed in the same way as in training (range of scores: 0 – 1).

Results.

As most of the variables were nominal, analyses were restricted to non-parametric tests. Children's performance on the FS and FB post-test tasks across training groups was first examined with within-participants analyses to check for any difference and association between the tasks. To examine whether the tasks are transferable, the post-test performance of the three training groups was compared with between-participants analyses to test whether the FB training group would show better performance on the FS post-test task and vice versa when comparing with the Control training group; and performance change through training to post-test was also investigated. Finally, age was subjected to correlation tests to examine whether it affected children's receptiveness of training and post-test performance.

Relation between the performance of the FS and FB post-test tasks.

Performance on each of the post-test tasks is shown in Table 2.7. If the FS and FB tasks were equivalent, children should perform similarly on both of the tasks. Two related-samples tests indicated that children's performance on the FS task did not significantly differ from their performance on both 'self' and 'other' questions of the

FB task, Wilcoxon $Z = -1.61, p = .17$ and Wilcoxon $Z = -1.34, p = .26$ respectively. Moreover, Chi-Square tests showed that performance on the FS task was related to the performance on the 'self' and 'other' questions of the FB task, $\chi^2(1) = 5.56, p < .05$ with a phi-coefficient of .31 and $\chi^2(1) = 4.31, p < .05$ with a phi-coefficient of .28 respectively. Hence, the first prediction that no difference but an association would be found between performance on the FS and FB post-test tasks was supported. Furthermore, this finding replicated those in Experiment 1.

Table 2.7

Number and Percentage of Children With Correct Scores on Each of the Post-Test Tasks

		Post-test task						
Training group	*n*	FS	FB - self	FB - other	VPT - self	VPT - other	AR	NC
FS	19	10 (53%)	9 (47%)	5 (26%)	17 (89%)	12 (63%)	4 (21%)	11 (58%)
FB	19	12 (63%)	5 (26%)	11 (58%)	17 (89%)	11 (58%)	7 (37%)	12 (63%)
Control	19	3 (16%)	4 (21%)	3 (16%)	17 (89%)	13 (68%)	5 (26%)	14 (74%)

Note. FS = False Sign; FB - self = the 'self' question of the False Belief task; FB - other = the 'other' question of the False Belief task; VPT - self = the 'self' question of the Visual Perspective-Taking task; VPT - other = the 'other' question of the Visual Perspective-Taking task; AR = Appearance-Reality; and NC = Number Conservation.

Post-test performance on all tasks.

Scores of the FS and FB post-test tasks and performance of the FS and FB training groups were collapsed to create a single false representation measure (range: 0 – 3) and a single False Representation training group respectively because children's performance on the two tasks were not significantly different and were related. A Mann-Whitney U test revealed a significant difference between the False Representation training group and the Control training group on the false representation measure ($U = 192.5, p < .01$ with an effect size of Cohen's $d = .91$), suggesting that children's understanding of false representations was enhanced when they were trained on false representation tasks in comparison to the NC task. However, no significant difference was found between the False Representation training group and the Control training group on the NC post-test task ($U = 313.5, p =$

.39). Training of the NC task did not enhance children's performance on the NC post-test task relative to those who had not been trained on it.

To further examine the effects of different training groups on relevant post-test tasks and whether the FS and FB tasks are transferable, the post-test performance of the three training groups was compared using a series of Chi-Square tests. First, there was a significant difference between the performance of the training groups on the FS post-test task, $\chi^2(2) = 9.55, p < .01$. The performance of both FS and FB training groups was significantly better than the Control training group on the FS post-test task, $\chi^2(1) = 5.73, p < .05$ with an effect size of phi (ϕ) = -.39 and $\chi^2(1) = 8.92, p < .01, \phi = -.49$ respectively, whereas the performance of the FS and FB training groups on the FS post-test task did not differ, $\chi^2(1) = .43, p = .74$. Second, there was no significant difference between the performance of the training groups on the 'self' question of the FB post-test task, $\chi^2(2) = 3.41, p = .28$. However, a significant difference was found between the performance of the training groups on the 'other' question of the FB post-test task, $\chi^2(2) = 8.21, p < .05$. The performance of the FB training group was significantly better than the Control training group on the 'other' question of the FB post-test task, $\chi^2(1) = 7.24, p < .05, \phi = -.44$, whereas the performance of both FB and Control training groups did not differ from the FS training group on that question, $\chi^2(1) = 3.89, p = .10$ and $\chi^2(1) = .63, p = .69$ respectively. Finally, there were no significant differences between the performance of the training groups on the other post-test tasks, $\chi^2(2) = .00, p = 1$ for the 'self' question of the VPT post-test task, $\chi^2(2) = .45, p = .94$ for the 'other' question of the VPT post-test task, $\chi^2(2) = 1.22, p = .66$ for the AR post-test task and $\chi^2(2) = 1.08, p = .69$ for the NC post-test task.

Hence, training using a location change version of FS and FB tasks enhanced performance on the relevant post-test task which involved contents change. That is, the FS training group trained on the location change FS task performed better on the contents change FS post-test task than the Control training group; and the FB training group trained on the location change FB task performed better on the 'other' question of the contents change FB post-test task than the Control training group. The FB training group also outperformed the Control training group on the FS post-test task, confirming half of the second prediction that the effect of false-belief training transferred to false-sign understanding. Children's performance change through training to post-test was further investigated in order to reveal whether false-sign training had some contribution to later false-belief understanding within participants and to shed light on the underlying mechanism of training representational understanding in children.

Performance change through training to post-test.

Performance on each of the training tasks for each training session and relevant post-test is shown in Table 2.8. Improvement was shown from training session 1 to post-test for the FS and Control training groups, Cochran's Q (2) = 8.17, $p < .05$ and Cochran's Q (2) = 8.6, $p < .05$ respectively. However, the increase in performance for the FB training group from training session 1 to post-test did not reach significance, Cochran's Q (2) = 5.17, $p = .08$. A similar result was also found in Slaughter (1998, Study 2). She suggested two possible explanations: (1) children benefited from the pre-test so that they had acquired some understanding of false belief already by the first training session; (2) her training tasks were somewhat easier than the pre- and post-tests. The data obtained in the current experiment supported Slaughter's first explanation. First, a significant difference was found when comparing children's performance on the pre-test and training session 1 for the FB training group, $\chi^2(1) = 4.17, p < .05, \phi = .47$. For the FS training group, children's performance on the pre-test and FB post-test, which was their second FB task, was significantly different for the 'self' question, $\chi^2(1) = 7.11, p < .01, \phi = .61$, and marginally different for the 'other' question, $\chi^2(1) = 3.20, p = .06, \phi = .41$. However, no significant difference was found between children's performance on the pre-test and FB post-test for the Control training group, $\chi^2(1) = 2.25, p = .13$ for the 'self' question and $\chi^2(1) = 1.33, p = .25$ for the 'other' question, but this might be due to the delay from the pre-test to post-test which could have washed out the potential practice effect. Despite the delay from the pre-test to post-test, a significant difference was found for the FS training group, suggesting that training on the FS task contributed to children's understanding of false belief, especially for the 'self' question. Second, the FS and Control training groups' performance on the 'other' question of the FB post-test task was not significantly different from the FB training group's performance on their first training session, which involved an 'other' question, $\chi^2(1) = 0.13, p = 1$ and $\chi^2(1) = 1.31, p = .45$ respectively. Taken together, these findings indicated that practice on the pre-test without feedback might be sufficient to increase later (short delay: no longer than a month) performance on the related task and the FB training task seemed to be of similar difficulty to the FB post-test task.

Table 2.8

Number and Percentage of Children With Correct Scores on Each of the Training Tasks for Each Training Session

Training group	n	Training session 1	2	Post-test
FS	19	7 (37%)	15 (79%)	10 (53%; false sign)
FB	19	6 (32%)	12 (63%)	11 (58%; false belief - 'other' question)
Control	19	8 (42%)	15 (79%)	14 (74%; number conservation)

This consolidating effect of extra practice on the FB pre-test might have enhanced the transfer effect from false-belief training to false-sign understanding. If children had been given a pre-test which included a FS task as well as a FB task so that the FS training group had the same amount of practice on the FS task as the FB training group had on the FB task, the effect of training on the FS task might have transferred to the FB post-test task in comparison to the Control training group. In order to investigate this possibility, the following analyses examined whether children in the FS training group benefited from training session 1, which was their first FS task, in the same way as those in the FB training group had from the pre-test. A significant increase in performance was found from training session 1 to 2 for the FS training group, $\chi^2(1) = 6.13, p < .01, \phi = .57$. Moreover, the improvement from pre-test to training session 1 for the FB training group was not significantly different from the improvement from training session 1 to 2 for the FS training group, $\chi^2(1) = .45, p = .74$, suggesting that an extra practice on the FS task might help in consolidating false-sign understanding in the same way that the pre-test has benefited the FB training group.

To further examine the transferability of the FB and FS tasks, the contingency between performance of the FB training group on training session 2 and the FS post-test was compared with the contingency between performance of the FS training group on training session 2 and the FB post-test. The number of children who passed FB training session 2 but failed the FS post-test task ($n = 4$) was not significantly different from the number of children who failed FB training session 2 but passed the

FS post-test task (n = 4), $\chi^2(1) = .13, p = 1$; and the number of children who passed FS training session 2 but failed the 'self' question of the FB post-test task (n = 7) was also not significantly different from the number of children who failed FS training session 2 but passed the 'self' question of the FB post-test task (n = 1), $\chi^2(1) = 3.13, p = .07$. Moreover, both contingencies showed the same numbers of children passing (n = 8) and failing (n = 3) both training session 2 and the post-test and the same number of children passing (n = 8) either training session 2 or the post-test. These findings suggested that training on the FS task probably had the same transfer effect to the 'self' question of the FB post-test task as training on the FB task had on the FS post-test task, confirming the other half of the second prediction.

Finally, with a wide age range of 3;3 – 5;4, it was possible that age may have played a role in how well the children responded to the training and subsequent success on the post-test tasks. To test this possibility, children who failed the first training session and passed the relevant post-test task were identified as "improvers" and then whether improvement and post-test performance correlated with age was examined. There were six improvers in each of the FS and FB training groups and seven in the Control training group. Spearman correlation tests revealed that age was not associated with improvement and post-test performance, r_ss (57) < .24, ps > .07[3], suggesting that age did not play a significant role in the success of training and post-test performance.

Discussion.

The main aim of this experiment was to critically examine the proposal that the FB and FS tasks draw on the same underlying conceptual demands in testing understanding of representations. If equivalence was found in a training study, this would provide more stringent evidence for this claim that understanding of false belief reflects a general understanding of representations rather than a domain-specific understanding of mind. Related to this the experiment aimed to find concept generalisation by using tasks that were not identical between training and post-test. Moreover, it was expected that children's age would not explain how well they responded to the training and performed in the post-test.

[3] The finding of r_s (57) = .24, p = .07 was gained between age and the 'self' question of the VPT post-test task. The other findings, showing age was not associated with improvement and post-test performance, were r_ss (57) < .21, ps > .12.

Main findings.

With respect to the proposal that understanding of false beliefs reflects a general understanding of representations, two main findings provided evidence for the equivalence of false-belief and false-sign understanding. First, it was found that children responded similarly on both FB and FS post-test tasks and the two tasks were related, suggesting that the two tasks were equivalent, supporting the results of previous studies (Leekam et al., 2008; Sabbagh, Moses, et al., 2006, Experiment 2). Second, it was found that training on the FB task enhanced performance of the FS post-test task, showing a transfer effect from false-belief training to false-sign understanding. Although between-participants analysis did not show a transfer effect from false-sign training to false-belief understanding, other within-participants analyses suggested that training on the FS task contributed to the understanding of one's own false beliefs and provided evidence for the equivalence/transferability between the two types of false representations.

Evidence for concept generalization was also found. In comparison to the Control training group, training on the location change FS task enhanced performance on the contents change FS post-test task for the FS training group and training on the location change FB task enhanced performance on the 'other' question of the contents change FB post-test task for the FB training group. This finding showed a generalisation of trained concepts to the post-test tasks that are structurally different from the training tasks. However, effects of training on non FB and FS post-test tasks were not found. Finally, age did not have a significant effect on the training and post-test performance. The implications of all these findings are outlined below.

Implications of the findings.

The main finding of transferability between the FB and FS tasks after training supports the claim for non-specificity of false-belief understanding. However, an alternative explanation might come from the overall better performance on both FB and FS post-test tasks shown by the FB training group (see Table 2.7). Instead of a transfer effect from false-belief training to false-sign understanding, one might argue that the children of the FB training group might have already possessed the concept of false-sign understanding in the first place given that they were not pre-tested on any FS task. Although this possibility cannot be ruled out, I suggest that it is unlikely based on the contingency demonstrated between the performance of the FB training group on the FS and the 'other' question of the FB post-test tasks. Only the 'other' question of the FB task was considered because it was the question on which the false-belief training was shown to be effective. There were 13 out of 19 children who showed consistent performance between the two post-test tasks. Out of these 13 children, 9 passed both tasks. In contrast, there were only 3 children in the FB training group who failed the FB task but passed the FS task. These 3 children showed no benefits from the false-belief training at all (failing the 'self' question of the FB task as well) but still performed well on the FS task, suggesting that they were the few who might have already possessed the concept of false-sign understanding in advance. Such pattern of contingency provided stronger evidence for a transference from false-belief training to false-sign understanding than a pre-existing knowledge of false-sign understanding.

The subsidiary finding of concept generalisation between different task forms supports the results of previous successful training studies (Appleton & Reddy, 1996; Slaughter, 1998). Appleton and Reddy found significantly improved performance on the contents change FB post-test tasks after training children on the location change FB tasks, whilst Slaughter demonstrated an elevated performance on the FB and VPT post-test tasks after children were trained on the FB task. However, the current experiment only showed an improvement on the 'other' question of the contents change FB post-test task whereas Appleton and Reddy showed improvements on both 'self' and 'other' questions of the contents change FB post-test tasks after training on the location change FB tasks. A possible reason for this difference between the two studies would be the difference in procedures used for the 'self' question of the contents change FB task. Appleton and Reddy's procedure for the 'self' question of the contents change FB task involved asking the question while the unexpected content of the container was still out of the container and on display in the experimenter's hand. The current experiment, instead, involved putting the unexpected content of the container back into the container before asking the 'self' question of the contents change FB task. As Appleton and Reddy have mentioned, their procedure for the 'self' question of the contents change FB task might have slightly reduced the incidence of error on the question by reducing the conflict with reality. Therefore, even though training on the location change FB tasks in both that and the present experiment only involved asking someone else's belief about the location of something (i.e., an 'other' question), only Appleton and Reddy's study found improvements on both 'self' and 'other' questions of the contents change FB post-test tasks.

With respect to the last finding of no training effects on other post-test tasks (i.e., the VPT, AR and NC tasks), there may be several explanations. First, the 'self' question of the VPT task might be too easy for the children so that a ceiling effect was demonstrated across the training groups. Second, a different training task was used for the FB training group in the current experiment in comparison to Slaughter's (1998) Study 2 and Slaughter and Gopnik (1996); this might be a possible reason that the performance on the VPT and AR post-test tasks was not improved by training. Both previous studies trained children with deceptive objects. This type of training is relevant to the VPT and AR post-test tasks. However, the current experiment trained on location change tasks that are less relevant to the cognitive demands of visual perspective-taking and appearance-reality distinction. This may explain why the two previous studies showed evidence that the FB training group outperformed the other two groups on the VPT and AR post-test task, whereas the current experiment did not. Third, another reason why there was no training effect on the VPT and NC post-test tasks might be because of the possibility that some of the children in the irrelevant training groups (i.e., the Control training group for the VPT post-test task; the FS and FB training group for the NC post-test task) already possessed the concepts of visual perspective-taking and number conservation. Previous studies have shown evidence that the ability of visual perspective-taking emerges prior to false-belief understanding (e.g., Flavell, Green, & Flavell, 1986; Gopnik & Slaughter, 1991).

Finally, the lack of pre-test on the NC task and of significant difference between the training groups' performance on the NC post-test task might lead one to argue that the training of the NC task was so off-putting that it did not have any benefits and further hampered the Control training group's performance on the other post-test tasks, resulting in the impression that the FS and FB training groups were seemingly better than the Control training group on those tasks. One might also suggest that the NC task might not be a good enough control task because it was completely different in structure to the FB and FS tasks (e.g., the NC task might not be subject to any practice effect but the FB and FS tasks were). This possible difference between tasks might create a difference between the training groups and affect the results. However, further analyses indicated that these possibilities were unlikely to be the case. First, there was a significant increase in performance from training session 1 to 2 for the

Control training group, $\chi^2(1) = 4, p < .05$. Improvement was also shown from training to post-test as mentioned in the Results section. Second, the improvement from training session 1 to 2 for the Control training group was not significantly different from that for the FS training group and the improvement from pre-test to training session 1 for the FB training group, $\chi^2(1) = .59, p = .84$, indicating that the NC task might be subjected to the same practice effect as the FS and FB tasks. Moreover, the number of "improvers" in each training group was similar (6 in each of the FS and FB training groups and 7 in the Control training group), suggesting that the training groups did not differ in terms of how well they benefited from the training tasks provided. Thus, children in the Control training group did benefit from the training and their poor performance on the other post-test tasks was due to their lack of competence.

Mechanism underlying the effects of training.

Training effects were successfully demonstrated in the current experiment and the potential mechanism underlying these effects was further investigated. Results revealed that practice with or without feedback seemed to be a key to enhance performance, whether it is on false representations or number conservation. More specifically, the first encounter with a task significantly increased later performance, provided that the delay in between is short. This practice effect shown from the first to second encounter of a task however does not imply an acquisition of conceptual understanding. Instead, a conceptual change in understanding representations was demonstrated by the concept generalisation from training on the location change tasks to post-test on the contents change tasks and the transfer effect between false belief and false sign shown in the current experiment. It is widely suggested that repeated exposure to the tasks together with feedback given according to participants' responses are crucial to effect this conceptual change. Useful kinds of feedback include counter-evidence through simple verbal feedback used in the current experiment, Slaughter (1998, Study 2), and Slaughter and Gopnik (1996), implicit feedback (Amsterlaw & Wellman, 2006), discourse (e.g., Appleton & Reddy, 1996; Lohmann & Tomasello, 2003), and explanations (e.g., Melot & Angeard, 2003).

Another possible key factor for successful training was the intensity of the training scheme. All successful false-belief training studies (e.g., Appleton & Reddy, 1996; Lohmann & Tomasello, 2003; Slaughter, 1998) employed a distributed scheme, having several sessions over a long period of time. Amsterlaw and Wellman's (2006) microgenetic study even demonstrated that a focal group which received a spread-out schedule of experiences on FB tasks showed improvement whereas a comparison group which had a more condensed schedule of experiences did not.

Other factors, such as age (e.g., see Wellman et al., 2001, for a meta-analysis) and language (e.g., Lohmann & Tomasello, 2003; see Milligan et al., 2007, for a meta-analysis), are important for children to make progress in false-belief understanding. However, the current experiment did not show that age played a significant role in the success of the training, suggesting that the 3-year-olds were no less reactive or sensitive to the training than the 5-year-olds. The training scheme used was as effective to young children as to older children if they were previously found to have not yet possessed the conceptual understanding of false belief. Unfortunately, the current experiment was not able to suggest whether language development has a contribution in the success of the training and the development of representational understanding. Further investigation is required to evaluate the influence of language development on the acquisition of representational understanding. For example, if children's level of language skills determines their understanding of representation, then reducing the language demands of the false representation tasks should enhance their performance. This issue is dealt with later in the thesis.

To conclude, a conceptual understanding of representations, rather than simply learned superficial strategies of passing the tasks, was developed gradually in the children in the current experiment through a distributed training scheme which involved repeated exposure to the tasks with appropriate feedback. Given that the current experiment is the first to include false-sign training, I have only been able to discuss the underlying mechanism in terms of previous false-belief training studies. However, due to the transferability that has been shown between false belief and false sign, I reason that the same mechanism which underlies false-belief training would also hold for false-sign training. I suggest that children's understanding of representations develops through a process of conceptual consolidation and restructuring which requires the experience of new evidence to be provided.

The interpretation for the training effect found in the current experiment was of a shared concept of representational understanding underlying both FB and FS tasks. However, an alternative explanation could also be proposed. As both FB and FS tasks confound the requirement of understanding representations with language and executive demands, the transferability found between the FB and FS tasks might be attributed to an increased ability of language and executive function rather than representational understanding after training. In the following chapter, the nature of the equivalence between the FB and FS tasks is further investigated by carefully controlling the demands of language and executive function of the tasks.

Chapter 3 Equivalence of Performance on False Mental and False Non-Mental Representation Tasks: Representational Understanding or Language and Executive Function?

As discussed earlier, both FB and FS tasks confound the requirement of understanding representations with language and executive demands. It is therefore possible that the association and transfer effect found between the two tasks in previous studies, including Experiment 1 and 2 in Chapter 2, were due to language and executive function rather than representational understanding. In order to address this possibility, Chapter 3 describes tasks that disentangle language and executive function from representational understanding, and two experiments that employed these tasks. One of these tasks is Apperly et al.'s (2004) non-verbal reality-unknown FB task which was adapted from Call and Tomasello (1999). Apperly et al.'s FB task was employed because it improved on Call and Tomasello's task by standardising the procedure with videos and carefully controlling confounding factors with intervened test, control and filler trials. I briefly outline Apperly et al.'s FB task below.

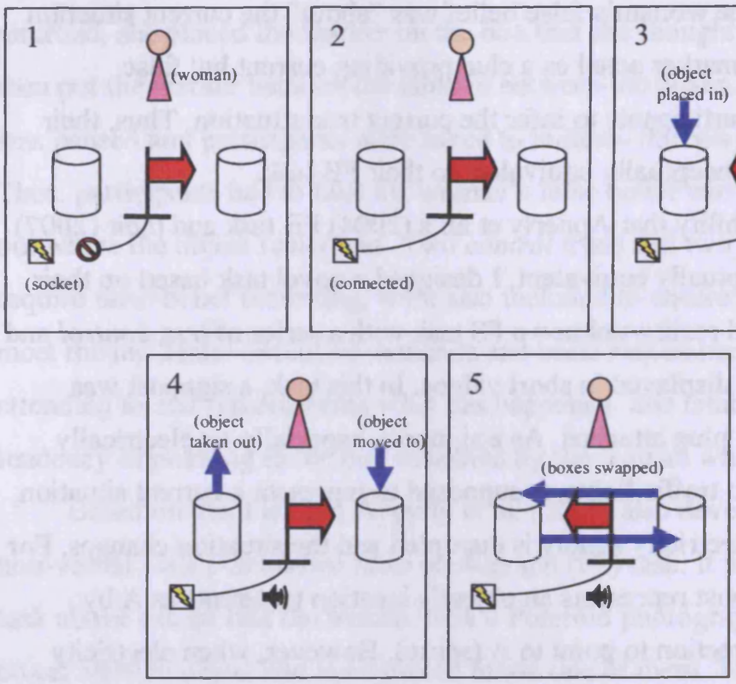
In Apperly et al.'s (2004) FB task, the confounding cognitive factors of language and executive function were minimised, eliminated or carefully controlled. First, the task was mainly non-verbal which used the minimum level of language. Second, the task was free from the requirement of cognitive inhibition of one's own knowledge about reality. Third, the task involved control trials in which participants were tested on incidental executive demands (i.e., working memory and response inhibition). Each trial of the task was displayed through a short video and participants were required to work out which of two identical boxes in the video contained a hidden object (i.e., a plastic block). Participants were instructed that a woman in the videos was there to help them find the object by placing a marker on top of a box to indicate the object's location. Participants' comprehension of the instructions given was checked with a number of warm-up trials on which corrective feedback was provided when necessary. All of the trials initially showed a man hiding an object in one of two boxes. This hiding action did not reveal to the participant which box actually contained the object. He then tipped the boxes toward the woman so she could have a look into both boxes. The participant him/herself could not see into the boxes at this stage. Therefore, both the man and woman in the videos knew which box the object was in, but the participant did not. For the *test* trial, the woman then left and the man swapped the boxes around, creating a false belief in the woman. When she returned, she placed the marker on the box that she thought contained the object and then put the marker back on the table in between the boxes. At this point, the video was paused and participants were asked to indicate the box that contained the object. Thus, participants had to take the woman's false belief into account in order to work out where the object really was. Two *control* trials and two *filler* trials, which did not require false-belief reasoning, were also included to ensure that participants could meet the incidental executive demands and basic requirements of the task such as attending to and remembering what has happened, and inhibiting the prepotent tendency of pointing to the box indicated by the woman when she was wrong.

Based on this FB task, Apperly et al. (2007) also developed a similar task: the non-verbal reality-unknown false photograph (FP) task. It followed closely to the FB task above except that the woman took a Polaroid photograph of the interior of the boxes after an object had been placed inside one of them. The participants could not see which box contained the object. The woman placed the photograph face-down in front of the boxes, and left without returning. For the *test* trial, the man swapped the boxes. He then helped participants to find the object by revealing the photograph which showed the object inside one of the boxes. He placed the photograph face-down again and the video was paused for participants to indicate the box that contained the object. Two *control* and two *filler* trials were also included. This FP task was claimed to be closely matched to the FB task in the aspects of conceptual and incidental executive demands.

According to Apperly et al. (2007), both of their FB and FP tasks required understanding of representations. They argued that both made the same conceptual demand in the sense that both tasks required participants to consider the woman's belief or the photograph in relation to the current situation in order to figure out the current location of the object. Hence, Apperly et al. suggested that their tasks answered Perner's (1995) argument against using the standard FP task (e.g., Zaitchik, 1990) as a non-mental comparison to the standard FB task (i.e., the photograph used in the standard FP task is a true representation of the situation at the time the photograph was taken; whereas a false belief is a misrepresentation of the current situation). However, it is still important to note that the content of a photograph is "about" an outdated situation although Apperly et al.'s FP task related it to the current situation. In other words, the photograph in their FP task acted as a clue providing outdated information which helped participants to infer the current situation. On the contrary, in their FB task, the woman's false belief was "about" the current situation and her indication with the marker acted as a clue providing current but false information which helped participants to infer the current true situation. Thus, their FP task might still not be conceptually equivalent to their FB task.

Considering the possibility that Apperly et al.'s (2004) FB task and their (2007) FP task might not be conceptually equivalent, I designed a novel task based on their tasks. This was a non-verbal reality-unknown FS task with a series of *test*, *control* and *filler* trials which were also displayed in short videos. 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 electricity supply is disrupted and the situation changes. For example, an electrical signpost represents an object's location (referent) as A by automatically turning its direction to point to A (sense). However, when 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. The object's location which is the referent of the signpost is now B but the sense of the signpost is still A so the signpost becomes false in that its reference is distinct from its sense. In this way, the false signpost matches with the woman's indication in Apperly et al.'s FB task, both of which act as clues providing current but false information which help participants to infer the current true situations. However, the signpost is non-mentally (electrically) operated whereas the woman's indication in Apperly et al.'s FB task is mentally generated.

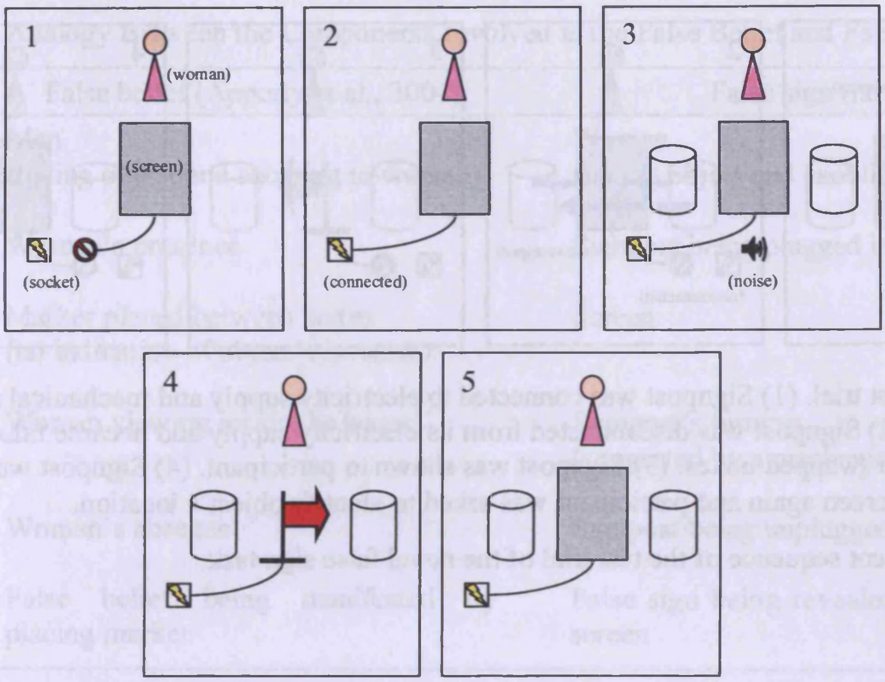
The newly devised FS task involved a short training video clip which presented the signpost as being electrically operated by means of electrical connection when plugged into a socket by a woman. It then turned left or right in between two identical boxes to show where an object was. A mechanical noise could be heard whenever the signpost was turning. Therefore, participants learned the association between the mechanical noise and the "turning" of the signpost during the training. It is important to note that the mechanical noise did not indicate direction, only the pointing of the signpost indicated direction. Moreover, the signpost turned not only when the object was moved from one box to the other box, but also when the two boxes were swapped around with the object inside one of them. Figure 3.1 presents the event sequence in the training video clip.



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 3.1. Event sequence in the training of the novel false sign task.

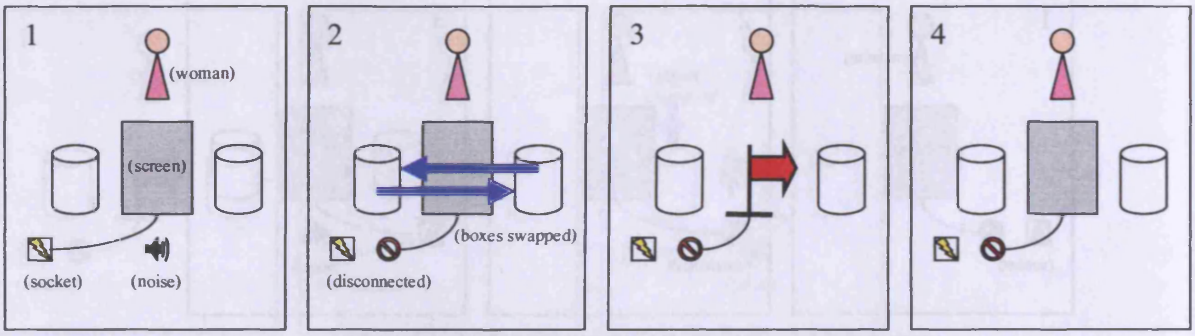
Following the training of the novel FS task, warm-up trials were given (see Figure 3.2). In the warm-up trials, the signpost was blocked from view by a screen but its electric plug was visible. The woman put the plug into a socket and placed the two boxes on the sides of the screen which covered the signpost only. It was important to note that participants knew that one of the two boxes contained the object but did not know which one. The mechanical noise was heard (suggesting the signpost was turning behind the screen) as soon as the two boxes were placed. When the noise stopped, the woman removed the screen to show the signpost as a true sign indicating the location of the object. Participants were then asked to identify the object's location. Following Apperly et al.'s (2004, 2007) FB and FP tasks, corrective feedback was provided in these warm-up trials when necessary.



Warm-up trial. (1) Signpost was blocked from 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 identify object's location.

Figure 3.2. Event sequence of the warm-up trial of the novel false sign task.

The *test* trial of the FS task was as follows (see Figure 3.3). (1) The signpost was blocked from view by a screen in the same way as the warm-up trials. The woman put the plug of the signpost into a socket. She then placed the two boxes (one contained the object but participants did not know which one) on the sides of the screen. The mechanical noise was heard (meaning the signpost was turning). (2) 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. (3) She then helped participants to find the object by taking away the screen to show the signpost. (4) Eventually, she put the screen back in front of the signpost. The video was paused and participants were asked to identify the object's location. Participants thus had to take the false sign into account in order to work out where the object really was. Table 3.1 shows the key stages of the test trials of these non-verbal reality-unknown FB, FP and FS tasks. Table 3.2 presents the analogy between the components involved in the FB and FS tasks.



False sign test trial. (1) Signpost was connected to electricity supply and mechanical noise was heard. (2) Signpost was disconnected from its electricity supply and became false when woman swapped boxes. (3) Signpost was shown to participant. (4) Signpost was blocked by screen again and participant was asked to identify object's location.

Figure 3.3. Event sequence of the test trial of the novel false sign task.

Table 3.1

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.

Table 3.2

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

With this newly devised FS task, the following two experiments investigated whether the equivalence between the standard FB and FS tasks are based on the same underlying representational understanding or other cognitive processes of language and an aspect of executive function, namely cognitive inhibition. The standard FB and FS tasks used in previous studies and Experiment 1 and 2 presented in Chapter 2 were verbal tasks in which the real location or real identity of an object was known by the participant. The novel FS task and Apperly et al.'s (2004, 2007) FB and FP tasks used in the following two experiments were non-verbal reality-unknown tasks in which the real location of an object was not known by the participant. Hence, language was not involved and cognitive inhibition was not required in these non-verbal reality-unknown tasks. If the association between the standard FB and FS 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 between the FB and FS tasks.

Experiment 3 provided the very first test of the novel FS task, investigating how it worked relative to Apperly et al.'s FB and FP tasks in typically developing children. As Apperly et al.'s tasks have not been used with typically developing children in any published study, Experiment 3 also extended the applicability of their tasks. Children with an age range of 3 to 7 instead of 5 years were recruited because the youngest mean age that Apperly et al.'s FB task has been established with 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 et al., 2007). If children's performance on these non-verbal reality-unknown tasks was similar to that of the standard tasks, the strong association previously found between the standard FB and FS tasks but not between the standard FB and FP tasks would be replicated in this non-verbal reality-unknown context. This replication would provide further support for the claim that the FB and FS tasks share a developmental factor which is not shared by the FP task. Moreover, this developmental factor would not be cognitive skills of language and cognitive inhibition but representational understanding, which allows a differentiation between reference and sense, as the non-verbal reality-unknown context did not involve the demands of language and cognitive inhibition.

Experiment 4 added a verbal reality-known version (corresponding to the standard version) of the FB and FS tasks to compare with the non-verbal reality-unknown version of the two tasks in a different group of typically developing children with the same age range of 3 to 7 years. The aim of this experiment was twofold. First, this experiment intended to further replicate the associations found between the FB and FS tasks in both standard and non-verbal reality-unknown versions within a single experiment. If a replication was illustrated again, it would evidently suggest that language and cognitive inhibition do not play a role in the association between the FB and FS tasks. Second, the experiment aimed to validate the novel FS task with the standard version of the task. If a validation was shown, it would suggest that the new FS task was virtually the standard FS task but could further be used to test populations who have language and cognitive inhibition problems such as atypically developing children.

Both Experiment 3 and 4 also measured children's language ability using the long form of the British Picture Vocabulary Scale - Second Edition (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997). This scale is a measure of receptive vocabulary which does not require any reading, speaking or writing. From this scale, children's verbal mental age was calculated. As both Experiment 3 and 4 aimed at finding associations between the FB and FS tasks and attempting to exclude language as a possible explanation for the associations found, controlling children's verbal mental age when examining the associations could further test the claim that language does not play a significant role in the equivalence between the two tasks. Ethics approval which applied to both Experiment 3 and 4 was granted by the Ethics Advisory Sub-Committee in the Department of Psychology in Durham University and the School of Psychology Ethics Committee in Cardiff University.

Experiment 3

The novel non-verbal reality-unknown FS task was compared to Apperly et al.'s (2004, 2007) non-verbal reality-unknown FB and FP task in 3- to 7-year-old children. Although associations have been examined between the standard FB, FP, and FS tasks in previous studies to date, comparisons between non-verbal reality-unknown versions of these tasks have not been made. If the strong association between the standard FB and FS tasks but not between the standard FB and FP tasks was replicated in the non-verbal reality-unknown context of the current experiment, this would suggest that language and cognitive inhibition cannot explain the association between the tasks.

Following the pattern of performance on the standard FB, FP and FS tasks, it was expected that performance on the non-verbal FS and FB tasks should correlate strongly whereas performance on the non-verbal FP and FB tasks should be less correlated. This strong correlation between the non-verbal FS and FB tasks should also hold, whether or not performance on the non-verbal FP task or children's verbal mental age was controlled. Moreover, older children were expected to perform better than younger children on the non-verbal FB, FP and FS tasks, especially with this wider range of age. These predictions were made regarding to children's performance on the test trials of these non-verbal tasks.

These non-verbal tasks also involve control trials, which did not require false-belief, pictorial and false-sign reasoning, to ensure that children could meet the incidental executive demands of the tasks. Instead of screening children with these control trials as done in Call and Tomasello (1999), the assurance of meeting incidental executive demands of the tasks was achieved by showing that children's performance on these control trials was significantly better than that on the test trials. Performance on these control trials across tasks also provided an additional test of whether these tasks were equivalent in terms of their incidental executive demands. Finally, performance on the filler trials was also analysed.

Method.

Participants.

Twenty 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. 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. None of the participants from North East England participated in Experiment 1 and 2 reported in Chapter 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 BPVS-II. 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. Parental informed consent was obtained before testing.

Design.

This was a mixed design, testing younger versus older participants with the non-verbal reality-unknown FB, FP and FS tasks. Participants were tested with three sessions at one- or two-week intervals. In each session, they were tested with three non-verbal reality-unknown tasks: false belief, false photograph and false sign. Each task in the test phase consisted of five trials: false representation, working memory control, response inhibition control, true representation filler and clue confirmation filler. The presentation order of the three tasks was counterbalanced across participants and sessions whilst the order of the five trials within each task was randomised in each of the three sessions. The purpose of giving three sessions was to gain more reliable data for each participant as each trial involves a single pass/fail response.

Materials and procedure.

Participants were tested individually in a room of their own school by an experimenter. The tasks were video based and were presented on a laptop computer using PowerPoint software. As the FB and FP tasks were originally designed for adults being administered in two different studies (Apperly et al., 2004, 2007), 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. This training procedure for the FP task was adapted from that in the standard FP task devised by Zaitchik (1990). As children by the age of 3 years already understand the relationship between seeing and knowing, no training would be required for the FB task. Second, calls for attention (e.g., calling participants by their names) and encouraging words (e.g., "Well done!") to boost motivation were required for children in each of the FB, FP and FS tasks. Third, a break was given between each of the three tasks.

The training video of the FP task showed a man who pressed the button of a Polaroid camera in front of some flowers and a picture came out of the camera; the picture then developed to show an accurate image of the flowers. The training video of the FS task showed a woman who manipulated an electrical signpost (see Figure 3.1). She put the plug that was attached to the signpost in a socket and then hid an object (i.e., a plastic block) in one of two identical boxes which were placed on the two sides of the signpost. The signpost automatically turned to show the location of the object. Each time the signpost turned, it made a mechanical noise. When the object was displaced visibly (i.e., the object was taken out from one box and put into the other box in full view of participants), the signpost turned to show the object's new location and the mechanical noise was heard while it was turning. Finally, the object was displaced invisibly (i.e., the object was in one of the boxes and the boxes were swapped). The signpost turned again to accurately show the object's current location and the mechanical noise was heard when the signpost turned.

Before each task started, the principles of each task were explained verbally with stills, e.g., "We will play a hiding-and-finding game and your job is to find a block. Here, the woman connects the sign by putting its electric plug into the socket. Then she hides the block here and the sign turns to help you to find the block." Warm-up trials were then given (see Figure 3.2). Participants were asked to identify the location of the object which was hidden in one of the two boxes. In each task, a woman provided a clue to help the participants to locate the object. For the FB task, a woman placed a marker on top of the box she thought contained the object. 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 object. 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 participants before they entered the test phase of the relevant task. None of the participants required more than four warm-up trials to achieve two consecutive correct responses.

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 FB, FP and FS tasks. The five trials of the FB and FP tasks were exactly the same as those in Apperly et al. (2004, 2007; see Table 3.3 and 3.4 respectively). For the FS task, its test, control and filler trials are listed in Table 3.5.

Table 3.3

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>

Table 3.4

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

<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 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.</td></tr>
<tr><td>The man swapped the boxes.</td><td>The man revealed the photograph and then covered it.</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 man revealed the photograph and then covered it.</td></tr>
<tr><td>He revealed the photograph and then covered it.</td><td>He swapped the boxes.</td><td colspan="2">He revealed the photograph and then covered it.</td><td>He took the object out of the box indicated by the photograph and put it into another box.</td></tr>
</table>

Table 3.5

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>

Only the *false representation test* trials required false-belief, pictorial or false-sign reasoning. The *control* trials involved the same demands of (1) holding the events that happened in mind while working out the object's location, and (2) inhibiting the tendency of pointing to the box indicated, as the test trials. In the *true representation filler* trials, the object's location was not changed so the woman's belief, photograph or signpost remained accurate in representing the object's location. This filler trial type was to check whether participants had adopted an incorrect strategy of always pointing to the opposite box from the one indicated by the woman, photograph or signpost. The *clue confirmation filler* trials revealed the woman's belief, photograph or signpost and then showed the object being taken out of the indicated box, demonstrating that the woman, photograph or signpost was not meant to be deceptive. Both types of these filler trials did not require false-belief, pictorial and false-sign reasoning, and demanded less executive skills than the test and control trials. Thus, a relatively good performance on these filler trials would reflect that participants had paid attention and not adopted incorrect strategies of guessing or always pointing to the opposite box from the one indicated. Feedback which showed the interior of the two boxes was always presented at the end of each trial after participants had responded.

Two more testing sessions followed, at one- or two-week intervals with different presentation orders of tasks and trials. In the second and third sessions, no training and warm-up trials was entailed. Each of the FB, FP, and FS tasks started with an explanation of its principle, followed by the test phase of the relevant task.

Results.

Children received a score of 1 for each of the trials if they correctly identified the object's location. Scores across the three sessions were summed and the mean scores for each trial type of the three tasks are shown in Table 3.6. The distribution of the data was negatively skewed for most of the dependent variables and data transformations, including square root, log and inverse transformations, did not improve normality. However, the outcomes of the following parametric analyses were corroborated by running non-parametric equivalent tests. Results from these non-parametric analyses were also reported for each set of the parametric analyses.

Performance on the false representation test trials of the three tasks was first analysed to investigate the equivalence of the tasks. These analyses included two parts. Part one took age difference into account by (1) comparing older and younger participants' performance against chance; (2) performing a mixed between- and within-participants ANOVA with age as between-participants factor and task as within-participants factor; and (3) determining whether older versus younger participants improved across the three sessions given that feedback was provided after each trial. This improvement across sessions was examined using non-parametric Friedman tests as the range of scores participants obtained in each session was 0 – 1. Part two involved the whole sample to (1) examine the degree of associations between the FB, FP and FS tasks with correlation tests; and (2) investigate the contingencies between participants' performance on any two of the three tasks in the first session with non-parametric McNemar tests.

To explore whether the three tasks were also equivalent in their incidental executive demands, performance on the control trials of the tasks was analysed by comparing against chance and carrying out a mixed ANOVA with age as between-participants factor and task as within-participants factor for each type of the control trials. Then, performance on the false representation test trials versus the control trials was investigated for each of the FB, FP and FS tasks, using a one-way repeated-measures ANOVA. The purpose of this analysis was to ensure that children's performance on the false representation test trials could not be accounted for by their ability in meeting the incidental executive demands of the tasks. Finally, performance on the filler trials of the tasks was examined. As the filler trials were relatively easy, a ceiling effect on these trials was expected if children have paid attention and have not adopted incorrect strategies.

Table 3.6

Mean Scores on Each Trial Type of the False Belief, False Photograph and False Sign Tasks in Experiment 3 (Standard Deviations are Shown in Brackets)

			Task (range of scores = 0 – 3)		
Trial	Age	N	False belief	False photo	False sign
False representation	All	40	1.45 (1.06)	1.5 (1.2)	1.5 (1.18)
	Younger	20	0.85 (0.81)	0.7 (0.8)	0.9 (1.12)
	Older	20	2.05 (0.94)	2.3 (0.98)	2.1 (0.91)
Response inhibition	All	40	2.7 (0.65)	2.6 (0.81)	2.73 (0.60)
	Younger	20	2.45 (0.83)	2.4 (0.99)	2.5 (0.76)
	Older	20	2.95 (0.22)	2.8 (0.52)	2.95 (0.22)
Working memory	All	40	2 (1.04)	2.45 (.78)	2.1 (1.03)
	Younger	20	1.6 (0.94)	2.25 (0.85)	1.7 (0.98)
	Older	20	2.4 (0.99)	2.65 (0.67)	2.5 (0.95)
True representation	All	40	2.73 (0.51)	2.73 (0.56)	2.8 (0.46)
	Younger	20	2.65 (0.59)	2.8 (0.41)	2.75 (0.55)
	Older	20	2.8 (0.41)	2.65 (0.67)	2.85 (0.37)
Clue confirmation	All	40	3 (0)	3 (0)	3 (0)
	Younger	20	3 (0)	3 (0)	3 (0)
	Older	20	3 (0)	3 (0)	3 (0)

False representation test trials.

The first part of the analyses looked at older versus younger participants' performance across the FB, FP and FS tasks. Older participants performed above chance on the test trials of the three tasks, $ts(19) > 2.60$, $ps < .05$ with effect sizes of $rs > .51$. However, younger participants performed below chance on the three test trials, $ts(19) < -2.40$, $ps < .05$, $rs > .48$. A 2 (age) X 3 (task) ANOVA revealed a main effect for age, $F(1, 38) = 28.69$, $p < .001$ with an effect size of $r = .66$, but no effect for task and no interaction, $Fs(2, 76) < 1.40$, $ps > .25$. Equivalent non-parametric analyses demonstrated identical results. Older participants performed consistently above chance, Wilcoxon $Zs < -2.27$, $ps < .05$ with effect sizes of $rs < -.36$. However, younger participants performed below chance level, Wilcoxon $Zs < -2.13$, $ps < .05$, $rs < -.34$. Mann-Whitney U tests showed that older participants performed better than younger participants on all three tasks ($Us < 85$, $ps < .01$ with effect sizes of $rs < -.51$). There was no significant difference between tasks, Friedman test: $\chi^2(2, N = 40) = .28$, $p = .89$.

To examine whether participants improved their performance across the three sessions, non-parametric Friedman tests were used. Older participants did not improve their performance on the FB and FS tasks across the three sessions, $\chi^2s(2, N = 20) < 3.82$, $ps > .24$, 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 was improved from Session 1 to 2 and from Session 1 to 3, Wilcoxon $Zs < -2.45$, $ps < .05$, $rs < -.39$, but no improvement was shown from Session 2 to 3, Wilcoxon $Z = -1$, $p = 1$. Younger participants did not improve their performance on all three tasks, $\chi^2s(2, N = 20) < 3.43$, $ps > .23$.

The second part of the analyses involved the whole sample to examine the correlations and contingencies between any two of the FB, FP and FS tasks. Bivariate correlations were calculated with total scores across the three sessions (range: 0 – 3), whereas contingencies were computed with scores from the first session only (range: 0 – 1). Both Pearson's and non-parametric Spearman's correlation tests showed that performance on the three tasks was significantly correlated. Pearson's correlations indicated that the FB task was significantly related to both FP and FS tasks, $r(40) = .67$, $p < .001$, and $r(40) = .76$, $p < .001$ respectively, and the FP task was correlated with the FS task, $r(40) = .71$, $p < .001$. Results from Spearman's correlation tests can be found in Table 3.7. The magnitude of the bivariate correlations between the three tasks remained strong when children's VMA was controlled: $pr(37) = .60, p < .001$ between the FB and FP tasks, $pr(37) = .71, p < .001$ between the FB and FS tasks, and $pr(37) = .66, p < .001$ between the FP and FS tasks. When performance on the third task was controlled, only the correlation between the FB and FS tasks remained highly significant; the FP and FS tasks became less correlated; and the correlation between the FB and FP tasks fell to marginal significance level ($p = .09$; see Table 3.7).

Table 3.7

Bivariate Raw Non-parametric Spearman's Correlations 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	-	.68** (40)	.77** (40)
False Photograph	[.28^{m}] (37)	-	.71** (40)
False Sign	[.55**] (37)	[.42*] (37)	-

Note. $^{**}p < .001$, $^{*}p < .01$, $^{m}p = .09$

The contingencies between performance on any two of the FB, FP and FS tasks in the first session are shown in Table 3.8. There was a majority of participants who performed consistently between any two of the tasks (i.e., they either answered correctly or incorrectly to both tasks: 27 in the FB-FP pair, 27 in the FB-FS pair, and 28 in the FP-FS pair). The other participants who passed only one task in any two of the tasks were subjected to McNemer tests. These tests indicated that the number of participants who passed one task but failed the other task (e.g., 7 passed the FB task but failed the FP task) was not significantly different from the number of participants who failed the former but passed the latter (e.g., 6 failed the FB task but passed the FP task), χ^2s$(1, N = 40) < .08$, $ps > .77$. These findings suggest that the three tasks were of similar difficulty for participants who first encountered them.

Table 3.8

Number of Participants With Correct Scores on any Pairs of the False Belief (FB), False Photograph (FP) and False Sign (FS) Tasks in the First Session

	Total in each pair of tasks		
Correct score	FB and FP	FB and FS	FP and FS
Contingency between			
Both tasks (+ +)	20	19	20
1st task only (+ -)	7	6	5
2nd task only (- +)	6	7	7
Neither task (- -)	7	8	8

Control trials.

Older children performed consistently above chance on the control trials of all three tasks, $ts(19) > 4.05$, $ps < .01$, $rs > .68$ (non-parametric tests showed the same results: Wilcoxon $Zs < -3.15$, $ps < .001$, $rs < -.50$). Younger children also performed above chance on the control trials of all three tasks, $ts(19) > 3.94$, $ps < .01$, $rs > .67$, except the working memory control trials of the FB and FS tasks on which they performed at chance level, $ts(19) < .91$, $ps > .37$. The same was found with non-parametric tests: younger children performed above chance on the control trials of the three tasks (Wilcoxon $Zs < -2.98$, $ps < .01$, $rs < -.47$), except the working memory control trials of the FB and FS tasks (Wilcoxon $Zs > -1.01$, $ps > .35$).

For the response inhibition control trials of the three tasks, Mauchly's (1940) test indicated a departure from sphericity and it was corrected using the Greenhouse-Geisser *epsilon* (Greenhouse & Geisser, 1959; corrected degrees of freedom are reported to one decimal place). There was a main effect of age, $F(1, 38) = 6.37$, $p < .05$, $r = .38$, but no effect for task and no interaction, $Fs(1.4, 76) < .98$, $ps > .36$. Non-parametric Mann-Whitney U tests showed that older children performed better than younger children on the FB and FS tasks ($Us < 138.50$, $ps < .05$, $rs < -.38$) but not on the FP task ($U = 158$, $p = .15$). No significant difference between tasks was showed, using non-parametric Friedman test: $\chi^2(2, N = 40) = 2.17$, $p = .44$.

For the working memory control trials of the three tasks, there were significant main effects of age and task, $F(1, 38) = 8.36$, $p < .01$, $r = .42$ and $F(2, 76) = 5.23$, $p < .01$ with an effect size of $\eta_p^2 = .12$ respectively, but no interaction, $F(2, 76) = 1.25$, $p = .29$. To spell out the difference between the three tasks, pairwise comparisons using Bonferroni adjustment were performed. Performance on the working memory control trials of the FB and FS tasks was showed to be worse than that of the FP task, $p < .01$ and $p = .07$ respectively, but no significant difference was found between that of the FB and FS tasks, $p = 1$. Non-parametric Mann-Whitney U tests revealed that older children performed better than younger children on the FB and FS tasks (Us < 110, ps $< .01$, rs $< -.41$) but not on the FP task ($U = 147.50$, $p = .13$). Friedman test also indicated a significant difference between tasks, $\chi^2(2, N = 40) = 10.33$, $p < .01$. Performance on the working memory control trials of the FB and FS tasks was worse than that of the FP task, Wilcoxon $Z = -2.81$, $p < .01$, $r = -.31$ and Wilcoxon $Z = -2.18$, $p < .05$, $r = -.24$ respectively, whereas performance of the FB and FS tasks did not differ, Wilcoxon $Z = -0.66$, $p = .60$.

False representation test trials versus control trials.

There were significant differences between performance on the test trials, response inhibition control trials, and working memory control trials, $F(2, 78) = 27.87$, $p < .001$, $\eta_p^2 = .42$ for the FB task, $F(2, 78) = 19.08$, $p < .001$, $\eta_p^2 = .33$ for the FP task, and $F(2, 78) = 23.74$, $p < .001$, $\eta_p^2 = .38$ for the FS task. For the FB task, performance on the test trials was significantly worse than that on the response inhibition control trials ($p < .001$) and working memory control trials ($p < .05$); whereas performance on the response inhibition control trials was significantly better than that on the working memory control trials ($p < .001$). The same pattern of performance was shown for the FS task: performance on the test trials was worse than that on the response inhibition control trials ($p < .001$) and working memory control trials ($p < .05$); whereas performance on the response inhibition control trials was better than that on the working memory control trials ($p < .01$). For the FP task, performance on the test trials was also worse than that on the response inhibition control trials and working memory control trials (ps $< .001$); whereas performance on the two types of control trials did not differ ($p = 1$). Similar results were found with equivalent non-parametric analyses: performance on both of the control trials was significantly better than that on the test trials for all three tasks (Wilcoxon Zs < -2.82, ps $< .01$, rs $< -.32$).

Filler trials.

Ceiling effects were shown for both clue confirmation and true representation filler trials of all three tasks (see Table 3.6). Moreover, there were only 11/120, 11/120, and 8/120 errors on the true representation filler trials of the FB, FP, and FS tasks respectively. 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 patterns of results suggested that participants were not likely using a strategy of always pointing to the opposite box from the one indicated by the clues.

Discussion.

The main goal of Experiment 3 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 in a group of 3- to 7-year-old children, providing the very first test of the novel FS task. Although both of Apperly et al.'s tasks have not been used in such a young age range in any published study, their tasks and the novel FS task were shown to be able to demonstrate a developmental change of false-belief, false-sign and pictorial understandings in consistent with previous studies (e.g., Leekam et al., 2008; Sabbagh, Moses, et al., 2006). Moreover, performance on the false representation test trials of the FB, FP and FS tasks was not significantly different from but was correlated with each other, even when VMA was controlled. However, only the test trials of the FS and FB tasks remained highly correlated when the third task was controlled. These findings were supported by Leekam et al. (2008) who found similar results using standard verbal FB, FP and FS tasks.

Neither younger nor older participants significantly improved their performance on the false representation test trials of the FB and FS tasks across the three sessions. However, older but not younger participants improved their performance on that of the FP task. Although the three tasks were shown to be of similar difficulty for all participants who first encountered them (in particular, older participants started off with the same level of performance: 11 passed each of the FB and FP tasks, and 10 passed the FS task), older participants improved on the FP task but not the FB and FS tasks. This finding suggests that the FP task was somehow easier and more subject to a practice effect than the FB and FS tasks for older participants, providing further support for the FP task's lack of strong associations with the FB and FS tasks.

Regarding the performance on the control trials, the FS task rather than the FP task was more comparable to the FB task in the aspect of incidental working memory demand. Comparisons between the two types of the control trials of the three tasks 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. This finding provides a possible explanation for why the FP task was easier. However, it does not sufficiently explain why only older participants improved on the FP task as younger participants were also performing above chance on the working memory control trials of the FP task. This inquiry might relate back to the conceptual problem that exists in the standard FP task. As the photographs used in the standard FP task and Apperly et al.'s (2007) FP task were true representations which represented the scenes at the time the photographs were taken (referent) as the way they were (sense), the photographs corresponded with their own referents and senses. Thus, participants might pass both of the tasks using their understanding of *correspondence* rather than representation. Although there is evidence showing that children can understand correspondence as early as 3 years of age, as mentioned in Chapter 1, younger participants probably failed Apperly et al.'s FP task because they tended to regard representations as corresponding to reality, regardless of whether it was known to them. However, older participants might realise that the task could be passed by using their understanding of correspondence from the first trial. With their sophisticated understanding of correspondence, older participants improved their performance significantly on later trials of the FP task but not the FB and FS tasks. Nevertheless, further investigations were required to evaluate this speculation. Taken together, all these findings suggested that the FB task was more equivalent to the new FS task rather than the FP task in both aspects of conceptual and incidental executive demands when applied in a population of typically developing children who are more sensitive to subtle differences in the demands of cognitive tasks than adults.

Finally, performance on both control and filler trials of the three tasks ensured that children's performance on the test trials could not be explained by an inability of meeting incidental executive demands and an adoption of incorrect strategies. Nevertheless, the non-verbal reality-unknown version of the FS task used in this experiment was new, therefore it was essential to replicate its findings and validate it with a task that corresponds to the standard FS task. Hence, Experiment 4 tested children with the non-verbal reality-unknown version and the verbal reality-known (standard) version of the FB and FS tasks.

Experiment 4

The first aim of Experiment 4 was to replicate the previous finding that the equivalence between the FB and FS tasks holds in both verbal reality-known (standard) and non-verbal reality-unknown versions within a single experiment with a different sample of typically developing children. If this was the case, it would provide further evidence that children's performance on the FB and FS tasks was supported by representational understanding and the association between the two tasks could not be explained by language and cognitive inhibition. Second, the experiment aimed to validate the new non-verbal reality-unknown FS task with the verbal reality-known FS task which was essentially a standard FS task.

In order to accomplish these two aims, Experiment 4 followed Call and Tomasello's (1999) design of administering a verbal reality-known (standard) version of the FB task in the context of the non-verbal reality-unknown version of the task series. The same procedure was also employed in the non-verbal reality-unknown FS task series. Hence, the FB and FS tasks could be compared in both versions within a single experiment. Based on the findings of previous studies and Experiment 3, no significant difference but instead significant correlations were expected between the FB and FS tasks in both verbal and non-verbal versions, regardless of whether VMA was controlled.

Given that an association has been found in Call and Tomasello (1999) between the non-verbal and verbal versions of the FB task, a similar finding was expected in the current experiment. However, whether an association would be shown between the two versions of the FS task was a focus of the current experiment. No significant difference but significant correlations between the non-verbal and verbal versions of the FB and FS tasks would validate the new non-verbal FS task.

Method.

Participants.

Thirteen children below age 5 years 2 months (aged 39 – 62 months; $M = 52.69$ months, $SD = 7.54$) and thirteen older children above age 5 years 2 months (aged 63 – 88 months; $M = 78.38$ months, $SD = 8.93$) were recruited from two schools in South East Wales. All were different children from those in Experiment 3. The sample had a mean VMA of 72.46 months (range = 37 – 100 months; $SD = 17.62$) which was calculated using the BPVS-II. 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. Parental informed consent was obtained before testing.

Design.

This was a mixed design, testing younger versus older participants with both non-verbal reality-unknown and verbal reality-known versions of the FB and FS tasks. Participants were tested with three sessions at one-week intervals. In each session, they were tested with two versions of the FB and FS tasks: one verbal and one non-verbal. Following the exact procedure of previous studies, the verbal version of the tasks consisted of two questions (a test question and a memory question); whereas the non-verbal version of the tasks consisted of five trials (false representation, working memory control, response inhibition control, true representation filler and clue confirmation filler). The presentation order of the FB and FS tasks was counterbalanced across participants. Moreover, half of the participants were tested with the verbal version of the tasks first whereas the other half of the participants had the non-verbal version first. For the verbal version of the tasks, the test questions were always asked before the memory questions. However, for the non-verbal version of the tasks, the order of the five trials within each task was randomised in each of the three sessions.

Materials and procedure.

Participants were tested individually in a room in their own school by an experimenter. The same non-verbal FB and FS task series of Experiment 3 were used. A break was given between the two task series. The only difference between Experiment 3 and 4 was a verbal version of the tasks was added in the context of the non-verbal task series. The verbal tasks were either presented consecutively after the warm-up trials of the non-verbal tasks or after the series of the test, control and filler trials of the non-verbal tasks. In this way, the presentation order of the verbal tasks and the non-verbal series of trials was counterbalanced across participants.

Similar to Call and Tomasello (1999), the verbal FB task was identical to the response inhibition control trial of the non-verbal FB task except that participants were asked about the woman's false belief regarding the location of the object when she returned to the room. That is, the test question, "Where does the woman think the block is?" After participants responded, they were asked the memory question, "Where is the block really?"

For the verbal FS task, participants saw the woman presented with the signpost covered by a screen but its plug was visible. The woman put the plug in the socket and placed the two boxes (one contained the object) on the table. The mechanical noise produced by the turning signpost was heard. The woman then revealed the signpost, took the plug out, and displaced the object visibly. Participants were then asked the test question, "Where does the arrow show the block is?", and the memory question, "Where is the block really?"

Following the non-verbal version of the tasks, feedback was provided in the verbal version of the tasks after participants responded but in a verbal way, e.g., "No, the arrow shows the block is in this box." Some participants looked confused if the feedback was unexpected but none of them disagreed with it. Calls for attention and encouraging words to boost motivation were used as in Experiment 3.

Results.

For both non-verbal and verbal versions of the FB and FS tasks, children received a score of 1 for each of the trials if they responded correctly. 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 to calculate a final score for each of the verbal FB and FS tasks. Scores across the three sessions were summed and the mean scores for each version of the tasks were shown in Table 3.9. The distribution of the data was negatively skewed and data transformations, including square root, log and inverse transformations, did not improve normality. However, as before, the outcomes of the following parametric analyses were corroborated by running non-parametric analyses. Results from these non-parametric analyses are also reported for each set of the parametric analyses.

Performance on the verbal FB and FS tasks and the false representation test trials of the non-verbal FB and FS tasks was analysed first. Performance of older and younger children was compared against chance and a three-way mixed design ANOVA with age as between-participants factor, and version and task as within-participants factors was administered. Correlation tests were then carried out to examine the degree of associations (1) between the FB and FS tasks in both verbal and non-verbal versions, and (2) between the two versions of the tasks.

Following the analyses of Experiment 3, performance on the false representation test trials of the non-verbal tasks was also subjected to non-parametric within-participants analyses to determine whether participants improved across the three sessions. Performance on the control trials of the non-verbal tasks was compared against chance and mixed ANOVAs were also carried out. Moreover, performance on the false representation test trials versus the control trials of the non-verbal tasks was compared, using one-way repeated-measures ANOVAs. Finally, performance on the filler trials was examined.

Table 3.9

Mean Scores on Each Version of the False Belief and False Sign Tasks in Experiment 4 (Standard Deviations are Shown in Brackets)

			Task (range of scores = 0 – 3)	
Version	Age	N	False belief	False sign
Verbal				
Test question	All	26	1.85 (1.19)	2.08 (1.06)
	Younger	13	1.31 (1.25)	1.38 (0.96)
	Older	13	2.38 (0.87)	2.77 (0.60)
Final score	All	26	1.73 (1.28)	1.89 (1.24)
	Younger	13	1.15 (1.28)	1.08 (1.12)
	Older	13	2.31 (1.03)	2.69 (0.75)
Non-verbal				
False representation	All	26	1.58 (1.33)	1.69 (1.23)
	Younger	13	0.77 (1.17)	0.85 (1.07)
	Older	13	2.38 (0.96)	2.54 (0.66)
Response inhibition	All	26	2.77 (0.51)	2.81 (0.40)
	Younger	13	2.54 (0.66)	2.62 (0.51)
	Older	13	3 (0)	3 (0)
Working memory	All	26	2.19 (1.02)	2.35 (0.94)
	Younger	13	1.69 (1.18)	1.92 (1.04)
	Older	13	2.69 (0.48)	2.77 (0.60)
True representation	All	26	2.42 (0.64)	2.69 (0.55)
	Younger	13	2.46 (0.78)	2.62 (0.51)
	Older	13	2.38 (0.51)	2.77 (0.60)
Clue confirmation	All	26	2.96 (0.20)	2.96 (0.20)
	Younger	13	2.92 (0.28)	2.92 (0.28)
	Older	13	3 (0)	3 (0)

Verbal tasks and false representation test trials of non-verbal tasks.

Final scores of the verbal FB and FS tasks were used in the following analyses. Older children performed consistently above chance on both verbal and non-verbal FB and FS tasks, $ts(12) > 2.82$, $ps < .05$, $rs > .63$. However, younger children performed below chance on the non-verbal FB and FS tasks, $ts(12) < -2.21$, $ps < .05$, $rs > .54$, but at chance on the verbal FB and FS tasks, $ts(12) > -1.37$, $ps > .20$. A 2 (age) X 2 (version) X 2 (task) ANOVA revealed a main effect for age, $F(1, 24) = 22.81$, $p < .001$, $r = .70$, but no effects for version and task, and no interactions, $Fs(1, 24) < 2.13$, $ps > .16$. This pattern of findings remained the same when the scores of the test questions of the verbal tasks were used.

Equivalent non-parametric analyses illustrated similar pattern of results. Older children performed consistently above chance on both verbal and non-verbal FB and FS tasks, Wilcoxon $Zs < -2.29$, $ps < .05$, $rs < -.45$. Younger children, on the other hand, performed marginally below chance on the non-verbal FB and FS tasks, Wilcoxon $Zs < -1.91$, $ps = .07$, $rs < -.37$, but at chance on the verbal FB and FS tasks, Wilcoxon $Zs > -1.33$, $ps > .23$. Mann-Whitney U tests showed that older children performed better than younger children on both verbal and non-verbal FB and FS tasks ($Us < 42$, $ps < .05$, $rs < -.45$). Moreover, no significant difference was found between the verbal FB task, the verbal FS task, the non-verbal FB task and the non-verbal FS task, Friedman test: $\chi^2(3, N = 26) < 3.98$, $p = .27$.

Pearson's and Spearman's correlations for the whole sample of children indicated that performance on the test trials of the non-verbal FB and FS tasks was significantly correlated, $r(26) = .87$, $p < .001$ and $r_s\ (26) = .86$, $p < .001$, even after controlling for VMA, $pr(23) = .79$, $p < .001$. Performance on the verbal FB and FS tasks was also correlated, $r(26) = .86$, $p < .001$ and $r_s\ (26) = .84$, $p < .001$, even after controlling for VMA, $pr(23) = .74$, $p < .001$. Significant correlations were also found between the verbal and non-verbal FB tasks, $r(26) = .61$, $p < .001$ and $r_s\ (26) = .61$, $p < .01$, and between the verbal and non-verbal FS tasks, $r(26) = .58$, $p < .01$ and $r_s\ (26) = .57$, $p < .01$. However, the two versions of the same task did not remain as significantly correlated after controlling for VMA, $pr(23) = .34$, $p = .10$ for the two versions of the FB task, and $pr(23) = .27$, $p = .20$ for the two versions of the FS task.

For both non-verbal FB and FS tasks, both older and younger children's performance was not improved across the three sessions, $\chi^2s(2, N = 13) < 5.20$, $ps >$

.17. Moreover, no significant improvement was shown by both older and younger children on both verbal FB and FS tasks across the three sessions, χ^2s(2, N = 13) < 6, ps > .07.

Control trials of non-verbal tasks.

Older children performed consistently above chance, $ts(12) > 7.64$, $ps < .001$, $rs > .91$ (non-parametric tests showed the same results: Wilcoxon $Zs < -3.31$, $ps < .001$, $rs < -.65$). Younger children also performed above chance, $ts(12) > 5.67$, $ps < .001$, $rs > .85$ (Wilcoxon $Zs < -3.07$, $ps < .01$, $rs < -.60$), except the working memory control trials on which they performed at chance level, $ts(12) < 1.47$, $ps > .17$ (Wilcoxon $Zs > -1.37$, $ps > .21$).

There were significant main effects of age, $F(1, 24) = 8.16$, $p < .01$, $r = .50$ for the response inhibition control trials and $F(1, 24) = 9.44$, $p < .01$, $r = .53$ for the working memory control trials. However, no effect for task and no interaction was found for both of the control trials, $Fs(1, 24) < .86$, $ps > .36$. Non-parametric tests also showed that older children performed better than younger children on both of the control trials (Mann-Whitney $Us < 52$, $ps < .05$, $rs = -.46 - -.48$) and there was no significant difference between tasks, Wilcoxon $Zs > -.92$, $ps > .48$.

Test versus control trials of non-verbal tasks.

There were significant differences between performance on the test trials, response inhibition control trials, and working memory control trials, $F(2, 50) = 14.07$, $p < .001$, $\eta_p^2 = .36$ for the FB task, and $F(2, 50) = 12.89$, $p < .001$, $\eta_p^2 = .34$ for the FS task. For the FB task, performance on the test trials was significantly worse than that on the response inhibition control trials ($p < .001$) and marginally less good than that on the working memory control trials ($p < .06$); whereas performance on the response inhibition control trials was significantly better than that on the working memory control trials ($p < .05$). The same pattern of performance was shown for the FS task: performance on the test trials was worse than that on the response inhibition control trials ($p < .001$) and working memory control trials ($p < .05$); whereas performance on the response inhibition control trials was better than that on the working memory control trials ($p < .05$). Similar results were found with equivalent non-parametric analyses: performance on both of the control trials was significantly better than that on the test trials (Wilcoxon $Zs < -2.29$, $ps < .05$, $rs < -.32$), and performance on the response inhibition control trials was also better than that on the working memory control trials (Wilcoxon *Zs* < -2.38, *ps* $< .05$, *rs* $< -.33$).

Filler trials of non-verbal tasks.

Ceiling effects were shown on both true representation and clue confirmation filler trials (see Table 3.9). Moreover, there were only 15/78 and 8/78 errors on the true representation filler trials of the non-verbal FB and FS tasks respectively. Out of the 86 correct responses to the test trials of the tasks, there were only 16 which were paired with an incorrect response to the true representation filler trial in the same task. Hence, participants were not likely using a strategy of always pointing to the opposite box from the one indicated by the clues.

Discussion.

The purpose of Experiment 4 was (1) to replicate the finding that the equivalence between the FB and FS tasks holds in both verbal reality-known and non-verbal reality-unknown versions, and (2) to validate the new non-verbal FS task. Results showed a replication of the association between the FB and FS tasks. Furthermore, the association remained significant regardless of whether the tasks are verbal reality-known or non-verbal reality-unknown and whether VMA was held constant. This finding suggests that the FB and FS tasks shared the same underlying requirement of representational understanding rather than language and cognitive inhibition. By comparing the two versions of the FB and FS tasks, the new non-verbal FS task was validated based on the findings that performance on the non-verbal FS task did not differ from that on the verbal FS task and the two FS tasks were strongly correlated ($r = .58$). Thus, the new non-verbal FS task was appropriate to be further employed to assess false-sign understanding in populations who have language and cognitive inhibition problems such as atypically developing children.

The two versions of the FB tasks were also shown to be comparable, replicating the results of Call and Tomasello's (1999). However, neither the two versions of the FB task nor the two versions of the FS task remained as significantly correlated after controlling for children's VMA. Unfortunately, none of the previous studies that have employed the same verbal and non-verbal FB tasks as those in the current experiment have measured participants' VMA (Call & Tomasello, 1999; Figueras-Costa & Harris, 2001). Hence, it was unclear whether or not the non-significant associations found between the two versions of the FB and FS tasks after controlling for VMA in the current experiment were unusual. Nevertheless, Figueras-Costa and Harris compared the performance of deaf children of hearing parents on their first non-verbal FB test trial versus their unique verbal FB test trial. There were eleven children who either passed or failed both trials whereas 10 children passed either one of the trials (9 passed the non-verbal trial whereas 1 passed the verbal trial). Using a Pearson's chi-square test, the two trials were shown to be not associated, $\chi^2(1, N = 21) = 1.05$, $p =$ .61 with a phi-coefficient of .22. This finding provided a support for the non-significant association found between the two versions of the FB task after controlling for VMA in the current experiment. On the contrary, both Call and Tomasello's study and the current experiment found associations between the two versions of the FB task before controlling for typically developing children's VMA. Taken together, these findings suggest that language ability has an effect on determining whether or not the two versions of the FB task were associated. The same probably held between the two versions of the FS task given that the FS and FB tasks were highly associated.

Finally, participants' performance on the non-verbal tasks could not be explained by their inability to meet the incidental executive demands of the tasks, learning through feedback across sessions or an adoption of incorrect strategies, replicating those findings in Experiment 3. Despite that, participants' motivation of doing their best in the tasks might not be sufficiently induced by just having verbal compliments from the experimenter in both Experiment 3 and 4. Although the tasks were embedded in a hiding-and-finding game context, finding a block was not a very interesting goal for children. Having children choose some stickers they preferred, asking them to find the stickers and receiving the stickers as rewards for success, Call and Tomasello's (1999) procedure definitely had a better chance of getting children motivated. However, I would suggest that the less motivating procedure used in both Experiment 3 and 4 would not be a critical problem based on the observation that participants in both experiments usually showed excitement when the feedback demonstrating the block's location was displayed.

In conclusion, both Experiment 3 and 4 successfully showed that the novel non-verbal reality-unknown FS task was valid in tapping false-sign understanding and closely related to Apperly et al.'s (2004) FB task. Having established the validity of these non-verbal false representation tasks in typically developing children, they can be applied in other young populations especially those have problems in language and cognitive inhibition. Children with Autism Spectrum Disorder (ASD) who are widely suggested to be specifically impaired in understanding minds as well as having general deficits in language and executive function would be a valuable population to re-examine the applicability of the tasks and the non-specificity claim of false-belief understanding. This is therefore the aim of the next chapter which reports an experiment investigating children with ASD's performance on the novel FS task as well as Apperly et al.'s (2004, 2007) FB and FP tasks.

Moreover, in this chapter, I suggested that Apperly et al.'s (2007) FP task, corresponding to the standard FP task, may be passed by employing an understanding of correspondence. This possibility would be further evaluated in children with ASD with the inclusion of Apperly et al.'s FP task in the experiment reported in Chapter 4. If children with ASD passed the standard FP task with the understanding of correspondence but failed the standard FB task due to an impaired understanding of representation, they would probably pass Apperly et al.'s FP task but fail their FB task and the novel FS task. Moreover, children with ASD were matched with typically developing children in Experiment 5. This enabled us to test whether the finding of Experiment 3, that older typically developing children improved on the FP task but not the FB and FS tasks across sessions, could be replicated.

Chapter 4 Autism Spectrum Disorder (ASD): A Special Case?

In Chapter 1, it was proposed that the question of whether understanding mental representations is underpinned by a general concept of representational understanding cannot be completely addressed by looking at typical development only. Although the previous empirical chapters of this thesis consistently suggested that children's understanding of representation underpins their performance on both false mental and false non-mental representation tasks, whether this finding can be generalised to atypical developing populations is unknown. The atypical case of Autism Spectrum Disorder (ASD) has been postulated as being caused primarily by a specific deficit in "theory of mind" (e.g., Baron-Cohen et al., 1985; for a review, see Tager-Flusberg, 2007). If individuals with ASD are selectively impaired in understanding mental representations, it would provide evidence for the existence of a domain-specific mechanism for understanding mental representations, contradicting the results found in typically developing children. Moreover, children with ASD are also known to have impairments in language and executive function which may confound with their difficulties in understanding mental representations. Therefore, Chapter 4 investigates whether ASD is a special case in which a dissociation of ability to understand mental versus non-mental representations may be seen, independent of other cognitive abilities, namely language and executive function.

Impaired Mental but Intact Non-Mental Representational Understanding?

Previous studies have shown that about 80% of children with ASD do not understand false beliefs when compared with typically developing children of a similar mental age range (4 to 6 years; e.g., Baron-Cohen, 1989b; Baron-Cohen et al., 1985; Leslie & Frith, 1988; Perner, Frith, Leslie, & Leekam, 1989; for meta-analyses, see Happé, 1995; Yirimiya et al., 1998). In addition to these false-belief studies, later studies, using tasks previously mentioned in Chapter 1 that tap the understanding of mental representations, such as the "level 2 visual perspective-taking" task (e.g., Masangkay et al., 1974) and the "appearance-reality distinction" task (e.g., Flavell et al., 1983), also showed that children with ASD performed significantly worse than children with typical development or mental retardation (e.g., Down's syndrome). Two examples of these studies are briefly described below.

Hamilton, Brindley, and Frith (2009) showed children a toy (e.g., a panda) which was placed on a square turntable and trained them to use pictures which were taken from the front, back, left and right side of the toy to answer questions. During the test phase, children were asked two types of questions. One type of question tests the understanding that visual perception is a representation which may represent the same referent (panda) in different senses (front, back, left or right side) depending on which viewpoint the referent is being viewed from (i.e., level 2 visual perspective-taking: "This is Susan. When I lift the pot, which panda will Susan see?"). Another type of question controls for the ability to imagine the rotation of the toy on the turntable (i.e., mental rotation: "When I lift the pot, which panda will you see?"). Hamilton et al. found that children with ASD have difficulty on the former type of question compared to the latter type of question, relative to typically developing children. This finding suggests that children with ASD have genuine impairments in level 2 visual perspective-taking which cannot be accounted for by problems in mental rotation.

Another example comes from the "appearance-reality distinction" task in which an object's appearance is distinct from its real identity. For example, a bottle of milk was shown to children, who were asked to name it and its colour. Then, an orange filter was placed in front of the bottle of milk and children were asked the appearance question, "Now what colour does the milk look?", and the reality question, "What colour is it really?" In order to answer both questions correctly, children need to understand that their visual perception represents the bottle of milk (referent/reality) as orange in colour (sense/appearance) and clearly distinguish between referent/reality and sense/appearance. Baron-Cohen (1989a) found that children with ASD performed worse than children with Down's syndrome and tended to commit more phenomenist errors by answering that the bottle of milk looked orange and really was orange.

This impairment in understanding mental representations in children with ASD was suggested to be specific due to the findings that children with ASD performed poorly on the FB task but well on a non-mental representation task (i.e., the FP task; Leekam & Perner, 1991; Leslie & Thaiss, 1992). The same pattern of dissociation between mental and pictorial understandings was also found in children with ASD, using the "false" drawing task (Charman & Baron-Cohen, 1992). More recent studies further investigated children with ASD's competence on a different type of non-mental representation task which has also been mentioned in Chapter 1 – the ambiguous figures task (Ropar, Mitchell, & Ackroyd, 2003; Sobel, Capps, & Gopnik, 2005; Wimmer & Doherty, 2010). Despite children with ASD's failures on mental representation tasks such as the FB task, these studies indicated that children with ASD could reverse an ambiguous figure by switching from one of its interpretations to the other after being informed about the ambiguity of the figure. Wimmer and Doherty (2010) even suggested that this reversal ability of children with ASD was underpinned by an understanding of pictorial representation.

However, as discussed in Chapter 1, the photographs and drawings used in the FP and "false" drawing tasks were not false and no distinction between reference and sense was required. The same issue also holds for the ambiguous figures task. Without the requirement of distinguishing reference from sense, it is not clear whether a satisfying performance on these tasks implies a genuine understanding of non-mental representations or a simple understanding of correspondence. Children with ASD might understand that a photograph/drawing corresponds to the scene at the time the photograph/drawing was taken/made, and an ambiguous figure corresponds to two different objects so they succeeded on these tasks. They might even learn that a non-mental representation (e.g., a word) corresponds to something through simple association. A recent study (Preissler, 2008), using the same paradigm as that in Preissler and Carey (2004) mentioned in Chapter 1, taught low-functioning children with ASD (IQ under 70) a novel word (e.g., 'whisk') for a picture of a novel object (e.g., a whisk) and asked them to indicate the word from a choice of the picture and a real object (e.g., a real whisk). Unlike typically developing children, Preissler found that low-functioning children with ASD made associative mappings between words and pictures and failed to transfer a word learned for a picture to their corresponding real referent. Thus, children with ASD's good performance on the FP, "false" drawing and ambiguous figures tasks could be explained by lower-level cognitive abilities rather than a real understanding of non-mental representations.

The only task, until now, that assessed children with ASD's true understanding of non-mental representations was the false signal task (Bowler et al., 2005). The signal in the task falsely represented a plane's location (referent) as a different location (sense). In order to correctly predict the destination of an automatic train which picked up cargo from planes based on the signal, participants had to differentiate the signal's sense from its referent. Bowler et al. showed that children with ASD's performance on the false signal task was comparable to their performance on the FB task. Moreover, strong associations (phi-values > .55) were found between the two tasks with two different samples of children with ASD in the two experiments of Bowler et al.'s study. Therefore, with a non-mental representation task that requires a distinction between reference and sense, children with ASD show the same deficit as they have in the FB task. The suggestion that ASD selectively impairs an individual's understanding of mental representations may require revision.

Language and Executive Dysfunctions or Deficit in Representational Understanding?

Previous research has illustrated that language abilities vary widely among individuals with ASD (e.g., Charman et al., 2003; Kjelgaard & Tager-Flusberg, 2001; Smith, Mirenda, & Zaidman-Zait, 2007). However, it is commonly found that both language comprehension and production are delayed in children with ASD (e.g., Charman et al., 2003; Charman, 2004; Mitchell et al., 2006). Moreover, two recent studies with large samples of children with ASD aged 24 to 59 months showed that receptive language ability is more severely delayed or impaired than expressive language ability in ASD (Ellis Weismer, Lord, & Esler, 2010; Hudry et al., 2010). There is also evidence demonstrating that the development of different aspects of language (i.e., pragmatic, lexical and grammatical) is strongly related to false-belief understanding in individuals with ASD across ages and levels of ability (e.g., Fisher, Happé, & Dunn, 2005; Paynter & Peterson, 2010; for a review, see Tager-Flusberg, 2000).

Individuals with ASD are also found to be impaired in executive function, including inhibitory control and working memory (e.g., Luna et al., 2007; Robinson et al., 2009; for a review, see Hill, 2004). Previous research has indicated that children with ASD who had difficulty on the FB task also showed impairments on tasks tapping executive function (e.g., Ozonoff, Pennington, & Rogers, 1991; Ozonoff & McEvoy, 1994; Pellicano, 2007; Zelazo, Jacques, Burack, & Frye, 2002). Russell (1997) even suggested executive function as a contributor to individuals with ASD's performance on the FB task. Russell, Saltmarsh, and Hill (1999) further argued that the difference between children with ASD's performance on the FB and FP/"false" drawing tasks was because of the different executive demands of the tasks. Although the FB and FP/"false" drawing tasks all elicit a prepotent response of acting in accordance with reality which is in conflict with the representation in question, a photograph/drawing might be more salient than a belief so that the requirement of resisting interference from reality would be less demanded in the FP/"false" drawing tasks than the FB task. To evaluate this proposal, Russell et al. designed a modified FP task in which a photograph of a screen was taken. An object was then placed in front of that screen. By having no object in the photograph, the photograph became less salient and thus this modified FP task might become as demanding as the FB task. Children with ASD were tested on this modified FP task and the standard FP and FB tasks. No better performance on the modified FP task than the FB task was found. However, better performance on the standard FP task than the other two tasks was indicated. Therefore, Russell et al. concluded that executive function was a critical factor, explaining children with ASD's failure on the FB and modified FP task as well as their success on the standard FP task.

Hence, it is possible that individuals with ASD's impairments in language and executive function rather than false-belief understanding determine their performance on the FB task. As mentioned in Chapter 1, although children with ASD's problem in understanding false beliefs was usually concluded from a comparison with a control group which matched the ASD group on verbal mental age and/or general intellectual ability, verbal mental age and general intellectual ability were measured by independent tests which did not measure the exact incidental cognitive demands of the FB task. It is thus possible that the ASD and control group are different in terms of their ability in meeting the incidental cognitive demands of the FB task. Furthermore, the impairments in executive function of children with ASD are found to be independent of verbal and general intellectual abilities (Robinson et al., 2009). 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.

In the following, I briefly review three recent studies that have attempted to disentangle children with ASD's problem with false-belief understanding from their impairments in language and executive function. Two of these studies used the non-verbal FB paradigm of Southgate et al. (2007) to investigate individuals with ASD's spontaneous anticipatory looking of a protagonist's searching location based on the false belief the protagonist was holding. The first study was Senju, Southgate, White, and Frith (2009) which tested adults with Asperger syndrome, whereas the other study was Senju et al. (2010) which tested children with ASD. Results of these two studies showed that individuals (both adults and children) with ASD failed to show correct action anticipation that is in line with the protagonist's false belief. Although the ASD group was not matched with the control group on verbal mental age in Senju et al. (2010), the group difference on anticipatory looking remained significant when verbal mental age was covaried out. These findings suggest that individuals with ASD were impaired in false-belief attribution which was independent of verbal ability. However, as mentioned in Chapter 1, looking behaviours are ambiguous and alternative explanations are possible. For example, individuals with ASD's lack of correct spontaneous anticipatory looking may be generally directed to social stimuli (Ruffman, Garnham, & Rideout, 2001) rather than specifically to false-belief attribution. Moreover, the control group's success could be explained by behaviour rules (see Chapter 1; Perner & Ruffman, 2005; Ruffman & Perner, 2005). Therefore, this anticipatory-looking paradigm does not necessarily provide evidence that a false belief is attributed or genuinely understood as a *false representation* which involves a difference between reference and sense.

Using a non-verbal FB task which was adapted from Call and Tomasello (1999), Colle et al. (2007) tapped the genuine understanding of mental representations in children with ASD and children with specific language impairment who were matched on their very low levels of language development. In this task, children were asked to find an object with the help of a protagonist who put a marker on the location where she thought the object was. Children needed to understand that the protagonist's false belief represented the object's location (referent) as the location where she put the marker (sense) and reason the object's location to be the alternative location as they had witnessed the switch of locations. Thus, a distinction between reference and sense was required while cognitive inhibition of one's knowledge about reality was not needed. Results showed that children with ASD were significantly impaired in false-belief understanding relative to children with specific language impairment, suggesting a fundamental deficit in false-belief understanding in children with ASD. Only with a task like this can one conclude that children with ASD have impairment in understanding false beliefs which cannot be fully explained by their language and executive dysfunctions.

However, the FB task and the false signal task used in Bowler et al. (2005) did not disentangle the demands of language and executive function from that of representational understanding. It is therefore possible that the association found between these two tasks in children with ASD was due to language and executive function rather than representational understanding. In the following section, I report an experiment to investigate whether ASD selectively impairs an individual's understanding of mental representations or generally impairs his/her representational understanding or primarily impairs his/her language and executive skills which mask their competence of understanding mental and non-mental representations in standard verbal tasks. Ethics approval was provided by the Ethics Advisory Sub-Committee in the Department of Psychology in Durham University and the School of Psychology Ethics Committee in Cardiff University.

Experiment 5

The aim of Experiment 5 was to examine whether a dissociation between understanding of mental versus non-mental representation will be seen in the case of children with ASD. In order to critically address this issue, it was important to demonstrate that this dissociation is a core characteristic of ASD which cannot be explained by other cognitive impairments namely language and executive function. Hence, mental and non-mental representation tasks that require clear distinction between reference and sense while minimising, eliminating, and controlling for language and executive demands were essential. The non-verbal reality-unknown FB and FS tasks used in Chapter 3 met the criterion and were thus employed in the current experiment. Although the non-verbal reality-unknown FP task used in Chapter 3 did not meet the criterion of distinguishing reference from sense and did not involve the same level of incidental working memory demand, it was also employed in the current experiment for the following reasons. First, the inclusion of this FP task allowed an evaluation of whether this FP task resembled the standard FP task which could be passed by children with ASD given their spared understanding of correspondence, as proposed in Chapter 3. Second, if children with ASD could perform well on the non-verbal FP task, it would then serve as a control task in the experiment. This FP task and the control trials of each of the FB and FS tasks would allow a validation of equal executive abilities between children with ASD and comparison children. If the two groups of children with and without ASD did not differ in their performance on the FP task and the control trials of the FB and FS tasks but significantly differed in their performance on the false representation test trials of both FB and FS tasks, it would clearly suggest that children with ASD suffer a genuine impairment in understanding representations.

Typically developing children were recruited as comparison children to match with children with ASD for the following reasons. First, 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 with a unitary system of 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. Second, if typically developing children improved across sessions on the FP task but not the FB and FS tasks, the finding of Experiment 3 would be replicated. This finding would suggest that children might realise that the FP task could be passed by using their understanding of correspondence from the first session and significantly improved their performance afterwards.

Method.

Participants.

Eighteen 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 of them have Statements of Special Education Needs which confirmed their formal diagnoses of autistic disorder or Asperger syndrome, given by qualified clinicians, according to internationally accepted criteria (American Psychiatric Association, 1994; World Health Organisation, 1993). 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 is a screening measure for the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, Le Couteur, 1994) and has high diagnostic validity (Berument, Rutter, Lord, Pickles, & Bailey, 1999). With 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; $M = 26.50$; $SD = 6.34$) based on their parents' reports on the SCQ.

A comparison group of 18 typically developing children were recruited from two schools in South East Wales. Five of these children took part in Experiment 3 reported in Chapter 3 and their data was carried over to the current experiment for matching with the ASD group. The ASD and comparison groups were matched in terms of verbal mental age (VMA) and non-verbal intelligence quotient (NVIQ), as measured with the long form of the British Picture Vocabulary Scale - Second Edition (BPVS-II; Dunn et al., 1997) and the Brief Intelligence Quotient composite of the Leiter International Performance Scale-Revised (Roid & Miller, 1997) respectively. 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. Participant characteristics are presented in Table 4.1. 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), $t(27.40) = 5.59, p < .001$ with an effect size of $r = .73$. 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. All participants were recruited with parental informed consent.

Table 4.1

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 3, except that the between-participants factor was group (children with ASD versus typical development) rather than age.

Results.

Children received a score of 1 for each of the trials if they correctly identified the object's location. Scores across the three sessions were summed and the mean scores for each trial type of the FB, FP and FS tasks are presented in Table 4.2. The distribution of the data was negatively skewed for most of the dependent variables and data transformations, including square root, log and inverse transformations, did not improve normality. However, the outcomes of the following parametric analyses were corroborated by running non-parametric analyses. Homogeneity of variance was violated for some of the variables according to Levene's test (Levene, 1960), but due to equal sample sizes this is unlikely to constitute a problem (Cardinal & Aitken, 2006).

Table 4.2

Mean Scores on Each Trial Type of the False Belief, False Photograph and False Sign Tasks (Standard Deviations are Shown in Brackets)

			Task (range of scores = 0 – 3)		
Trial	Group	N	False Belief	False Photo	False Sign
False representation	ASD	18	1.67 (1.08)	2.28 (.89)	1.89 (1.13)
	Comparison	18	2.61 (.70)	2.67 (.49)	2.72 (.57)
Response inhibition	ASD	18	2.89 (.47)	2.83 (.38)	3 (0)
	Comparison	18	3 (0)	2.94 (.24)	3 (0)
Working memory	ASD	18	2.56 (.78)	2.67 (.77)	2.67 (.59)
	Comparison	18	2.28 (1.02)	2.72 (.46)	2.56 (.92)
True representation	ASD	18	2.67 (.77)	2.94 (.24)	2.72 (.46)
	Comparison	18	2.61 (.61)	2.94 (.24)	2.89 (.47)
Clue confirmation	ASD	18	3 (0)	3 (0)	3 (0)
	Comparison	18	3 (0)	3 (0)	3 (0)

Note. ASD = Autism Spectrum Disorder

Data from the false representation test trials and the control trials were analysed separately. Performance on each of these trials was compared against chance and tested using a mixed between- and within-participants ANOVA with group (ASD or comparison) as between-participants factor and task as within-participants factor. Performance on the false representation test trials of each group was also subjected to (1) planned contrasts to test whether the ASD group performed significantly worse than the comparison group on the FB and FS tasks but not on the FP task; (2) correlation tests to examine the degree of associations between the FB, FP and FS tasks; and (3) within-participants analyses to determine whether participants improved across the three sessions given that feedback was provided after each trial. As the range of scores participants obtained in each session was 0 to 1, non-parametric Friedman tests were used to examine whether there were improvements across sessions. Finally, performance on the filler trials was examined to exclude the possibility that participants' performance was determined by an adoption of incorrect strategies such as always pointing to the opposite box from the one indicated by the clues provided in the tasks. Results from non-parametric equivalent tests were also reported for each set of the parametric analyses.

False representation test trials.

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, $t(17) = 3.69, p < .01$ with an effect size of $r = .67$. On the contrary, performance on all three tasks was significantly above chance for the comparison group, $ts(17) > 6.76, ps < .001, rs > .85$.

A 2 (group) X 3 (task) ANOVA revealed significant main effects of group and task, $F(1, 34) = 9.30, p < .01$ with an effect size of $r = .46$ and $F(2, 68) = 3.12, p = .05$ with an effect size of $\eta_p^2 = .08$ respectively, but no significant interaction between group and task, $F(2, 68) = 2.43, p = .10$. This pattern of findings remained the same, except the main effect of task was not significant, when chronological age, VMA, or NVIQ was included as a covariate. Planned contrasts revealed that the groups were different on the FB and FS tasks, $t(34) = -3.11, p < .01, r = .47$ and $t(34) = -2.79, p < .01, r = .43$ respectively, but not on the FP task, $t(34) = -1.62, p = .11$. To spell out the difference between the three tasks, pairwise comparisons using Bonferroni adjustment were performed. 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.

Equivalent non-parametric analyses indicated identical results. Children with ASD performed at chance level on the FB and FS tasks, Wilcoxon $Z = -.65, p = .58$ and Wilcoxon $Z = -1.41, p = .19$ respectively. However, they performed above chance on the FP task, Wilcoxon $Z = -2.83, p < .01$ with an effect size of $r = -.47$. On the contrary, the comparison group performed consistently above chance, Wilcoxon $Zs <$

-3.63, *ps* < .001, *rs* < -.61. Mann-Whitney U tests showed that the comparison group performed better than the ASD group on the FB and FS tasks (*U*s < 91, *ps* < .05 with effect sizes of *rs* < -.42) but their performance on the FP task did not differ ($U = 126$, $p = .23$). There was no significant difference between tasks, Friedman test: $\chi^2(2, N = 36) = 4.36, p = .12$.

The bivariate correlations between the FB, FP and FS tasks found with the ASD group are presented in Table 4.3 whereas those found with the comparison group are shown in Table 4.4. For the ASD group, both Pearson's and non-parametric Spearman's correlation tests showed that performance on the three tasks was significantly correlated. However, only the correlation between the FB and FS tasks remained significant when performance on the third task was controlled. When VMA was controlled, the association between the FB and FS tasks remained significant, $pr(15) = .61, p < .05$, but the correlations between the FB and FP and that between the FP and FS tasks fell to marginally significant, $pr(15) = .46, p = .06$ and $pr(15) = .43, p = .08$ respectively.

Table 4.3

Bivariate Raw Pearson's/Non-parametric Spearman's Correlations and [Third-Task-Controlled Correlations] (With *N*) Between the False Belief, False Photograph and False Sign Tasks for the ASD Group

	False Belief	False Photograph	False Sign
False Belief	-	.53*/.54* (18)	.69**/.66** (18)
False Photograph	[.29] (15)	-	.50*/.47* (18)
False Sign	[.58*] (15)	[.22] (15)	-

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

Table 4.4 indicated that performance of the comparison group on the three tasks was also significantly correlated, except the correlation between the FB and FP tasks fell to non-significance when the non-parametric Spearman's test was used. When the third task was controlled for, only the correlation between the FB and FS tasks stayed significant. The same pattern was also obtained when VMA was controlled for: $pr(15) = .51, p < .05$ between the FB and FS tasks, $pr(15) = .21, p = .42$ between the FB and FP tasks, and $pr(15) = .37, p = .15$ between the FP and FS tasks.

Table 4.4

Bivariate Raw Pearson's/Non-parametric Spearman's Correlations and [Third-Task-Controlled Correlations] (With N) Between the False Belief, False Photograph and False Sign Tasks for the Comparison Group

	False Belief	False Photograph	False Sign
False Belief	-	.46*/.40 (18)	.60**/.57* (18)
False Photograph	[.24] (15)	-	.49*/.49* (18)
False Sign	[.48*] (15)	[.30] (15)	-

Note. ** $p < .01$, * p = or $< .05$

Children with ASD improved their performance across the three sessions, $\chi^2(2, N = 18) = 10.89, p < .01$ for the FB task, $\chi^2(2, N = 18) = 7.75, p < .05$ for the FP task, and $\chi^2(2, N = 18) = 10.75, p < .01$ for the FS task. Further analyses showed that improvement from Session 1 to 2 and from Session 1 to 3 was significant 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 $.06$, $rs < -.35$. No improvement was shown from Session 2 to 3 for all three tasks, Wilcoxon $Zs > -.58$, $ps = 1$. 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, $\chi^2(2, N = 18) = 12, p < .01$. Further analyses showed that typically developing children's performance on the FP task was improved from Session 1 to 2 and from Session 1 to 3, Wilcoxon $Zs = -2.45$, $ps = .05$, $rs < -.41$, but no improvement was shown from Session 2 to 3, Wilcoxon $Z = 0, p = 1$.

Control trials.

Both groups performed above chance on the two types of control trials of the FB, FP and FS tasks, $ts(17) > 5.71$, $ps < .001$, $rs < -.81$ for the ASD group, and $ts(17) > 3.24$, $ps < .01$, $rs < -.62$ for the comparison group. The same was found with non-parametric tests: Wilcoxon $Zs < -3.49$, $ps < .001$, $rs < -.58$ for the ASD group, and Wilcoxon $Zs < -2.67$, $ps < .01$, $rs < -.45$ for the comparison group.

A mixed ANOVA with group as between-participants factor and task as within-participants factor was carried out for each type of the control trials. Mauchly's

(1940) tests were inspected for both types of the control trials and only the response inhibition control trials showed a departure from sphericity which was corrected using the Greenhouse-Geisser *epsilon* (Greenhouse & Geisser, 1959). For both types of the control trials, there were no significant main effects of group and task, Fs < 2.31, ps > .14, and no significant interactions between group and task, Fs < .64, ps > .53. Non-parametric tests revealed the same results: no significant difference between groups, Us < 140.50, ps > .47, and tasks, Friedman test: χ^2s(2, N = 36) > 4.77, ps > .10, for both types of the control trials.

Filler trials.

Ceiling effects on the two types of the filler trials were shown for all three tasks (see Table 4.2). Moreover, there were only 13/108, 2/108, and 7/108 errors (6/54, 1/54, and 5/54 were from the ASD group) on the true representation filler trials of the FB, FP, and FS tasks respectively. Out of the 250 correct responses to the test trials of all three tasks (106 were from the ASD group), there were only 16 (7 were from the ASD group) which were paired with an incorrect response to the true representation filler trial in the same task. These patterns of results suggested that participants were not likely using a strategy of always pointing to the opposite box from the one indicated by the clues and the two groups were not using different strategies.

Discussion.

Experiment 5 was the first experiment which successfully applied the novel non-verbal FS task devised for this thesis as well as Apperly et al.'s (2004, 2007) FB and FP tasks in a group of children with ASD. It investigated whether ASD is a special case in which there is a dissociation of an individual's ability to understand mental versus non-mental representations, independent of other impairments in language and executive function. If this was the case, it would provide evidence for the existence of a domain-specific mechanism for understanding mental representations which was impaired in ASD. Results showed that children with ASD performed worse on the false representation test trials of both FB and FS tasks but not on that of the FP task relative to typically developing children. No significant difference between the false representation test trials of these three tasks was found. Moreover, the association between the false representation trials of the FB and FS tasks remained significant for the ASD group no matter whether VMA or the FP task was controlled or not. These findings suggest that children with ASD were not selectively impaired in understanding mental representations but generally impaired in understanding mental and non-mental representations. Moreover, the same as typically developing children, both mental and non-mental representations were processed with a unitary system of representational understanding, which was however impaired, in ASD. Despite that, children with ASD were capable of improving their performance over sessions, showing significant improvement on the FB and FS task but marginally significant improvement on the FP task. This might be due to an already good performance on the FP task (10 out of 18 children with ASD passed) but relatively poor performance on the FB and FS tasks (7 out of 18 passed in each) in the first session. Their better performance on the FP task was consistently found in previous studies (e.g., Leekam & Perner, 1991), probably due to the lack of requirement of distinguishing reference from sense and could be passed by employing a simple understanding of correspondence. However, children with ASD's improvement on the FB task was in contrast with Grant et al.'s (2007) finding that children with fragile X syndrome or intellectual disability did not improve their performance on the FB task. This difference in finding might be due to the different intervals between trials employed by the current experiment versus Grant et al.'s. The current experiment had a week in between the trials whereas Grant et al. presented the trials consecutively. This might suggest that atypically developing children may require time to show an improvement in performance. There is a recent 3-year longitudinal study by Pellicano (2010a) which indicated that children with ASD made substantial progress over time in false-belief understanding.

Although children with ASD performed well on the standard FP task and the non-verbal FP task, previous evidence has shown that they failed Russell et al.'s (1999) modified FP task in which the photograph used was also a true representation that corresponded with its referent and sense. If children with ASD succeeded on the former two FP tasks with their spared understanding of correspondence, why would they fail the latter? I would suggest that children with ASD failed Russell et al.'s modified FP task because of its higher executive demands and its unnatural nature of taking photograph of a screen, as Russell et al. has suggested. Children with ASD's performance on the modified FP task was no better than that on the FB task in Russell et al.'s study, suggesting that the two tasks were of similar executive demands. However, no association was found between these two tasks in their study, suggesting that the two tasks were probably not tapping the same conceptual understanding. The modified FP task probably tested an understanding of correspondence, whereas the FB task tested an understanding of representation.

For the comparison group, the bivariate correlations found between the false representation test trials of the three tasks, including those controlled for VMA or the third task, were less significant than those found in Experiment 3 reported in Chapter 3. However, this could be explained by the smaller sample size of children with a narrower age range that was recruited in the current experiment compared to Experiment 3. Regarding the improvement shown across sessions, the finding of improvement on the false representation test trials of the FP task but not on those of the FB and FS tasks in the older children of Experiment 3 was replicated in the current experiment. As in Experiment 3, children started off with the same level of performance on the false representation test trials across the three tasks in the first session (14 passed the FB task, 12 passed the FP task, and 14 passed the FS task). Both children with ASD and older children with typical development showed a different degree of improvement on the FP task compared to the FB and FS tasks, providing further evidence for the FP task's lack of strong associations with the FB and FS tasks.

Finally, children with ASD performed above chance on the control trials of the FB, FP and FS tasks and their performance was not different from that of the comparison group, suggesting that they were capable of meeting the incidental cognitive demands of the tasks. Furthermore, the patterns of the results found in the filler trials of the tasks suggested that children did not use an incorrect strategy of always pointing to the opposite box and the two groups were not using different strategies. Taken together, children's performance on the false representation test trials reflects their genuine understanding of representations rather than other cognitive skills and incorrect strategy use.

With the FP task and the control trials of the tasks, the two groups were verified as comparable in their ability to meet the incidental requirements of the tasks although they were not able to be matched on their chronological age. However, further studies were required to apply these tasks in different atypically developing populations, such as children with intellectual disability, to compare with children with ASD as well as children with typical development. This would extend the current research in testing the non-specificity claim of false-belief understanding and provide an evaluation of the findings reported in this thesis.

To conclude, ASD is not a special case in which there is a dissociation of an individual's ability to understand mental versus non-mental representations. Rather, children with ASD also process mental and non-mental representations with a unitary system of representational understanding. However, this system of representational understanding in ASD is impaired. Moreover, this impairment in ASD is independent of its language and executive dysfunctions. The possible underlying causes of this impairment are discussed in the following chapter.

Chapter 5 General Discussion

In this chapter, I first summarise the results from the previous three empirical chapters and discuss them in terms of the theoretical issues outlined in Chapter 1. Chapter 1 provided a basis upon which I built up this thesis's theme: Children's understanding of false representations. The most frequently investigated false representation is false belief. However, whether children's understanding of false beliefs is underpinned by a general system for understanding representations or a domain-specific mechanism for processing mental-related information is continuously debated. This debate is the focus of this thesis and it has two implications. One contributes to the theoretical account of cognitive development and the other extends to our understanding of the human brain. Results from the empirical chapters of this thesis are thus discussed regarding to these two implications. Based on this discussion, suggestions for future research are made in the last part of this chapter.

Summary of the Main Findings

In order to address the debate in question, the previous empirical chapters introduced new FS tasks as non-mental comparisons of different FB tasks to investigate whether children's understanding of false beliefs is equivalent to their understanding of false signs through different experimental approaches. If equivalence was to be found that could not be explained by other alternatives except for a genuine understanding of representations, it would evidently support the general account of understanding false beliefs which suggests that both mental and non-mental representations are processed with a unitary system of understanding representations. Different experimental approaches include (1) manipulating children's understanding of false beliefs versus false signs; (2) controlling confounding factors of language and executive function; and (3) comparing typical with atypical development in understanding false beliefs versus false signs.

Chapter 2 examined the equivalence by manipulating false-belief versus false-sign understanding through training and testing for transferability after training. Both Experiment 1 and 2 used the FB and FS tasks that were verbal and reality-known. These FB and FS tasks were found to be associated in both experiments. However, this association could be explained by a developmental coincidence between two distinct concepts, one for understanding mental representations and the other for understanding non-mental representations. Experiment 2 raised evidence against this possibility by showing a transfer effect from false-belief training to false-sign understanding and the contribution of false-sign training to the understanding of one's own false beliefs.

Chapter 3 dealt with the equivalence by excluding the possible contribution of other cognitive skills that are normally required in both standard FB and FS tasks. Experiment 3 and 4 indicated that the association between tasks testing false beliefs and false signs remained significant when the demands of language and cognitive inhibition were minimised or eliminated. Furthermore, the association between the two tasks was also found to be present whether or not participants' verbal mental age (VMA) was held constant, providing further evidence that language ability was not determining the relationship between the two tasks. Therefore, the association between the tasks probably reflects the same underlying representational understanding rather than language and cognitive inhibition processing.

Chapter 4 further explored the equivalence by testing its existence in a well-known type of atypical development – Autism Spectrum Disorder (ASD), which impairs an individual's mental-related processing, language and executive function, relative to that found in typical development. Experiment 5 showed that children with ASD were impaired in both non-verbal reality-unknown FB and FS tasks but the association between the two tasks remained significant no matter whether VMA or the non-verbal reality-unknown FP task was controlled. This suggests that children with ASD were not selectively impaired in understanding mental representations but generally impaired in understanding representations relative to that of typical developing children while general language ability, cognitive and executive functions, and ability in meeting task requirements were controlled with different measures.

To conclude, an association between the FB and FS tasks was consistently found in the series of experiments reported in Chapter 2 to 4. This association was also found in previous studies with different versions of FB and FS tasks and different samples of children (Bowler et al., 2005; Leekam et al., 2008; Parkin, 1994; Sabbagh, Moses, et al., 2006, Experiment 2). Hence, it has important implications for the theoretical account and neurological basis of cognitive development.

Developmental Accounts of Understanding False Belief: Domain-Specific or not?

In Chapter 1, I outlined two developmental accounts for the understanding of false beliefs. One is the domain-specific account which indicates that false-belief understanding relies on an innate specific mechanism. This mechanism matures according to its own timetable and is dedicated to mental-related information only (e.g., Leslie, 1987, 1994). As children possess this mechanism early in life, their failure on the FB task is due to their lack of other general cognitive skills required in the task, such as language and executive function. In contrast, the other account is the general system approach which suggests that children's understanding of false beliefs depends on a general theory of representation (e.g., Perner, 1991; Wellman, 1990). This general theory develops with age through experience and it supports children's understanding of both mental and non-mental representations.

The consistent association between the FB and FS tasks found in previous studies and the series of experiments in this thesis obviously supports the general system approach. However, the experiments in this thesis added new evidence to the literature: (1) the general system for understanding representations develops gradually through experience; and (2) it sufficiently explains the equivalence between the FB and FS tasks, independent of other shared cognitive skills namely language and cognitive inhibition.

The argument in this thesis is that experience is key to development rather than natural maturation. Furthermore, other cognitive skills cannot explain the associated performance on the FB and FS tasks. Therefore, the position taken is that of a general system account. I discuss these two issues below in separate sections. I then extend the discussion on the general account of understanding mental representations to the conception of "theory of mind" and ASD.

Experience: Key of the development of metarepresentation.

In Chapter 2, Experiment 2 demonstrated a conceptual change in understanding representations through training and transferability between the FB and FS tasks. Evidently, the experience of repeated exposure of tasks and feedback that children gained over time from pre-test to training sessions is crucial to effect the conceptual change in understanding representations. Previous training studies on false-belief understanding also converged on the same conclusion that experience of repeated exposure of tasks and feedback that is incorporated and integrated over several sessions is important for the development of false-belief understanding (e.g., Amsterlaw & Wellman, 2006; Appleton & Reddy, 1996; Lohmann & Tomasello, 2003; Slaughter, 1998, Study 2).

The experience of repeated exposure of tasks and feedback also involved a social child-adult interaction component which may facilitate children's conceptual change in understanding representations. Previous studies have demonstrated that children's experience of social interaction has an influence on their development of false-belief understanding. For example, the well-known "sibling effect" suggests that children who have experience in interacting with older siblings illustrate an earlier understanding of false beliefs (e.g., Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, &Berridge, 1996; Ruffman, Perner, Naito, Parkin, & Clements, 1998). Longitudinal studies also indicated that family talk about mental states was related to later false-belief understanding (e.g., Brown, Donelan-McCall, & Dunn, 1996; Ruffman, Slade, & Crowe, 2002). Although the social interaction between a child and the experimenter in Experiment 2 was much less extensive than that between the child and his/her family members in everyday life, it was a possible factor of experience which children have gained through training and might have helped in enhancing their understanding of representation.

Although the experiments reported in Chapter 3 and 4 also involved several sessions with tasks, feedback and social interaction between children and the experimenter, no significant improvement was shown in typically developing children, especially on the FB and FS tasks. One might argue that this finding contradicted that of Experiment 2. However, there is a fundamental difference between Experiment 2 and the experiments in Chapter 3 and 4 that may explain the difference in the effect of experience shown. That is, the experiments in Chapter 3 and 4 did not screen children to ensure their lack of false-belief understanding before inclusion. Without this manipulation, some typically developing children in these experiments may already have some kind of representational understanding so a significant improvement across sessions was less likely to occur. On the contrary, children with ASD in Experiment 5 reported in Chapter 4 significantly improved their performance across sessions, especially on the FB and FS tasks. Despite that, their mean performance across sessions was significantly worse than that of typically developing children, suggesting a fundamental impairment in understanding representations in ASD. All these findings, taken together, might suggest that experience is important for children who initially lack a conceptual understanding of representation, such as the typically developing children included in training in Experiment 2 and children with ASD, to gradually develop this understanding or improve task performance (as the findings in Experiment 5 did not suggest a conceptual change of representational understanding in ASD).

Other cognitive skills: Language and executive function.

The experiments in Chapter 3 and 4 indicated that the equivalence between tasks testing false beliefs and false signs was not due to shared demands of language and cognitive inhibition. Especially, the association between the two tasks remained, whether or not VMA was controlled, and regardless of whether the sample had a diagnosis of ASD. There is other evidence that language and cognitive inhibition might not be the crucial sources of difficulty for children in a false representation task based on the following logic. If language and cognitive inhibition placed a critical demand on children when they were tested with the standard FB/FS task, better performance on the non-verbal reality-unknown version of the task which did not require the comprehension of linguistic narratives and cognitive inhibition of knowledge about reality should have resulted. However, this was not the case. Experiment 4 in Chapter 3 showed no significant difference but correlation between the verbal reality-known and non-verbal reality-unknown versions of both FB and FS tasks. Call and Tomasello (1999) also found that children passed the non-verbal reality-unknown FB task at the same age as they passed the verbal reality-known FB task. Another study (Grant et al., 2007) which tested children with fragile X syndrome and children with intellectual disability yielded no significant difference between three types of FB tasks: (1) the non-verbal reality-unknown FB task, (2) a standard verbal reality-known FB tasks, and (3) Samson et al.'s (2005) non-verbal reality-known FB task.

Colle et al. (2007) further demonstrated a dissociation between children with specific language impairment's language competence (very low level) and their performance on the non-verbal reality-unknown FB task (76% correct), suggesting that language and false-belief understanding are relatively independent. Other studies found that eliminating the demand of cognitive inhibition in the FB task did not facilitate 3-year-olds' performance (Russell, 1996). In addition to language and cognitive inhibition, response inhibition and working memory might also not be an alternative explanation for children's performance on false representation tasks. Results of Experiment 3 and 4 showed that where participants failed the false representation test trials of the tasks it was neither because of their difficulty in inhibiting their tendency to act according to the clues nor because of their difficulty in remembering and processing the sequence of the events while working out the location of the object. In a microgenetic study, children tested with a battery of response inhibition and false belief tasks every four weeks for six phases of testing showed good performance on response inhibition before having a good understanding of false beliefs (Flynn, O'Malley, & Wood, 2004). Moreover, Sabbagh, Xu, Carlson, Moses, and Lee (2006) showed that Chinese preschoolers had advanced executive skills but this advantage did not lead to advanced performance on the FB task. All these findings suggest that children's performance on false representation tasks could not be reduced to their ability in language and executive function, but rather it is determined by a general conceptual understanding in representation.

The control trials of the non-verbal reality-unknown false representation tasks described in Chapter 3 and 4 teased apart the incidental executive demands of the tasks. This design of the tasks had three advantages. First, it allowed a check of whether children could meet the incidental executive demands. Second, it served as a measure of executive abilities to ensure an equality of confounding variables between children with ASD with the comparison group. Third, it provided a means to evaluate the equivalence between the tasks in terms of their incidental executive demands. Performance on these control trials showed that the FB and FS tasks imposed similar levels of response inhibition and working memory demands on children. In this sense, response inhibition and working memory might explain part of the equivalence between the FB and FS tasks although language and cognitive inhibition have been found to be not involved. This is reasonable because the false representation test trials of these tasks involved the demands of response inhibition and working memory but not language and cognitive inhibition.

Compared to the non-verbal reality-unknown FB and FS tasks, the non-verbal reality-unknown FP task involved less working memory demand. This might be related to the different functions of photographs versus beliefs and signs. Photographs show situations at the time they are taken, whereas beliefs and signs represent whatever they are supposed to (usually up-to-date with respect to reality). Due to the changeability of beliefs and signs (especially mechanical signs like the one used in the non-verbal reality-unknown FS task and traffic lights), more working memory capacity might be required when reasoning about their false version than "false" photographs, which are fixed once they have been developed. Sabbagh, Moses, et al. (2006) also suggested that executive functioning may be required to think about representations that are intended to be up-to-date representations of reality, i.e., false beliefs and false signs but not "false" photographs.

To conclude, children may need language and executive function to reason about false representations but their success in reasoning about false representations is not determined by language and executive functions. The findings presented in this thesis consistently suggest that representational understanding is probably the key variable in dealing with false representations for children.

Theory of mind: An umbrella concept.

Understanding of false belief has been considered as a crucial aspect of "theory of mind" which has been broadened to comprise processing of all types of inner, mental and emotional states. Hence, the argument that understanding of false belief is a general ability of representational understanding would be contradictory to the notion that "theory of mind" is a domain-specific mechanism which is limited to mental-related information processing (e.g., Leslie, 1987, 1994; Leslie et al., 2004). However, this thesis underlines that understanding of false belief involves not only mental but also representational characteristics so it should be regarded as a separate aspect of "theory of mind" as previously suggested by Leekam et al. (2008), Perner (1991), and Wellman (1990). 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. "Non-representational theory of mind" may thus include the following examples: gaze monitoring, joint attention, detection of intentional contingencies, and emotion processing, etc. Contrasting to "representational theory of mind", "non-representational theory of mind" might be domain-specific as they are all restricted to social stimuli and seem to be dependent on specific neural circuitry (e.g., Blakemore, et al., 2003; Campbell, Heywood, Cowey, Regard, & Landis, 1990; Stone, 2005).

Further, other researchers suggest that "theory of mind" goes beyond the understanding of false belief to what has been called an "interpretive theory of mind" (e.g., Carpendale & Chandler, 1996; Chandler & Lalonde, 1996; Lalonde & Chandler, 2002). In addition to the appreciation that different people will hold different beliefs if they have access to different information (e.g., a false belief situation in which A has access to the previous location of an object while B has access to its current location), "interpretive theory of mind" involves the recognition that different people can make different interpretations of exactly the same information. For example, children by the age of seven to eight years come to understand that different people can interpret an ambiguous figure differently (e.g., A would say that  is a duck while B would say that it is a rabbit; Carpendale & Chandler, 1996). In other words, children by this stage do not only view the mind as passively accommodating to the world, but also view it as actively interpreting the world and constructively assimilating experience to its own knowledge structures. This more complex appreciation of the mind was also found to be related to executive function (Lagattuta, Sayfan, & Blattman, 2010).

The very broad set of "theory of mind" abilities, as Stone and Gerrans (2006a, 2006b) and Gerrans and Stone (2008) has recently proposed, may thus depend on the interaction of domain-general mechanisms, such as metarepresentation, language and executive function, with domain-specific low-level cognitive mechanisms, such as gaze monitoring and joint attention. This argument against the domain-specific account of "theory of mind" seems to be in line with the findings and their implications discussed in this chapter.

ASD: What is its core cognitive impairment?

The differentiation between "representational theory of mind" and "non-representational theory of mind" not only has theoretical implication to the conception of "theory of mind" but also illuminates our understanding of atypical development. 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 mental 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 (Baron-Cohen, 1989c; Hobson, 1984; Leslie & Frith, 1988), 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). Castelli (2005) also indicated that individuals with ASD are able to recognize basic emotions (i.e., happiness, sadness, anger, fear, disgust, and surprise). Furthermore, there are several studies demonstrating that children with ASD are capable of understanding goals and intentions (see Hamilton, 2009, for a recent review), although they were also found to give explanations that were inappropriate to animated geometric shapes which moved interactively with implied intentions (Abell, Happé, & Frith, 2000; Castelli, Frith, Happé, & Frith, 2002). As a result, individuals with ASD may not be as impaired in "non-representational theory of mind" as they are in "representational theory of mind".

As mentioned earlier, language is also impaired in ASD (e.g., Charman et al., 2003; Lord & Paul, 1997), especially receptive language ability which is severely delayed (Ellis Weismer et al., 2010; Hudry et al., 2010). Could language delay be a central impairment which causes ASD's deficit in "representational theory of mind"? Some evidence from research with deaf children who are raised by hearing parents and mastered sign language belatedly at school (late signers) in comparison to children with ASD provide support for this possibility. Deaf preshoolers with hearing parents show delays and deficiencies on the FB task which are comparable to those found in children with ASD (e.g., Peterson & Siegal, 1995; Peterson, 2004). By contrast, deaf children who are raised by deaf parents and have acquired sign language early in life do not show such delays. These findings suggest that a delay in language development affects the development of understanding false beliefs in both deaf children of hearing parents and children with ASD. However, counterevidence comes from further research by Peterson, Wellman and Liu (2005), using a "theory of mind" scale devised by Wellman and Liu (2004). Peterson et al. illustrated that late-signing deaf children, children with ASD, and typically developing children all followed the same sequence of early "theory of mind" development, which includes the understanding of (1) diverse desires (e.g., two persons have different desires for the same object), (2) diverse beliefs (e.g., two persons have different beliefs about the same situation), and (3) perceptual access to knowledge. Late-signing deaf children and typically developing children continued to progress from understanding (4) false beliefs to (5) hidden emotions, whereas children with ASD deviated in these later steps of development. Children with ASD showed a reversed pattern of having greater difficulty in understanding false beliefs than hidden emotions. This finding suggests that late-signing deaf children simply delay in "theory of mind" development due to their delay in language development. However, children with ASD demonstrate a different pattern of development, showing a central difficulty in understanding false beliefs which could not be simply explained only by language delay.

Individuals with ASD are also found to have executive dysfunction which has even been suggested to be the explanation for the manifestation of autistic symptoms by some researchers (e.g., McEvoy, Rogers, & Pennington, 1993; Russell, 1997). Could executive dysfunction underline the ASD deficit in "representational theory of mind"? Pellicano (2007) examined the pattern of false-belief and executive function impairments in children with ASD. Children who are impaired in false-belief understanding but not in executive function were found. However, there was no child who showed impaired executive function but possessed intact false-belief understanding. Pellicano (2010b) further investigated the longitudinal relationship between false-belief understanding and executive function in ASD. Consistent with her earlier study, the asymmetric relationship between false-belief understanding and executive function (i.e., early-emerging executive skills were associated with the progress of false-belief understanding but not the reverse) remained over a longitudinal period of 3 years. Taken together, these findings suggest that intact executive function is necessary, but not sufficient, for the development of false-belief understanding in ASD.

Using a non-verbal reality-unknown FB task adapted from Call and Tomasello (1999), Colle et al. (2007) also indicated that children with ASD were significantly impaired in false-belief understanding relative to children with specific language impairment despite the fact that the demands of language and cognitive inhibition in the task were minimised and eliminated. Experiment 5 of this thesis presented in Chapter 4, extended this finding to false-sign understanding, suggesting a fundamental deficit in representational understanding in children with ASD which cannot be explained by their language and executive dysfunctions. However, how this metarepresentational deficit arises has not yet been established.

The Human Brain: Functional and Developmental Issues

Research on children's understanding of mental and non-mental representations is also informative to the investigation of the human brain, including its function and its development. For example, following the findings from behavioural studies that children's performance on mental representation task is equivalent to that on non-mental representation task, a question has been raised about whether the brain shows similar equivalence in processing mental and non-mental representations (e.g., Perner et al., 2006). Given that language and executive function are often confounded in standard mental and non-mental representation tasks, I formulated the question mentioned in Chapter 1: Is there a specific neural mechanism which selectively responds to mental rather than non-mental representations that is independent of language and executive demands? Further questions regarding development should also be investigated, e.g., does the brain change its way of processing mental and non-mental representations across the lifespan? Having outlined these questions, this section attempts to extend this thesis which is based on behavioural experiments by incorporating recent evidence from previous neuroscience studies.

As mentioned in Chapter 1, the temporo-parietal junction (TPJ) has been found to be crucial for understanding both mental and non-mental representations in adults (Aichhorn et al., 2009; Perner et al., 2006; Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005). More specifically, in studies with adults only, the right TPJ is more important for the understanding of false beliefs whereas the left TPJ is associated with both false beliefs and false signs (Aichhorn et al., 2009; Perner et al., 2006). This finding might be interpreted as showing a neural equivalence in processing mental and non-mental representations as well as a specific neural mechanism which selectively responds to mental representations. However, I would interpret this finding as in line with the transfer effect between the FB and FS tasks reported in Experiment 2 of this thesis.

Before expanding on this further, another intriguing finding by Dahlin, Neely, Larsson, Backman, and Nyberg (2008) has to be mentioned. Dahlin et al. indicated that training on one cognitive task, which recruits the same brain region as another cognitive task, will enhance activations in that region and benefit its recruitment in the latter task, aligning with a behavioural transfer effect after training. Given that the FB and FS tasks are associated with the same brain region namely the left TPJ, it is possible that training on the FB task increases activations in the left TPJ which benefit its recruitment in the FS task, resulting in the transfer effect from false-belief training to false-sign understanding demonstrated in Experiment 2 of this thesis. However, the FB task is not only associated with the left TPJ but also the right TPJ. Training on the FS task would probably increase activations in the left TPJ only which may not provide enough help for processing false-belief information, especially in processing information about *other people's* false belief. Reasoning about others' cognitive mental states and distinguishing oneself from others has consistently been shown to result in activating the right TPJ (Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005; see Decety & Lamm, 2007, for a meta-analysis). This speculation backed up by related evidence offers a reason why Experiment 2 of this thesis found that false-sign training contributed to the understanding of one's own false beliefs (i.e., the 'self' question of the FB task) but not other's false beliefs. Another reason may come from the findings of Burianova, McIntosh, and Grady's (2010) study, which suggested that the left TPJ is also involved in declarative memory retrieval. As the 'self' question of the FB task does not only require the understanding of one's representational change but also places demands on the retrieval of one's previous representation, the increased activation of the left TPJ resulting from false-sign training may benefit its recruitment in answering the 'self' question rather than the 'other' question of the FB task.

Although the above is mainly a speculation based on the findings of recent brain imaging studies with adults only, there is one issue which is more certain. That is, both right and left TPJ are activated not only when processing false beliefs and false signs, but also during other cognitive processes. The left TPJ is also engaged in declarative memory retrieval, whereas the right TPJ is also involved in attentional reorientation (e.g., Mitchell, 2008; Serences et al., 2005). Thus, there might not be a specific neural mechanism which selectively responds to mental or non-mental representations that is independent of other cognitive processes. Although one could isolate the cognitive demands involved in the FB and FS tasks in a way described in Chapter 3, one might not be able to find a brain mechanism that is specifically activated for understanding false representations but not other cognitive processes.

Other brain regions that are activated in "theory of mind" tasks, most of which involve false-belief scenario, are not specifically social either. For example, the posterior superior temporal sulcus (for meta-analyses, see Frith & Frith, 1999, 2003; for a recent review, see Carrington & Bailey, 2009). On one hand, the posterior superior temporal sulcus has been suggested to be concerned with biological motion (e.g., Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005; Saxe et al., 2009). On the other hand, it is also illustrated to participate in predicting complex movement trajectories of any kind, e.g., two balls that moved in correlated trajectories which were determined by a mathematical algorithm (Schultz, Friston, O'Doherty, Wolpert,

& Frith, 2005; Schultz, Imamizu, Kawato, & Frith, 2004; see also Kawawaki, Shibata, Goda, Doya, & Kawato, 2006).

Thus far, each brain region that has been mentioned is shown to be involved in a wide range of cognitive processes. Brain regions are also suggested to influence and be influenced by each other as well as its external environment across development (Friston & Price, 2001; Mareschal et al., 2007). The medial prefrontal cortex (mPFC) is an example. When engaging in tasks requiring mentalising, the mPFC is activated more extensively in children than adults, who however show greater activity in posterior temporal cortical areas (e.g., Blakemore, den Ouden, Choudhury, & Frith, 2007; Kobayashi et al., 2007; Wang, Lee, Sigman, & Dapretto, 2006). Another study using event-related potentials (ERPs) also indicated that a left frontal negative slow wave which presumably reflects the prefrontal cortex's activity was observed in adults and children who succeeded in belief reasoning, but these children's activity had a more diffuse frontal scalp distribution than adults (Liu, Sabbagh, Gehring, & Wellman, 2009). Liu, Meltzoff, and Wellman (2009) further illustrated that a mid-frontal late slow wave and a right-posterior late slow wave were associated with belief judgments in adults. According to the interactive specialisation view (see Johnson, 2005), the mPFC may start off by being extensively activated in a wide range of different contexts. However, the mPFC may change its response properties as it interacts with other brain regions in this wide range of different contexts during development. Johnson, Grossmann, and Cohen Kadosh (2009) further suggested that the mPFC may play a role in orchestrating the combinations of other posterior brain regions activated for dealing with a given mentalising task. Once the mPFC has learned the appropriate pattern of activity in the posterior regions for succeeding in that task, the activity in the mPFC becomes less extensive while greater activity can be found in the posterior regions.

A developmental change in response properties of a posterior brain region was also demonstrated in one recent brain imaging study. Saxe et al. (2009) scanned children aged 6 to 11 years and found that the mPFC, TPJ, and posterior cingulate were recruited while they listened to sections of a story describing a protagonist's thoughts. However, only the right TPJ showed a developmental change in that it was recruited equally for both mental and physical facts about people in younger children, but only for mental facts in older children. Although the response profile in the mPFC did not change over this period of age, previously mentioned studies suggest the diffuse to focal shift in the activity of the mPFC occurs later in development which extends to adulthood.

It is also important to note that studies using non-mental tasks also demonstrated similar developmental patterns of (1) shifting from diffuse to focal activity in the prefrontal cortex (e.g., Durston et al., 2006), and (2) changing response properties of posterior cortical regions (e.g., Brown et al., 2005). As a result, any functional selectivity of a certain brain region observed in adults may not be innate but gradually become sophisticated in a narrower range of cognitive processes through experience during development (e.g., the right TPJ; Saxe et al., 2009). Moreover, it is more likely that several brain regions interact to support a particular set of cognitive functions.

These suggestions are also consistent with the evidence from the neuroscience studies in neurodevelopmental disorders such as ASD. Instead of having a deficit localised in a specific brain region, abnormalities are indicated in several cortical and subcortical regions in ASD (e.g., Rumsey & Ernst, 2000; McAlonan et al. 2005). Importantly, an abnormal brain overgrowth trajectory was found as early as the first two years of life (e.g., Courchesne et al., 2001; Schumann et al., 2010; Webb et al., 2007) which precedes the age of 2 years at which symptoms usually become apparent. In addition to grey matter (i.e., nerve cell bodies in the brain), white matter (i.e., connections between different locations of nerve cell bodies in the brain) development is also anomalous (e.g., Herbert et al., 2004; Ben Bashat et al., 2007; Waiter et al., 2005). Moreover, there is a growing body of evidence demonstrating abnormalities in long-range, functional connectivity in ASD (e.g., Castelli et al., 2002; Cherkassky, Kana, Keller, & Just, 2006; Just, Cherkassky, Keller, Kana, & Minshew, 2007; Kana, Keller, Cherkassky, Minshew & Just, 2006; Kennedy & Courchesne, 2008; Kleinhans et al., 2008; Koshino et al., 2005, 2008). A more recent brain imaging study further showed that individuals with ASD continue to show developmental differences in interhemispheric connectivity into early adulthood (Anderson et al., 2010). Thus, any initial brain abnormalities that affect the processing of external stimuli early in life may further affect broader patterns of connectivity between brain regions during development. This may offer a neural framework explaining how the metarepresentational deficit found in ASD arises.

It has been suggested that an early abnormality in the brain that leads to diminished attention to social stimuli, resulting in decreased relevant input and experience, has a cascading influence on the development of other cortical areas in

ASD (e.g., Dawson et al., 2005; Johnson et al., 2005; Schultz, 2005). This suggestion was supported by recent studies which examined infant siblings of children diagnosed with ASD and found behavioural as well as neurological differences on eye-gaze, face, and object processing relative to infants with no family history of ASD (e.g., Elsabbagh et al., 2009; McCleery, Akshoomoff, Dobkins, & Carver, 2009; Merin, Young, Ozonoff, & Rogers, 2007; Presmanes, Walden, Stone, & Yoder, 2007). Given that social stimuli are often relevant to all sorts of social and cognitive learning, young infants who later received a diagnosis of ASD might have missed out various amount of input that is important for normal development of brain circuitry which subserves representational understanding. However, it is important to note that brain connectivity may change in relation to its interaction with external contexts across development. Therefore, any early training or intervention that may improve the affected processing of external stimuli may allow a reallocation of connectivity between brain regions.

In sum, there might not be a specific neural mechanism which selectively responds to mental or non-mental representations that is independent of other cognitive processes. Brain regions act together for a set of cognitive functions and may change functionally in connections with each other and in relation to external contexts during development. Children may initially rely on several brain regions to a greater extent than adults for a given cognitive task. However, the activation and function of these brain regions could become progressively specific probably through the interactions between these regions and external contexts in a developmental process. If any brain region, connectivity or aspect of external contexts alters during development, so may the others and their associated functions change subsequently.

Suggestions for Future Research

In this final section, I make suggestions for further research based on the broad discussion on the implications of the empirical work reported. This thesis addresses the importance and the scope of representational understanding, which include other essential cognitive processes such as mentalising, language and executive function, in the context of typical and atypical development. Further investigation is required to clarify the acquisition and development of representational understanding. Although Chapter 2 presented a training study which suggests that experience is the key of the acquisition and development of metarepresentation, the kind of experience gained was not specified (e.g., repeated exposure of representations, conflicting information, verbal instructions, or social interaction). Microgenetic studies are thus essential to investigate the change over the course of transition in representational understanding, especially on how the change occurs and what mechanisms or specific kinds of experience underpin the change. Although there are microgenetic studies that have been done on false-belief understanding (Flynn et al., 2004; Amsterlaw, & Wellman, 2006), neither of them has attempted to identify the specific kinds of experience that children have received in the study underpin the change of understanding. Moreover, according to the constructive viewpoint suggested by Carpendale and Lewis (2004), experience of social interaction has a significant influence on children's development of understanding minds. The link between social interaction experience and the development of metarepresentation is therefore an important issue to be investigated. It would have implications not only in cognitive development but also in functional brain development given that the brain changes in relation to its interaction with its external environment.

The interactive specialisation view of functional brain development also suggested that each brain region may change its response properties as it interacts with each other (see Johnson, 2005). Hence, the brain network that underpins representational understanding may become more sophisticated in processing representations or may even exhibit a less extensive activation for a specific type of representations as a child develops into an adult. However, this is just a speculation based on a specific view of functional brain development. Extensive research would have to be devoted to it before it could be established. Especially, whether it is the *interactions* between brain regions that affect the developmental changes in functional brain activity requires investigation.

Regarding atypical development of representational understanding, this thesis provides an example in examining children with ASD with a cross-sectional methodological approach which lacked a contribution in addressing the developmental change occurs in this particular type of atypical development. In particular, the question of how the metarepresentational deficit in ASD arises on behavioural, cognitive and neurological levels is essential to the conception and intervention of the disorder. I have suggested a neural framework which parallels to other models proposed by other researchers (e.g., Dawson et al., 2005) to explain the emergence of the social impairments and metarepresentational deficit shown in ASD.

Again, it is a suggestion which requires further investigation. As a result, neuroscience and brain imaging opens a wide window into the functional brain development that is involved in understanding representations in both typical and atypical development.

In addition to individuals with ASD, exploring representational understanding in other populations with atypical development would offer an advanced understanding of that particular type of atypical development and extend the current research on representational understanding. Given that atypical development, as mentioned in Chapter 1, may factor apart the cognitive components which usually entangle with representational understanding, the influence of mentalising, language and executive function development on the acquisition and development of representational understanding can therefore examined more clearly. For example, having established a new non-verbal paradigm in Chapter 3 and 4, further investigation can be carried out with atypically developing children, such as deaf children with hearing parents to specifically explore the effect of having a delay in language development on the acquisition and development of representational understanding. As deaf children of hearing parents have been found to show better performance on the non-verbal FB task than the standard one (Figueras-Costa & Harris, 2001), it is expected that the same would be found on the non-verbal FS task and it would be significantly correlated with the non-verbal FB task. If this is the case, it would provide support for the findings and their implications presented in Chapter 3 and 4.

Conclusion

This thesis addresses the long-standing dispute between domain-specificity and domain-generality in understanding mental representations by critically investigating whether mental and non-mental representations are both underlined by the same competence of representational understanding. Novel tasks for assessing children's understanding of non-mental representations were developed in comparison to mental representation tasks. In a series of experiments, children's understanding of mental representations was consistently demonstrated to be equivalent to their understanding of non-mental representations. The equivalence exists in a stringent methodology, testing transferability after training, as well as a correlational methodology, with and without the involvement of language and cognitive inhibition. Moreover, it was also found in a group of children with ASD. These findings provide supports to the generality claim that understanding of mental representation is an ability of representational understanding which includes, but is not limited to, mental aspect.

I suggest that children's understanding of representations develops through experience during a process of conceptual consolidation and restructuring. On the contrary, children with ASD are impaired in representational understanding probably due to their early abnormality in the brain that affects social orienting. Both children with typical development and ASD may need language and executive function to reason about false representations but their success in reasoning about false representations is not determined by language and executive functions. In other words, language and executive function are necessary but not sufficient for understanding false representations in both typically developing children and children with ASD. Although the findings reported in this thesis favour the generality claim of understanding mental representation, there is a possibility that children after the age of 4 to 5 years may rely on a general understanding of representation in dealing with mental representations but then this general understanding may become progressively specific through further experience and brain development. Future research focusing on the development of understanding representations and functional brain changes are needed. Detailed investigations of the nature and scope of functional brain development will help develop cognitive and neural models of representational understanding and its relation to other social and cognitive processes.

References

Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*, 1-16.

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. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Amsterdam, B. K. (1972). Mirror self-image reactions before age two. *Developmental Psychology, 5*, 297-305.

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development, 7*, 139-172.

Anderson, J. S., Druzgal, T. J., Froehlich, A., DuBray, M. B., Lange, N., Alexander, A. L., ... Lainhart, J. E. (2010). Decreased interhemispheric functional connectivity in autism. *Cerebral Cortex*. Advance online publication. Retrieved from http://cercor.oxfordjournals.org/content/early/2010/10/12/cercor.bhq190.full

Apperly, I., Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact 1st and 2nd order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience, 1*, 334-348.

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.

Apperly, I., Samson, D., & Humphreys, G. (2009). Studies of adults can inform account of theory of mind development. *Developmental Psychology, 45*, 190-201.

Appleton, M., & Reddy, V. (1996). Teaching 3-year olds to pass false belief tests: A conversational approach. *Social Development, 5*, 275-291.

Astington, J. W. (1993). *The child's discovery of the mind.* Cambridge, MA: Harvard University Press.

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

Baron-Cohen, S. (1989a). Are autistic children behaviourists? An examination of their mental-physical and appearance-reality distinctions. *Journal of Autism and Developmental Disorder, 19*, 579-600.

Baron-Cohen, S. (1989b). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30*, 285-297.

Baron-Cohen, S. (1989c). 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.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed.). Oxford, England: Oxford University Press.

Barrera, M. E., & Maurer, D. (1981). Recognition of mother's photographed face by the three-month-old infant. *Child Development, 52*, 714-716.

Baumgaertner, A., Buccino, G., Lange, R., McNamara, A., & Binkofski, F. (2007). Polymodal conceptual processing of human biological actions in the left inferior frontal lobe. *European Journal of Neuroscience, 25*, 881-889.

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.

Ben Bashat, D., Kronfeld-Duenias, V., Zachor, D. A., Ekstein, P. M., Hendler, T., Tarrasch, R., ... Ben Sira, L. (2007). Accelerated maturation of white matter in young children with autism: A high b value DWI study. *NeuroImage, 37*, 40-47.

Bennett, T., Szatmari, P., Bryson, S., Volden, J., Zwaigenbaum, L., Vaccarella, L., ... Boyle, M. (2008). Differentiating autism and Asperger syndrome on the basis of language delay or impairment. *Journal of Autism and Developmental Disorders, 38*, 616-625.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*, 444-451.

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

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*, 382-386.

Blakemore, S. J., Boyer, P., Pachot-Clouchard, M., Meltzoff, A., Segebarth, C., & Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cerebral Cortex, 13*, 837-844.

Blakemore, S. J., den Ouden, H., Choudhury, S., & Frith, C. D. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive & Affective Neuroscience, 2*, 130-139.

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.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development, 67*, 836-884.

Brown, T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cerebral Cortex, 15*, 275-290.

Burianova, H., McIntosh, A. R., & Grady, C. L. (2010). A common functional brain network for autobiographical, episodic, and semantic memory retrieval. *NeuroImage, 49*, 865-874.

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

Campbell, R., Heywood, C., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia, 28*, 1123-1142.

Cardinal, R. N., & Aitken, M. R. F. (2006). *ANOVA for the Behavioural Sciences Researcher.* New Jersey, USA: Lawrence Erlbaum Associates.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032-1053.

Carpendale, J. I. M., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*, 1686-1706.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27*, 79-151.

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping, 30*, 2313-2335.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism, 9*, 428-449.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*, 1839-1849.

Chandler, M. J., & Lalonde, C. (1996). Shifting to an interpretive theory of mind: 5- to 7-year-olds' changing conceptions of mental life. In A. Sameroff & M. Haith (Eds.), *Reason and responsibility: The passage through childhood.* Chicago: University of Chicago Press.

Charman, T. (2004). Matching preschool children with autism spectrum disorders and comparison children for language ability: methodological challenges. *Journal of Autism and Developmental Disorders, 34*, 59-64.

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.

Cherkassky, V. L., Kana, R. K., Keller, T. A., & Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport, 17*, 1687-1690.

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science, 3*, 81-92.

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.

Cohen, L. B. (2004). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development, 13*, 349-352.

Colle, L., Baron-Cohen, S., Hill, J. (2007). Do children with autism have a theory of mind? A non-verbal test of autism vs. specific language impairment. *Journal of Autism and Developmental Disorders, 37*, 716-723.

Courchesne, E., Karns, C., Davis, H. R., Ziccardi, R., Carper, R., Tigue, Z., … Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology, 57*, 245-254.

Dahlin, E., Neely, A. S., Larsson, A., Backman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320*, 1510-1512.

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25-31.

Dawson, G., Webb, S. J., Wijsman, E., Schellenberg, G., Estes, A., Munson, J., & Faja, S. (2005). Neurocognitive and electrophysiological evidence of altered face processing in parents of children with autism: Implications for a model of abnormal development of social brain circuitry in autism. *Development and Psychopathology, 17*, 679-697.

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist, 13*, 580-593.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62*, 736-752.

DeLoache, J. S., & Burns, N. M. (1994). Early understanding of the representational function of pictures. *Cognition, 52*, 83-110.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science, 9*, 205-210.

DeLoache, J. S., Simcock, G., & Marzolf, D. P. (2004). Transfer by very young children in the symbolic retrieval task. *Child Development, 75*, 1708-1718.

Denis, M. (1994). *Image and cognition* (2nd ed.). Paris: Presses Universitaires de France.

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

Dirks, J. R., & Gibson, E. (1977). Infants' perception of similarity between live people and their photographs. *Child Development, 48*, 124-130.

Doherty, M. J., & Wimmer, M. C. (2005). Children's understanding of ambiguous figures: Which cognitive developments are necessary to experience reversal? *Cognitive Development, 20*, 407-421.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature, 384*, 159-161.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test-revised: Form M*. Circle Pines, MN: American Guidance Service.

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

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fosella, J. A., Casey, B. J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science, 9*, 1-20.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., & Ong, B. (1996). Comparison of clinical symptoms in autism and Asperger disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 1523-1531.

Ellis Weismer, S., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders, 40*, 1259-1273.

Elsabbagh, M., Volein, A., Csibra, G., Holmboe, K., Garwood, M., Tucker, L., ... Johnson, M. H. (2009). Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biological Psychiatry, 65*, 31-38.

Fadiga, L., Craighero, L., & Roy, A. C. (2006). Broca's region: A speech area? In Y. Grodzinsky & K. Amunts (Eds.), *Broca's region* (pp. 137-152). New York, NY: Oxford University Press.

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., ... Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain, 132*, 1980-1988.

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.

Fisher, N., Happé, F., & Dunn, J. (2005). The relationship between vocabulary, grammar, and false belief task performance in children with autistic spectrum disorders and children with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 46*, 401-419.

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

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology, 15*, 95-120.

Flavell, J. H., Green, F., & Flavell, E. R. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development, 51* (1, Serial No. 212).

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7*, 103-115.

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283-296.

Friston, K. J., & Price, C. J. (2001). Dynamic representations and generative models of brain function. *Brain Research Bulletin, 54*, 275-285.

Frith, C. D., & Frith, U. (1999). Interacting minds – A biological basis. *Science, 286*, 1692-1695.

Frith, U. (1991). Translation and annotation of "Autistic psychopathy" in childhood, by H. Asperger. In U. Frith (Ed.), *Autism and Asperger Syndrome* (pp. 37-92). Cambridge: Cambridge University Press.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, Series B, 358*, 685-694.

Ganea, P. A., Bloom-Pickard, M., & DeLoache, J. S. (2008). Transfer between picture books and the real world by very young children. *Journal of Cognition and Development, 9*, 46-66.

Ganea, P. A., Preissler, M. A., Butler, L., Carey, S., & DeLoache, J. S. (2009). Toddlers' referential understanding of pictures. *Journal of Experimental Child Psychology, 104*, 283-295.

Gerrans, P., & Stone, V. E. (2008). Generous or parsimonious cognitive architecture? Cognitive neuroscience and theory of mind. *The British Journal for the Philosophy of Science, 59*, 121-141.

Goodman, N. (1976). *Languages of art*. Indianapolis, IN: Hackett Publishing Co.

Gopnik, A., & Rosati, A. (2001). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science, 4*, 175-183.

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62*, 98-110.

Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70-83.

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.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-112.

Haith, M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behaviour & Development, 21*, 167-179.

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.

Hamilton, A. F. C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition, 113*, 37-44.

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.

Harris, P. L., Kavanaugh, R. D., & Dowson, L. (1997). The depiction of imaginary transformations: Early comprehension of a symbolic function. *Cognitive Development, 12*, 1-19.

Herbert, M. R., Ziegler, D. A., Makris, N., Filipek, P. A., Kemper, T. L., Normandin, J. J., ... Caviness Jr, V. S. (2004). Localization of white matter volume increase in autism and developmental language disorder. *Annals of Neurology, 55*, 530-540.

Hill, E. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8*, 26-32.

Hobson, R. P. (1984). Early childhood autism and the question of egocentrism. *Journal of Autism & Developmental Disorders, 14*, 85-104.

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 57*, 567-582.

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., ... the PACT Consortium (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders, 45*, 681-690.

Johnson, M. H. (2005). *Developmental cognitive neuroscience: An introduction* (2nd ed.). Oxford: Blackwell.

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., ... Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology, 17*, 599-619.

Johnson, M. H., Grossmann, T., & Cohen Kadosh, K. (2009). Mapping functional brain development: Building a social brain through interactive specialization. *Developmental Psychology, 45*, 151-159.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex, 17*, 951-961.

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2006). Sentence comprehension in autism: Thinking in pictures with decreased functional connectivity. *Brain, 129*, 2484-2493.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217-250.

Kawawaki, D., Shibata, T., Goda, N., Doya, K., & Kawato, M. (2006). Anterior and superior lateral occipito-temporal cortex responsible for target motion prediction during overt and covert visual pursuit. *Neuroscience Research, 54*, 112-123.

Kennedy, D. P., & Courchesne, E. (2008). The intrinsic functional organization of the brain is altered in autism. *NeuroImage, 39*, 1877-1885.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*, 25-41.

Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16*, 287-308.

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research, 78*, 85-123.

Klauer, K. J., Willmes, K., & Phye, G. D. (2002). Inducing inductive reasoning: Does it transfer to fluid intelligence? *Contemporary Educational Psychology, 27*, 1-25.

Kleinhans, N. M., Richards, T., Sterling, L., Stegbauer, K. C., Mahurin, R., Johnson, L. C., ... Aylward, E. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain, 131*, 1000-1012.

Knoll, M., & Charman, T. (2000). Teaching false belief and visual perspective taking skills in young children: Can a theory of mind be trained? *Child Study Journal, 30*, 273-304.

Kobayashi, C., Glover, G. H., & Temple, E. (2007). Children's and adults' neural bases of verbal and nonverbal "theory of mind." *Neuropsychologia, 45*, 1522-1532.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490-507.

Koshino, H., Carpenter, P. A., Minshew, N. J., Cherkassky, V. L., Keller, T. A., & Just, M. A. (2005). Functional connectivity in an fMRI working memory task in high-functioning autism. *NeuroImage, 24*, 810-821.

Koshino, H., Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2008). fMRI investigation of working memory for faces in autism: Visual coding and underconnectivity with frontal areas. *Cerebral Cortex, 18*, 289-300.

Lagattuta, K. H., Sayfan, L., & Blattman, A. J. (2010). Forgetting common ground: Six- to seven-year-olds have an overinterpretive theory of mind. *Developmental Psychology, 46*, 1417-1432.

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology, 20*, 163-198.

Leekam, S., Libby, S., Wing, L., Gould, J., & Gillberg, C. (2000). Comparison of ICD-10 and Gillberg's criteria for Asperger syndrome. *Autism, 4*, 11-28.

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. (1987). Pretence and representation: The origins of "theory of mind". *Psychological Review, 94*, 412-426.

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50*, 211-238.

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., & Frith, U. (1988). Autistic children's understanding of seeing, knowing, and believing. *British Journal of Developmental Psychology, 6*, 315-324.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*, 225-251.

Levene, H. (1960). Robust Test for Equality of Variances. In I. Olkin (Ed.), *Contributions to Probability and Statistics* (pp. 278-292). California, United States: Stanford University Press.

Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. M. (1996). Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child Development, 67*, 2930-2947.

Lewis, C. N., Freeman, N. H., & Smith, C. (1992). *Dissociation of inferences about beliefs and pictures in preschoolers.* Unpublished manuscript, University of Lancaster, United Kingdom.

Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self.* New York: Plenum.

Liu, D., Meltzoff, A. N., & Wellman, H. M. (2009). Neural correlates of belief- and desire-reasoning. *Child Development, 80*, 1163-1171.

Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2009). Neural correlates of children's theory of mind. *Child Development, 80*, 318-326.

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development, 74*, 1130-1144.

Lord, C., & Paul, R. (1997). Language and communication in autism. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed.). New York: John Wiley & Sons.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders, 24*, 659-685.

Luna, B., Doll, S. K., Hegedus, S. J., Minshew, N. J., & Sweeney, J. A. (2007). Maturation of executive function in autism. *Biological Psychiatry, 61*, 474-481.

Mareschal, D., Johnson, M. H., Sirios, S., Spratling, M., Thomas, M., & Westermann, G. (2007). *Neuroconstructivism: Vol. I. How the brain constructs cognition.* Oxford: Oxford University Press.

Marzolf, D. P., & DeLoache, J. S. (1994). Transfer in young children's understanding of spatial representations. *Child Development, 65*, 1-15.

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

Mauchly, J. W. (1940). Significance test for sphericity of a normal *n*-variate distribution. *Annuals of Mathematical Statistics, 11*, 204-209.

Mayes, S. D., & Calhoun, S. L. (2003). Relationship between Asperger syndrome and high-functioning autism. In M. Prior (Ed.), *Learning and behaviour problems in Asperger syndrome*. New York: Guildford Press.

McAlonan, G. M., Cheung, V., Cheung, C., Suckling, J., Lam, G. Y., Tai, K. S., … Chua, S. E. (2005). Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain, 128*, 268-276.

McCleery, J. P., Akshoomoff, N., Dobkins, K. R., & Carver, L. J. (2009). Atypical face versus object processing and hemispheric asymmetries in 10-month old infants at risk for autism. *Biological Psychiatry, 66*, 950-957.

McEvoy, R., Rogers, S., & Pennington, B. (1993). Executive function and social communication deficits in young autistic children. *Journal of Child Psychology and Psychiatry, 34*, 563-578.

Melot, A.-M., & Angeard, N. (2003). Theory of mind: Is training contagious? *Developmental Science, 6*, 178-184.

Merin, N., Young, G. S., Ozonoff, S., & Rogers, S. J. (2007). Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-monthold infants at-risk for autism from comparison infants. *Journal of Autism and Developmental Disorders, 37*, 108-121.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*, 622-646.

Milne, A. A. (1926). *Winnie-the-Pooh*. London: Methuen & Co. Ltd.

Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory of mind. *Cerebral Cortex, 18*, 262-271.

Mitchell, P., & Lacohee, H. (1991). Children's early understanding of false belief. *Cognition, 39*, 107-127.

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Developmental and Behavioral Pediatrics, 27*, S69-S78.

Nielsen, M., & Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: A longitudinal investigation through the second year. *Infant Behavior & Development, 27*, 342-365.

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

Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology, 6*, 415-431.

Ozonoff, S., Pennington, B., & Rogers, S. (1991). Executive function deficits in high-functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*, 1081-1106.

Parkin, L. J. (1994). *Children's understanding of misrepresentation* (Unpublished doctoral thesis). University of Sussex, United Kingdom.

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. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*, 974-990.

Pellicano, E. (2010a). The development of core cognitive skills in autism: A 3-year prospective study. *Child Development, 81*, 1400-1416.

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

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex, 15*, 1866-1876.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

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

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

Perner, J., 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., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development, 60*, 689-700.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child Development, 73*, 752-767.

Perner, J., & Leekam, S. (2008). The curious incident of the photo that was accused of being false: Issues of domain specificity in development autism and brain imaging. *Quarterly Journal of Experimental Psychology, 61*, 70-89.

Perner, J., Leekam, S. R., Myers, D., Davis, S., & Odgers, N. (1993). *Misrepresentation and referential confusion: children's difficulty with false beliefs and outdated photographs*. Unpublished manuscript, Laboratory of Experimental Psychology, University of Sussex, United Kingdom.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125-137.

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

Peterson, C. C. (2004). Theory-of-mind development in oral deaf children with cochlear implants or conventional hearing aids. *Journal of Child Psychology and Psychiatry, 45*, 1096-1106.

Peterson, C. C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry, 36*, 459-474.

Peterson, C. C., & Siegal, M. (1998). Changing focus on the representational mind: Deaf, autistic and normal children's concepts of false photos, false drawings and false beliefs. *British Journal of Developmental Psychology, 16*, 301-320.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*, 502-517.

Piaget, J. (1951). *Play, dreams and imitation in childhood.* New York: Norton.

Pierroutsakos, S. L., & DeLoache, J. S. (2003). Infants' manual exploration of pictured objects varying in realism. *Infancy, 4*, 141-156.

Pierroutsakos, S. L., & Troseth, G. L. (2003). Video verite: infants' manual investigation of objects on video. *Infant Behaviour and Development, 177*, 1-17.

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.

Preissler, M. A. (2008). Associative learning of pictures and words by low-functioning children with autism. *Autism, 12*, 229-246.

Preissler, M. A., & Carey, S. (2004). Do both pictures and words function as symbols for 18- and 24-month-old children? *Journal of Cognition and Development, 5*, 185-212.

Presmanes, A. G., Walden, T. A., Stone, W. L., & Yoder, P. J. (2007). Effects of different attentional cues on responding to joint attention in younger siblings of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 133-144.

Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? *The Behavioral and Brain Sciences, 1*, 592-593.

Reed, M., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly, 30*, 131-147.

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

Rock, I., Gopnik, A., & Hall, S. (1994). Do young children reverse ambiguous figures? *Perception, 23*, 635-644.

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

Ropar, D., Mitchell, P., & Ackroyd, K. (2003). Do children with autism find it difficult to offer alternative interpretations to ambiguous figures? *British Journal of Developmental Psychology, 21*, 387-395.

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.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*, 161-174.

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

Rumsey, J. M., & Ernst, M. (2000). Functional neuroimaging of autistic disorders. *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 171-179.

Russell, J. (1996). *Agency: Its role in mental development.* Hove, UK: Erlbaum.

Russell, J. (Ed.). (1997). *Autism as an executive disorder.* Oxford: Oxford University Press.

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.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*, 74-81.

Salsa, A. M., & de Mendoza, O. P. (2007). Routes to symbolization: Intentionality and correspondence in early understanding of pictures. *Journal of Cognition and Development, 8*, 79-92.

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

Samson, D., Apperly, I., & Humphreys, G. (2007). Error analysis in brain-damaged patients with perspective taking deficits: A window to the social mind and brain. *Neuropsychologia, 45*, 2561-2569.

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

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

Saxe, R., & Powell, J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science, 17*, 692-699.

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

Saxe, R., Whitfield-Gabrieli, S., Scholz, J., & Pelphrey, K. A. (2009). Brain regions for perceiving and reasoning about other people in school-aged children. *Child Development, 80*, 1197-1209.

Schäffler, L., Lueders, H. O., & Beck, G. J. (1996). Quantitative comparison of language deficits produced by extraoperative electrical stimulation of Broca's, Wernicke's, and Basal temporal language areas. *Epilepsia, 37*, 463-475.

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron, 45*, 625-635.

Schultz, J., Imamizu, H., Kawato, M., & Frith, C. D. (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of Cognitive Neuroscience, 16*, 1695-1705.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*, 125-141.

Schumann, C. M., Bloss, C. S., Barnes, C. C., Wideman, G. M., Carper, R. A., Akshoomoff, N., … Courchesne, E. (2010). Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *Journal of Neuroscience, 30*, 4419-4427.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about identity at 18 months. *Child Development, 80*, 1172-1196.

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.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science, 16*, 114-122.

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

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development, 67*, 2967-2988.

Smith, V., Mirenda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research, 50*, 149-160.

Sobel, D. M., Capps, L. M., & Gopnik, A. (2005). Ambiguous figure perception and theory of mind understanding in children with autistic spectrum disorders. *British Journal of Developmental Psychology, 23*, 159-174.

Sommer, M., Dohnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *NeuroImage, 35*, 1378-1384.

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

Stone, V. E. (2005). Theory of mind and the evolution of social intelligence. In J. Cacciopo (Ed.), *Social Neuroscience: People Thinking About People.* Cambridge, MA: MIT Press.

Stone, V. E., & Gerrans, P. (2006a). Does the normal brain have a theory of mind? *Trends in Cognitive Science, 10*, 3-4.

Stone, V. E., & Gerrans, P. (2006b). What's domain-specific about theory of mind? *Social Neuroscience, 1*, 309-319.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs to 13-month-old infants. *Psychological Science, 18*, 580-586.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed., pp. 124-149). Oxford: Oxford University Press.

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science, 16*, 311-315.

Waiter, G. D., Williams, J. H., Murray, A. D., Gilchrist, A., Perrett, D. I., & Whiten, A. (2005). Structural white matter deficits in high-functioning individuals with autistic spectrum disorder: A voxel-based investigation. *NeuroImage, 24*, 455-461.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social Cognitive & Affective Neuroscience, 1*, 107-121.

Webb, S. J., Nalty, T., Munson, J., Brock, C., Abbott, R., & Dawson, G. (2007). Rate of head circumference growth as a function of autism diagnosis and history of autistic regression. *Journal of Child Neurology, 22*, 1182-1190.

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

Wellman, H. M. (2010). Developing a theory of mind. In U. Goswami (Ed.), *The Blackwell handbook of cognitive development* (2nd ed.). Oxford, UK: Blackwell.

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.

Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development, 75*, 523-541.

Wilkinson, K. (1998). Profiles of language and communication skills in autism. *Mental Retardation and Developmental Disabilities Research Reviews, 4*, 73-79.

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.

Wimmer, M. C., & Doherty, M. J. (2010). Children with autism's perception and understanding of ambiguous figures: Evidence for pictorial metarepresentation, a research note. *British Journal of Developmental Psychology, 28*, 627-641.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine, 11*, 115-130.

Wing, L. (1988). The autistic continuum. In L. Wing (Ed.), *Aspects of autism: biological research*. London: Gaskell/Royal College of Psychiatrists.

Wing, L. (1996). *The autistic spectrum: A guide for parents and professionals.* London: Constable.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*, 11-29.

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.

Zelazo, P. D., Jacques, C., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant and Child Development, 11*, 171-195.