

2007

# COUNTERFACTUAL THINKING AND THE FALSE BELIEF TASK: A DEVELOPMENTAL STUDY

ARRECKX, FREDERIQUE

<http://hdl.handle.net/10026.1/1758>

---

<http://dx.doi.org/10.24382/4506>

University of Plymouth

---

*All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.*

**COUNTERFACTUAL THINKING  
AND  
THE FALSE BELIEF TASK:  
A DEVELOPMENTAL STUDY**

by

FREDERIQUE ARRECKX

A thesis submitted to the University of Plymouth in partial fulfilment  
for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology  
Faculty of Science

November 2007

University of Plymouth Library
9007951518
THESIS 155.4134

3 ARR .

# **ABSTRACT**

Name: Frederique ARRECKX

Title: Counterfactual thinking and the false belief task: a developmental study

The main aim of this thesis was to investigate children's reasoning abilities, especially their ability to solve counterfactual tasks. This thesis studied counterfactual thinking as an independent ability but also in relation to both types of the false belief tasks that measure theory of mind and contrary-to-fact syllogisms. Four experiments were conducted with typically developing children between the ages of 3 to 5. The first experiment studied the difference between affirmative and negative counterfactual questions and found that the former was easier than the latter. The second experiment replicated these results and investigated the link between counterfactual thinking and the false belief task but failed to find a link between the two. The third experiment found a link between counterfactual and both types of false belief tasks; false belief to others and false belief to self task. This experiment also identified a discrepancy between counterfactual thinking and contrary-to-fact syllogisms. When examining contrary-to-fact syllogisms it was found that the ones requiring a yes answer were easier than the ones requiring a no answer. This discrepancy between counterfactual thinking and contrary-to-fact syllogisms was confirmed in experiment 4 as a fantasy context influenced performance on these tasks differently. Fantasy context increased performance on contrary-to-fact syllogisms but decreased performance on counterfactual tasks.

In conclusion, this thesis confirmed a link between counterfactual thinking and the false belief task. It identified an important difference between counterfactual thinking and contrary-to-fact reasoning. It also highlighted the importance of the polarity (affirmative versus negative) of questions as well as answers on performance.

# LIST OF CONTENTS

Copyright Statement.....	1
Abstract.....	3
List of Contents.....	4
List of Tables.....	7
List of Figures.....	8
Acknowledgements.....	9
Author's declaration.....	10
Chapter 1 Literature Review.....	11
1 Theory of mind.....	12
1.1 Definition.....	12
1.2 Mental states.....	14
1.3 The function of having a theory of mind.....	15
1.4 Theoretical background of theory of mind.....	17
2 Assessing Theory of Mind: The False Belief Task.....	21
2.1 Description of the false belief task.....	23
2.2 Main findings.....	27
2.3 Clinical populations and the false belief task.....	29
2.4 Evaluation of the false belief task.....	32
3 Reasoning.....	35
3.1 Development of thinking and reasoning in children.....	35
3.2 Syllogisms and Contrary-to-fact reasoning.....	38
3.3 Counterfactual reasoning.....	42
3.4 Counterfactual thinking and the false belief task.....	49
4 Conclusion.....	55
Chapter 2 Experiment 1: Affirmative and negative counterfactual questions.....	57
1 Introduction.....	57
2 Method.....	64
2.1 Participants.....	64
2.2 Material.....	64
2.3 Design.....	66
2.4 Procedure.....	66
3 Results.....	67
3.1 Comparison of the different types of questions.....	67
3.2 Analysis of mistakes.....	69
4 Discussion.....	70
Chapter 3 Experiment 2: Affirmative and negative counterfactual questions and the false belief task.....	75
1 Introduction.....	75
2 Method.....	83
2.1 Participants.....	83
2.2 Material.....	83
2.3 Design.....	85
2.4 Procedure.....	85
3 Results.....	85
3.1 Counterfactual Questions.....	85

3.2 Analysis of mistakes for the counterfactual tasks:	86
3.3 False belief tasks for the whole sample	87
3.4 Analysis of mistakes for the false belief task	88
3.5 Comparison between performance on counterfactual tasks and false belief tasks	89
3.6 Correlations between counterfactual thinking and false belief task for the whole sample	90
4 Discussion	92
<b>Chapter 4 Experiment 3: False belief tasks, counterfactual thinking and contrary to fact syllogisms</b>	<b>100</b>
1 Introduction	100
2 Method	107
2.1 Participants	107
2.2 Materials	108
2.3 Design	113
2.4 Procedure	113
3 Results	113
3.1 Control questions and biases	113
3.2 Counterfactual task, false belief to others tasks and false belief to self tasks	115
3.3 Contrary to Fact Syllogisms	117
3.4 CNRep	117
3.5 Correlations	118
4 Discussion	122
<b>Chapter 5 Experiment 4: the effect of context on counterfactuals and contrary-to-fact syllogisms</b>	<b>132</b>
1 Introduction	132
2 Method	140
2.1 Participants	140
2.2 Material	140
2.3 Design	142
2.4 Procedure	143
3 Results	143
4 Discussion	146
<b>Chapter 6 Discussion</b>	<b>153</b>
1 Introduction	153
2 Main goals	153
3 A review of the studies conducted in this thesis	154
4 Main issues	158
4.1 Counterfactual thinking	158
4.2 Counterfactual thinking and the false belief task	160
4.3 Contrary-to-fact syllogisms	161
4.4 Counterfactual thinking and contrary-to-fact syllogisms	162
4.5 Control questions: do they have their place?	163
4.6 The role of imagination	164
5 Theoretical background to the findings	164
6 Future work	167
6.1 Conclusion	171
<b>Appendices</b>	<b>173</b>
<b>Appendix A Recording Sheet for Experiment 1</b>	<b>174</b>

<b>Appendix B Other counterfactual and future hypothetical scenarios used in experiment 1.....</b>	<b>176</b>
<b>Appendix C Recording sheet for experiment 2.....</b>	<b>178</b>
<b>Appendix D Description of the tasks used in experiment 2.....</b>	<b>194</b>
<b>Appendix E Experiment 2: Analysis with the reduced sample.....</b>	<b>202</b>
<b>Appendix F Counterfactual scenarios use in experiment 3.....</b>	<b>206</b>
<b>Appendix G False belief to others task used in experiment 3.....</b>	<b>208</b>
<b>Appendix H False belief to self tasks used in experiment 3.....</b>	<b>210</b>
<b>Appendix I Contrary-to-Fact Syllogisms used in experiment 3.....</b>	<b>211</b>
<b>Appendix J Counterfactual scenarios used in experiment 4.....</b>	<b>212</b>
<b>Appendix K Second analysis for experiment 4.....</b>	<b>214</b>
<b>References.....</b>	<b>215</b>

## LIST OF TABLES

Table 1: Number of right answers for each type of question (n=61).....	67
Table 2: Total number and percentage of correct answers (n=122).....	68
Table 3: Number and percentage of types of errors for each type of question.....	69
Table 4: Means and standard deviations for affirmative counterfactual questions and negative counterfactual questions (n=41) – Maximum = 2.....	86
Table 5: Total number and percentages of each type of errors for counterfactual affirmative and counterfactual negative questions (n=41). ....	87
Table 6: Total number of each type of errors for theory of mind questions (n=41). ....	88
Table 7: Means and standard deviations for the total number of correct answers for counterfactual questions and false belief questions (n=41). ....	89
Table 8:Correlations between counterfactual thinking and false belief task (n=41). ....	90
Table 9: Correlations between the number of realist errors and performance of false belief questions (n=41). ....	91
Table 10: Correlations between the number of realist errors and performance on counterfactual and theory of mind tasks. ....	91
Table 11 . Crombach's alpha for each task.....	115
Table 12. Mean and standard deviation for each task (n=71).....	116
Table 13:Correlations between all the tasks separating yes and no syllogisms.....	118
Table 14:Rotated factor matrix in factor analysis .....	122
Table 15 Subtractive MD over the false belief to others task.....	126
Table 16. Subtractive MD over the counterfactual task .....	126
Table 17. Subtractive MD over false belief to self task.....	126
Table 18. Mean and standard deviation of the scores for each type of tasks in the two different contexts. ....	144
Table 19.Means and standard deviation for yes and no syllogisms. ....	145
Table 20.Means and standard deviation for each type of counterfactual question.....	145
Table 21. Correlations. ....	146



## **LIST OF FIGURES**

Figure 1:Percentage of correct answers for each type of questions.....	88
--	----

## **ACKNOWLEDGEMENTS**

I would like to thank all of those who have given me support and assistance in the process of completing this study:

- Both my supervisors, Professor John Clibbens and Professor Simon Handley for their enthusiasm and their excellent academic support during this PhD, but also before that, to be exact for the last 9 years altogether.
- The University of Plymouth for the funding of the research.
- The Speech and Language Therapy team at the College of Saint Mark and Saint John for their support, especially Anne Ayre and Maggie Cooper.
- My friends and family who have been very patient and supportive during these last few years... especially Guillaume as well as Pierre and Ulysse who both came into the world during this PhD.
- Helen Greathead who could really understand what writing is all about and shared my loneliness at times.
- Johnny Hallyday who helped me through the most stressful times, when nobody else could!

## **AUTHOR'S DECLARATION**

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

This study was financed with the aid of a studentship from the University Of Plymouth.

This thesis is the result of the author's own investigation, carried out under the guidance of her supervisor. Other sources are explicitly referenced.

Relevant seminars and conferences attended at which the work was presented:

### **Conferences Attended**

Arreckx, F, Clibbens, J., and Handley, S. (2005). Counterfactual reasoning and contrary-to-fact syllogisms: the effect of fantasy context. Paper presented as part of a symposium on counterfactual thinking in children, European Society for Developmental Psychology, University of La Laguna, Tenerife. (August 2005)

Arreckx, F., Clibbens, J. & Handley, S. (2004). Counterfactual reasoning, theory of mind and contrary-to-fact syllogisms Paper presented at the Child Language Seminar, Bristol (July 2004).

Arreckx, F., Clibbens, J. & Handley, S. (2003). Counterfactual reasoning and theory of mind. Paper presented at the Child Language Seminar, Newcastle (July 2003).

Word count of main body of thesis: 48,230

Signed 

Date.....17/11/07.....

# **CHAPTER 1**

## **LITERATURE REVIEW**

The main goal of this research project was to investigate children's reasoning abilities, especially to solve counterfactual tasks. The nature of counterfactual thinking itself was explored, in particular in relation to future hypothetical thinking and to the negative or positive formulation of counterfactual questions. Counterfactual thinking was also studied as a possible sub-component of the false belief task that is usually used to assess children's theory of mind. Finally this thesis also explored the effect of context on children's performance on counterfactual and syllogistic reasoning tasks.

The first chapter will give a review of the literature on theory of mind and the false belief task, and on reasoning in children, looking at the development of conditional reasoning, counterfactual reasoning and contrary-to-fact reasoning.

Chapters two, three, four and five will report the studies conducted and the results of these studies. Chapter six will be the final discussion.

This first chapter will review the literature on theory of mind and its methods of assessment; in particular the false belief task. Secondly the literature on reasoning development in children will be reviewed. This part will focus on the development of conditional reasoning, counterfactual reasoning and the ability to solve contrary-to-fact syllogisms in children. The link between all the types of reasoning mentioned above and the false belief task will be reviewed as well.

# 1 Theory of mind

## 1.1 Definition

Theory of mind has been at the centre of a considerable number of research projects in the last 25 years. Most research on theory of mind has been conducted with typically developing pre-school children with the view to understand how children develop a social understanding in the first few years of their life. However, theory of mind has also been studied with older children, adults with acquired brain injury and nonhuman primates, and a vast literature reports studies conducted with special populations especially deaf children and children with autism. This literature with older and/or special populations was conducted with the aim to contribute to the debate about the theoretical backgrounds to social understanding as well as understanding some specific disorders.

The pioneers in the area of theory of mind were Premack and Woodruff (1978) who were the first ones to use the term 'theory of mind' for which they gave the following definition: 'In saying that an individual has a theory of mind, it means that the individual imputes mental states to himself and others (...)' (p.515). Having a theory of mind is the ability to understand that people have different mental states than our own and to understand what these mental states are.

Premack and Woodruff's (1978) research examined whether chimpanzees have a theory of mind. They reported an experiment where a chimpanzee was shown a video of a human being trying to solve a problem and was then asked to pick a still photo that showed the accurate solution to the problem. They used a variety of videos and problems. Some of them were simple and were the type of problems the chimpanzee could often be confronted with, such as inaccessible bananas. Others were more complex, more the type of problems humans are confronted with; for example, in one of the scenarios, the experimenter tried to

play a gramophone that was not plugged in. The 14-year-old chimpanzee tested in this study performed well above chance level answering correctly 22 of the 24 scenarios presented to her for the simple problem and performed again above chance level for the more complex problems.

Premack and Woodruff (1978) argued that the chimpanzee was able to solve these problems by attributing mental states to the human actor. For example, regarding the simple problems where the human actor was confronted with a physically inaccessible banana, Premack and Woodruff (1978) explained that in order to solve that problem the chimpanzee needed to make sense of the situation, understanding that the human actor wanted the banana but could not reach it. She also needed to assume that the human actor knew what he needed to do to reach it, so when she was presented with photographs presenting solutions to the problem she would choose the correct one. They therefore concluded that the main mental states involved in that task were intention, or purpose, and knowledge, or belief. They argued that theory of mind is not a sophisticated act but a primitive one, as a chimpanzee can show signs of having a theory of mind, and that it seems to be innate rather than learned. However, human beings have a much more developed theory of mind as they can understand a vast number of mental states in a wide variety of situations. It is also worth noting that the chimpanzee that took part in this experiment was a highly trained chimpanzee that was used to human beings. It is very difficult to generalise these findings to chimpanzees and animals as a whole. The specific aspects of humans' theory of mind will be discussed in this chapter.

## **1.2 Mental states**

Having a theory of mind is the ability to attribute mental states to individuals, but what are these mental states? Mental states are by definition not directly observable (Premack and Woodruff, 1978) as they are people's thoughts and feelings. However, they can be detected through observable behaviours and through the mental verbs people use such as to believe, to think, to know, to like, to guess, to doubt, to pretend, to promise, to trust. The variety of mental verbs reflects the variety of mental states but Premack and Woodruff (1978) stated that the mental states we impute most often are 'purpose' and 'intention'. Bartsch and Wellman (1989) identified two main categories of mental states: 'desires' and 'beliefs' and those two main categories of mental states will drive a person's behaviour. It also can be said that 'purpose' and 'intention' drive people's behaviour. Human behaviour is very complex and mental states are just one of the factors that influence it. A large variety of other mental states also exist, such as pretence and imagination (Slaughter and Repacholi, 2003), wishes, hopes and ideas (Wellman and Woolley, 1990).

The mental states described above are the ones adults will impute. Although they are all present in adulthood they do not all develop at the same time in children (Wellman and Woolley, 1990; Wellman and Liu, 2004). Wellman and Liu (2004) conducted a meta-analysis of 45 different studies and found that children can correctly understand people's desires before they can understand their beliefs. This had previously been suggested by Wellman and Woolley's (1990) findings that showed that older two-year-olds can predict characters' behaviour/action according to their desires. This could be explained by the fact that desires are more important in children's life than beliefs are.

As mentioned above, mental states do not all develop at the same time but it can also be noticed that each particular mental state develops in a particular manner. Wellman and Bartsch (1994) described children as 'Desire Psychologists' as they tend to use desire mental state verbs (e.g.: want, wish, hope...) before belief mental state verbs (e.g.: think, know...). Wellman and Liu (2004) reported the development of desire in children. They explained that children can first judge that two different people can have different beliefs, then they can judge how a person will act according to their own beliefs, and finally they can judge whether a person's belief is true or false.

Bartsch and Wellman (1989) showed that children have some appreciation of the existence and implication of false beliefs. They can explain a character's action according to their false belief, which is easier than predicting behaviour from false belief as at three years of age the satisfaction of desires is more important than beliefs.

### **1.3 The function of having a theory of mind**

Having a theory of mind, being able to perceive other people's mental state, is extremely important to human beings. It is an essential component of social interactions. It allows us to make sense of ourselves and others (Wellman, 1990). It can help us understand the reasons why people behave in the way they do as behaviour is a product of mental states and attitudes (Wellman and Woolley, 1990). Wellman (1990) explained that the conception of mind relies on the interaction between belief, desire (two mental states) and action, with belief and desire preceding the action. If a child wants some chocolate (desire), he will look where he believes the chocolate is (belief), and then the search can start (action). In this particular example action is a product of desire and belief. Therefore, belief and action can help explain and even predict people's behaviours. Being able to



interpret our own as well as other people's behaviour is very useful but we cannot be sure that our interpretation is the right one. People are not always aware of the reasons that make them behave in a certain way. Interpreting other people's behaviour can be even more uncertain.

In order to understand people we need a theory of mind. Wellman and Lagattuta (2000) claimed that when we think of people we do it in terms of mental states, in terms of their beliefs, desires, hopes. Having a theory of mind is important for social interactions and communication.

Having a theory of mind has an implication for every day social interactions, and delays or deficits that can be observed in theory of mind can have important consequences for social development (Keenan, 2003). Being able to understand other people's mental states is essential to be integrated in human society. There is a positive correlation between success on the false belief task (a task that measures theory of mind) and quality of social development (see Keenan, 2003, for a review). However, it is difficult to know which one comes first. Does having a good theory of mind facilitate social development or does social development influence theory of mind? We need to understand people's thoughts and feelings to be integrated, to be 'popular' and Mitchell (1997) made a very strong claim saying that a human being who is not able to do so could become a social outcast.

Understanding people's mental states allows us to communicate more efficiently with them. Having a theory of mind helps us better understand what people really mean when they communicate. Knowing what the speaker's factual beliefs are allows us to interpret his, or her, discourse correctly (Mitchell, 1997); the communication then becomes clearer. The same discourse can be interpreted differently according to who the speaker is and what his, or her, own beliefs are. Most human interactions are very subtle. What people say with language is not always a true translation of their thoughts and beliefs. We need to understand

what people really mean to form friendships and love relationships that are essential to human beings (Mitchell, 1997). At a more sophisticated level having a theory of mind allows us to understand nonliteral language such as sarcasm and irony (Keenan, 2003). In order to understand sarcasm the listener needs to recognise through non verbal cues and context that the speaker's intention is to mean the opposite of what he said. For example, if the speaker says 'What a nice car' referring to an old car in poor condition he actually means the opposite of what was said. In order to understand this sarcasm the listener needs to have a good theory of mind. Understanding sarcasm is a sophisticated skill and children do not seem to be able to understand it until the age of 6 or 10 (Keenan, 2003). This shows that theory of mind develops progressively in childhood.

#### **1.4 Theoretical background of theory of mind**

The vast literature on theory of mind has studied the development of theory of mind with young children and has also tried to give a theoretical explanation to that development. The theoretical debate has mainly been centred domain general versus domain specific approaches, but more theories within each approach have also been developed.

In domain general approaches knowledge of the mind is seen as conceptual and conceptual knowledge arises from general processes (Wellman, 1992). The first domain general approach suggests that executive functions are necessary to build up a theory of mind but then, once this theory is acquired, they are no longer necessary (McKinnon and Moscovitch, 2007), therefore executive functions are necessary during the developmental phase of theory of mind only. The executive functions required are planning, inhibition, self-regulation and cognitive flexibility (Wellman, 1992). The importance of inhibition was demonstrated by Carlson and

Moses (2001) who found a link between performance on the false belief task and inhibitory control.

Another domain general approach is the representational-mind account described by Perner (1991). According to Perner (1991) children eventually understand mental representation after developing representations and multi-models of the world. Children need to be able to understand that multi-models are some representations of the world and they need to be able to manipulate their models to understand false beliefs. False beliefs understanding requires them to compare their own model of the situations with somebody else's. Perner (1991) suggested that understanding mental representations could be just a part of domain-general understanding of representation.

In contrast to domain general approaches, domain specific approaches have also been developed. These approaches assume that human understanding of mental states stems from specific knowledge, processes and mechanisms; that there is a domain specific mechanism that governs social cognition (Wellman, 1992). One of these approaches is the dominant account in theory of mind: the theory-theory account (Morton, 1980). The theory-theory has mainly been developed over the years by Wellman. Wellman and colleagues argued that in order to explain people's behaviour and talk children need to formulate a domain specific theory (Wellman, 1990). This approach sees the child as a scientist who is going to develop a theory for a domain (Wellman, 1990; Perner, 1991) which means that children actually need to learn theory of mind. The children need to modify their theory in light of new evidence. In this approach belief and desires fit into a wider causal network where each piece of knowledge has its place within a larger conceptual framework in a domain (Carpendale and Lewis, 2006).

Another domain specific approach that was developed alongside the theory-theory is the simulation approach. According to the simulation view children

develop an understanding of mental states through projecting themselves into other people's situations. They can then imagine the experience and then simulate how they would feel and finally generalise their perspective to others (Harris, 1989).

These approaches contrast with the modular approach that was influenced by Fodor's (1983) work. Fodor (1983) revived the idea of the modularity of the mind. He argued for a nativist perspective where the cognitive system would be modular, or in other words, the cognitive system would be made of subsets of cognitive systems called modules. In this theory modules are information-specific units that are domain specific. The modules are highly specialised, autonomous, automatic and fast. Fodor's work has been widely influential across psychology and the concept of modularity was applied to theory of mind suggesting that theory of mind arises from an innate, encapsulated and domain specific module (Scholl and Leslie, 1999).

Scholl and Leslie (1999; 2001) argued more specifically for a ToMM (Theory of Mind Mechanism) that is an innate module specialised in computing mental states. ToMM is an innate meta-representational basis that incorporates innate concepts such as belief and pretence. Being innate this module is available to children before they have developed general problem-solving resources and it allows children to have an insight into others' mind very early on. According to Scholl and Leslie (2001) this module allows children to spontaneously process behaviours and understand mental states. This module is very specific to theory of mind and does not apply to any other cognitive domain. However, Scholl and Leslie (1999; 2001) also explained that not all the theory of mind is modular; theory of mind has a specific innate basis that is ToMM, but it also includes a general executive process that they called SP (Selection Processing). The SP is non modular and is necessary to impute false beliefs. It allows the children to

inhibit the unwanted responses, such as, for example, a false belief. The SP is domain general support the hypothesis that executive functions are necessary to solve a false belief task. The difference between this theory and the domain general theory suggesting that inhibition is necessary to build up a theory of mind is the timeframe. In the domain general approach inhibition is required at the beginning, while the child is developing a theory of mind whereas in the domain specific approach inhibition is necessary in the long term as it is a selection process required for at least some of the theory of mind tasks. This domain general theory was further supported by subsequent studies conducted by McKinnon and Moscovitch (2007) who found that even with adult and older adult central processing is required to solve some complex false belief tasks.

The concept of modularity was also supported by the studies conducted with special populations, in particular children with autism. It appears that children with autism have a specific impairment of theory of mind (Baron-Cohen, 1995; Happé, 1990) with a specific difficulty in understanding thought and feeling of others while other abilities, such as inhibition, are spared. This suggests that the module ToMM is specifically impaired in autism.

In summary, no consensus has been reached and the debate between whether theory of mind is domain general or domain specific is on-going. Scholl and Leslie (1999; 2001) tried to reconcile the two by arguing that both types of abilities are necessary to solve theory of mind situations.

## **2 Assessing Theory of Mind: The False Belief Task**

In order to study theory of mind in the most valid and reliable way possible, a wide range of tasks has been developed to assess children, from naturalistic types of tasks involving free play and free speech to very structured tasks following a clearly defined protocol.

Naturalistic language transcripts have been used in the literature to analyse the different mental states children can represent (see Bartsch and Wellman, 1995). This activity involves sampling children's spontaneous speech and recording the use of mental state verbs by the children. This allows the researcher to investigate whether the children show signs of having a theory of mind, and use it when playing freely. However, these naturalistic tasks are not the most commonly used tasks as they tend to be difficult to put in place, time consuming and can lack reliability. It is more for controlled tasks to be used.

Most tasks assessing theory of mind require children to explain and predict a character's behaviour, assuming that they are going to act according to their mental states and beliefs. For example, if a character wants to retrieve a target object the children need to understand this character is going to look for the object where he, or she, thinks it is even if it is not where the object really is. The most widely used task measuring theory of mind is the false belief task in which 'accurate predictions about another person's behaviour can be derived by referring to that person's mental (mis)representation' (Repacholi and Slaughter, 2003, p. 4). We will consider this task in detail shortly.

Other controlled tasks have been used too, especially with high functioning children and teenagers with autism. Brent, Rios, Happé and Charman (2004) used advanced theory of mind tasks to test high-functioning 6- to 12-year-old children

with autism: Strange Stories and Cartoons that were first designed by Happé (1994b). In the context of Strange Stories the participants were presented with some stories involving mental states and some physical stories without any mental states as a control. In order to understand the first type of stories the participants needed to identify the underlying intention of the character. These stories assessed whether the participants could understand lies, white lies, double bluff, persuasion and misunderstanding. These concepts are more complex than simple false beliefs. Regarding the Cartoons, half of them required the participants to understand characters' mental states and the other half were based on mechanical/behavioural humour and did not involve any mental states.

A more advanced stage of theory of mind is the 'second order theory of mind' which is the ability to understand a person's thought about another person's thought (Perner and Wimmer, 1985). Steele, Joseph and Tager-Flusberg (2003) presented children with autism with second-order false belief tasks in the format of picture stories. In one of them a child character is about to receive a surprise gift from a parent. Without the parents being aware of it the child inadvertently finds the present. The autistic children were then asked second-order ignorance, belief and justification questions. They needed to show that they were aware of what the parent knew, or thought, about what the child knew, or thought. These tasks are much more complex than the simple false belief task.

The purpose of this thesis was to study the classic false belief task in relation to counterfactual thinking so the next part will analyse the structure of this task in more detail.

## **2.1 Description of the false belief task**

The prolific amount of research conducted in the domain of theory of mind has seen an important increase in the number of tasks used. One of the main ways of assessing theory of mind is the false belief task.

A variety of false belief tasks have been used in the literature but most follow a similar framework to that was described by Wimmer and Perner (1983). In that framework, the participant is aware that another character is aware of a particular state of affairs, X. This other character leaves the scene and as he/she is away the participant witnesses a change of state of affairs from X to Y. The participant knows that the present state of affairs is now Y but that the other person still believes that it is X. In order to pass the false belief task the participant needs to acknowledge the character's false belief regarding the state of affairs. The change of state of affairs usually consists of the change of location of a particular object.

Most false belief tasks include two control questions at the end: the reality question and the memory question. The reality question checks that the child is really aware of the final state of affairs and the memory question checks that the child remembers the first state of affairs. It is argued that if children fail either of these control questions, their answer to the critical question should be discarded as it then seems that the child is not fully aware of the situation in front of him/her (e.g.: Baron-Cohen, Leslie and Frith, 1985). These short stories are usually acted out with puppets and toys, but in some cases the story is presented on paper as a story board.

One of the first controlled false belief tasks reported in the literature was the Maxi task that was designed by Wimmer and Perner (1983). The Maxi task related a story involving two characters, a little boy called Maxi and his mother. The following story was told to the children and acted out with dolls and props. Maxi's



mother comes back from the shop. She has bought some chocolate to bake a cake. Maxi helps her put the things away. He asks his mum where he should put the chocolate and his mother replies: 'in the blue cupboard'. Maxi puts the chocolate in the blue cupboard. The children are then told that Maxi remembers exactly where he put the chocolate so that he can get some later, and that he loves chocolate. Then Maxi goes to the playground. At this stage, the boy doll is removed from the scene. While Maxi is away the mother starts preparing the cake. She removes the chocolate from the blue cupboard. She grates some chocolate to put in the mix and then she puts it back in the green cupboard instead of the blue one. She then realises that she forgot to buy some eggs so she goes to her neighbour's house for some eggs. Then Maxi comes back from the playground; he is hungry and he wants some chocolate. The boy doll is then brought back to the scene. The children are reminded that Maxi still remembers where he had put the chocolate and they are asked the false belief question: 'Where will Maxi look for the chocolate?' The correct answer is 'in the blue cupboard' as Maxi is not aware of the change of location. The answer can be verbal or the child can just point to the right location. The child is then asked two control questions: the reality question and the memory question. The reality question is: 'Where is the chocolate really?' This control question is meant to assess whether the child is aware of the real situation, the real location of the chocolate. The memory question is: 'Do you remember where Maxi put the chocolate in the beginning?' This question is meant to assess whether the child can remember the first state of affairs. This scenario described the basic Maxi task. In their paper, Wimmer and Perner (1983) presented the children with two different versions of this task: the cooperative story version and the competitive story version. In the cooperative version, the cupboards are fixed high on the wall and Mummy needs to lift up Maxi so he can put the chocolate in the cupboard, therefore when Maxi comes back he needs

somebody to help him retrieve the chocolate. He asks his grandfather to help him. His grandfather asks him where the chocolate is and the children are asked the following question: 'Where will Maxi say the chocolate is?' The children need to acknowledge that Maxi is not aware of the change of location and that he is going to tell the truth to his grandfather as he wants the chocolate. In the competitive version the cupboards are fixed low on the wall. In this story when Maxi comes back from the park, his big brother is already in the kitchen and is looking for the chocolate. The children are told that Maxi wants to tell something wrong to his brother so he does not find the chocolate. So when the children are asked 'where will Maxi say the chocolate is?' they should say in the green cupboard. This version seems more complicated because the child needs to understand that Maxi will say what he believes to be the wrong location which is actually the true location of the chocolate. The results showed that when the children answered the belief questions correctly there was then no difference between the competitive and the cooperative versions.

The Sally and Ann task (Baron-Cohen, Leslie and Frith, 1985), another extremely widely used task, was developed following the paradigm described by Wimmer and Perner (1983). This task was primarily designed to test theory of mind in children with autism but has also been used with typically developing children. The task involves two dolls, Sally and Anne. Sally has a basket and Anne has a box. Sally places a marble in her basket and then leaves the scene. While she is away, Anne takes the marble and puts it in her box. Sally then comes back and the child is asked the critical question: 'Where will Sally look for her marble?' In order to pass this task the child needs to point to the previous location of the marble: Sally's basket. As previously two control questions are then asked; the reality question: 'Where is the marble really?', and the memory question: 'Where

was the marble at the beginning?' This task follows a simpler scenario than the Maxi task and includes fewer narrative elements.

Other tasks have been developed more recently. An example of a more recent false belief task is the Post Office story created by Riggs, Peterson, Robinson and Mitchell (1998). In this story Sally and Peter (two little characters) are in their house. Peter is not feeling very well so he goes to bed. Then, Sally goes to the shops to get some medicine. While she is away the man from the post office rings Peter to ask him to come to help put out a fire. Peter gets out of bed and goes to the post office. The children are then asked whether Sally knows where Peter is, and they are corrected if they give the wrong answer. Sally then comes back and the children are asked the critical question: 'Where does Sally think Peter is?' This story again was acted out with toys. In the particular task the authors did not mention the use of control questions.

All these tasks follow the same pattern regarding the change in the state of affairs described at the beginning of this section. These tasks can all be described as 'false belief to others' tasks because the participants are requested to attribute false beliefs to another person. In the literature there is also a range of tasks testing false belief to self, where the children have to attribute false beliefs to themselves. The false belief to others tasks described previously are all based on a change of location of the target object whereas the false belief to self tasks are based on a change of identity or content.

The main false belief to self task is the Smartie task which was designed by Perner, Leekam and Wimmer (1987). Prior to testing, the content of a Smartie container is replaced by a pencil, without the participants witnessing this change. The participants are shown the Smartie box and are asked 'What's inside the box?' This container is highly familiar to children so they generally give the right answer 'Smarties' or 'sweeties'. They are then shown that the box actually

contains a pencil, and not Smarties. The pencil is then put back into the box and the box closed again. At this stage the children are asked a control question to check if they can remember the actual content of the box: 'Can you remember what's inside here?', and then the test question about their previous belief: 'But what did you think was in here?' If the children are able to acknowledge their own false belief they should say: 'Smarties'. Other versions of the task have been used in the literature. However, the false belief to others task has been used more frequently in the literature than the false belief to self task. At this stage the reason why is not clear. These two tasks will be studied in more depth later on in this thesis.

## **2.2 Main findings**

The false belief task has been used in a vast number of studies looking at different aspects of theory of mind. The link between theory of mind and language or social environment has been studied in particular, as well as the components of the task itself. Wellman, Cross and Watson (2001) ran a meta-analysis of theory of mind development and the false belief task that included 178 separate studies, published or unpublished up to January 1998 across seven different countries. This meta-analysis included 479 conditions and 4 000 typically developing children. The size itself of this meta-analysis makes it more reliable than individual studies.

Wellman et al. (2001) looked at many conditions individually and found that some of them did not have any effect on the children's performance. For example, the year of publication of the studies did not influence the results nor did the medium of presentation (i.e.: toys, pictures or video). Whether the false belief task is presented with real people, puppets, storybooks or videotapes the performance is the same. This finding gives some freedom to the experimenter regarding the

choice of material used. It therefore allows for a greater variety in the support material to be used, which allows the experimenter to specifically adapt the material to the population to be tested especially regarding their age, abilities or special needs. However, Wellman et al. (2001) reported five variables that have a significant effect on performance; they were motive, participation, salience, real presence and country. For example, stating deception explicitly during the presentation of the false belief task enhanced the performance of children of all ages. If the children made the transformation themselves, for example, they moved the target object from one location to the other, it improved their performance. If the false belief of the character is made more salient, i.e. clearly stated, it improved the children's performance but only for the younger children, whereas the real presence of the object improved the performance of the older children. Wellman et al.'s (2001) meta-analysis included studies from different countries from Europe, North America, South America, East Asia, and Africa and from Australia. They found that the pattern of development was very similar across countries despite some slight difference in performance between countries. They used the United States and the United Kingdom as a baseline as they represented the largest sample. The meta-analysis showed that children in Korea performed as well as the baseline but that children in Australia and Canada performed better than the baseline, whereas children from Austria and Japan performed worse.

Their results also showed a strong age effect. The performance consistently increased with age, as could be expected, with most children failing the task before the age of 3 but most of them passing it at the age of 4, therefore four seems to be the critical age to be able to solve the false belief task. The research suggested that there is an important developmental change between the ages of 3 and 5, with younger children performing below chance level and older children

above chance level. The false belief task also showed good reliability across the different studies, with the results being highly comparable between studies.

Wellman et al. (2001) also compared the false belief to others and the false belief to self task and they did not find any difference in performance between those two types of task, with the younger children performing exactly at the same level. Attributing false belief to themselves or to another character does not seem to be easier or more difficult.

The false belief task has also been used with non typically developing children such as children with autism, children with learning disability, children with Down syndrome and Deaf children. It appears that some children experience difficulty in developing a theory of mind. This development can be simply delayed or sometimes children do not fully develop a theory of mind. The performance on the false belief task with special populations will be discussed in the next section.

### **2.3 Clinical populations and the false belief task**

The work on autism and theory of mind has been very important in the past 20 years. Autism is a severe childhood disorder that appears in the first three years of life (American Psychiatric Association, 1987). It is often described as a psychosis as the children affected seem to be isolated in their own world (Baron-Cohen, 1994). Children with autism present three fundamental impairments captured in Wing's triad: they have a qualitative impairment of social interaction, a qualitative impairment of communication and a restricted repertoire of activities and interests (Happé, 1994a). These impairments directly affect the way children with autism interact with other people and the way they play. They display very little pretend play and cannot engage with complex cooperative play with their peers because of their communication difficulties.

Baron-Cohen, Leslie and Frith (1985) were at the origin of this important research area on autism. They predicted that autistic children lacked a theory of mind. They designed the now famous 'Sally and Anne task' (as described above) to test their hypothesis. In this study they found that the autistic children consistently pointed to the location where the marble really was, rather than where Sally thought it was. They concluded that these children could not pass this task because they did not distinguish their own beliefs from the doll's. They could not see the situation from the doll's perspective. Baron-Cohen et al. (1985) explained these results by saying that autistic children were not able to represent mental states. They argued that this is a specific deficit that is attributable to autism, and that cannot be attributed to general intellectual level, as the children with Down syndrome included as a control in their study did not show this deficit. This specific deficit potentially explains the lack of pretend play and social impairment in autism. If children with autism are not able to represent mental states then they cannot share pretence with other children and communication becomes very difficult as they do not know what knowledge they share, or do not share, with their interlocutor.

More evidence tends to suggest that performance on the false belief task is not directly linked to intelligence as even high functioning older children with autism have been shown to be impaired in advanced theory of mind ability in comparison to typically developing children. (Brent, Rios, Happé and Charman, 2004; Steele et al., 2003). It seems that theory of mind impairment is specific to autism and tends to still be present with high functioning autistic children and continues even in older children who perform at a lower level than control groups (typically developing children matched on intelligence and verbal abilities) on more advanced theory of mind tasks. However, it appears that children with autism can improve their performance on theory of mind tasks with age (Steele et al., 2003).

All these findings suggesting a specific theory of mind impairment in autism support the hypothesis presented by Scholl and Leslie (1999; 2001) that theory of mind is, at least partially, an innate module.

The modular hypothesis is also supported by some data collected with patient with acquired neurological disorders. It has been shown, for example, that adults with no anterior difficulty in theory of mind can present some problems in this area following a stroke, especially when the right hemisphere is affected. Happé, Brownwell and Winner (1999) tested brain damaged patients after a stroke in the right hemisphere using theory of mind tasks that did not necessarily require the acknowledgement of false beliefs. The tasks used were stories or cartoons involving some double bluff, persuasion or white lies. Happé et al. (1999) showed that these patients had a specific impairment in understanding stories and cartoons requiring mental state attribution, but they did not find this impairment in patients suffering from damage in the left hemisphere. This tends to suggest that the right hemisphere of the brain might be involved in theory of mind understanding. This supports Fodor's (1983) claim that modules are a fixed specialised portion of the neurological architecture.

The false belief task has also been studied with other special populations which this time highlighted the role of the environment. For example there is a large literature looking at deaf children and the false belief task. The results from these studies overall tend to show that language and communication are crucial to the development of theory of mind. The first studies published on this topic suggested a significant delay in the development of theory of mind in deaf children. For example, Peterson and Siegal (1995) found that only 35% of deaf children, aged 8 to 13, could pass the false belief task. These deaf children performed as badly as children with autism in the control group. Deaf children from hearing families seemed to show a deficit in theory of mind due to the lack of



communication with their family and their lack of familiarity with mental state conversation, whereas deaf children from deaf families, i.e. with a rich linguistic environment, did not have any difficulty with theory of mind, especially when tested in their preferred language: Sign Language. (e.g.: Courtin and Melot, 1998; Peterson and Siegal, 1999). These findings can still be coherent with the modular hypothesis as modules need to be triggered by the environment (Leslie and Scholl, 2001).

Following the work with special populations, it seems that theory of mind can be affected by developmental and acquired disorders. Some specific disorders such as autism or brain injuries in the right hemisphere of the brain have an impact but environmental factors such as lack of early rich interactions can influence its development too.

As the false belief task is one of the most commonly used tasks to assess theory of mind it will be assessed in the next section.

## **2.4 Evaluation of the false belief task**

The false belief task, as described earlier in this section, is widely used in theory of mind research, probably because it presents several advantages. First of all it is quick and easy to administer. Running one false belief task just takes a few minutes, and only some simple material, such as everyday toys, is necessary. It is also very engaging for the children who like being told short stories and being shown toys. Children usually enjoy doing the task. Finally it requires a simple answer. In most tasks, in particular in the change of location type task, such as the Sally and Anne task for example, the children do not even need to give a verbal answer; they just need to point to a particular location. This is always an advantage when testing very young children as they can be very shy and not inclined to speak to a stranger.

The false belief task has also shown good reliability across the literature: the results are robust and consistent, with most children passing this task from the age of four (Wellman, Cross and Watson, 2001); however, its validity has been questioned more recently. It has been argued that younger children do not fail the task because of their lack of conceptual competence but because of the cognitive demands of the task (Wellman et al., 2001). Children under the age of four show signs of having a theory of mind, by using mental state verbs for example, but cannot pass the false belief task (Wimmer and Perner, 1983).

Bloom and German (2000) even claimed that the false belief task should not be used to test theory of mind. They argued that the false belief task is a complex task requiring other skills than just having a theory of mind, and that theory of mind cannot be only restricted to passing the false belief task.

Astington (2003) supported this view by saying that the literature showed a difference between competence and performance on the false belief task. The complexity of the task and the other skills involved (memory, language, inhibition...) can prevent children from passing the false belief task although they do have a theory of mind, and they can understand people's mental states in a real world situation. The opposite situation is also possible; some high functioning individuals can pass the false belief task but cannot use their skills in a real life situation. However, Astington (2003) stated that the false belief task has some ecological validity as it relates to individual social behaviours in real life situations but it does not always reflect the children's social interaction in the real world. Some children with autism have shown ceiling effects in the traditional false belief task but their everyday communication was still odd. They could not use their full theory of mind in a communication situation. Although they could pass the false belief task they could not fully use this ability when communicating with others. The false belief task only assesses a basic theory of mind ability; everyday

communication requires a more sophisticated theory of mind. Everyday communication is complex and it involves dealing with several mental states simultaneously.

Some authors have claimed that other cognitive abilities than theory of mind, such as inhibition for example, are important to pass the false belief task. Inhibition is defined by Flynn, O'Malley and Wood (2004) as 'the ability to stop oneself from performing an action that is dictated by a prepotent response' (p.103). Inhibition develops between the age of 3 and 5 (Flynn et al., 2004), therefore at the same time as children start being able to pass the false belief task. Flynn et al. (2004) conducted a longitudinal study with children from the age of 3 to 5 to investigate the parallel development of inhibition and false belief understanding. The children were tested every four to six weeks for six sessions. Each session included two false belief tasks and two inhibitory tasks. The initial and the final sessions also included a measure of receptive vocabulary using the British Picture Vocabulary Scale. No significant correlations were found between BPVS scores and change in scores of inhibition and false belief understanding. However, they found that more children performed well on the executive inhibition task before they could understand false beliefs. When looking at children who were just starting to pass the false belief task they found that these children had good inhibition skills before they began to pass the false belief task. It therefore seems that inhibition could be a prerequisite to the false belief task. Leslie and Polizzi (2000) argued that inhibition is necessary to pass the false belief task. Children need to inhibit the true-belief target in order to pass this task, they need to be able to activate the SP (Scholl and Leslie, 1999; 2001).

It has also been argued that the ability to pass the false belief task is closely linked to the ability to solve counterfactual problems. The domain general hypothesis suggested that learning theory of mind is linked to executive functions

including reasoning abilities (Wellman, 1992). Riggs et al. (1998) argued that the classic false belief task includes a counterfactual element, so in order to pass the false belief task children need to be able to reason counterfactually. This link between the false belief task and counterfactual reasoning will be discussed in more depth later on in the thesis. Before exploring this relationship in depth it is necessary to have a good understanding of reasoning skills and their development in children, and in particular of counterfactual reasoning. The next part will review the literature published in these areas.

### **3 Reasoning**

#### **3.1 Development of thinking and reasoning in children**

The term 'thinking' embraces a variety of cognitive skills such as conceptual understanding, remembering, language, problem solving and reasoning, amongst other cognitive abilities (Siegler & Wagner Alibali, 2005). Thinking develops mainly from birth to adolescence (Siegler & Wagner Alibali, 2005) but its development is extremely rapid in the first few years of life. The main focus of this thesis is on typically developing children from the ages of 3 to 5, and by this age children have very good communication skills and show some logical reasoning abilities.

These rapid changes in children's thinking can be explained by two factors: biological changes in children and the social environment. Biological developments such as changes in the brain as a whole, changes in the structures within the brain and changes in neurones underpin this evolution in children's thinking (Siegler & Wagner Alibali, 2005). The contribution of the social environment is also of paramount importance as Vygotsky suggested in his theory. Vygotsky claimed that development occurs in social interaction, especially between the child and an adult or a more advanced peer (Siegler & Wagner

Alibali, 2005). This thesis focuses on one particular aspect of thinking: reasoning. Therefore the next section will address the development of reasoning in children.

Research on children's reasoning first started with Piaget in the middle of the 20<sup>th</sup> century. He was a pioneer in research in cognitive development in children from birth to adolescence. His work mainly explored how knowledge grows in children, and how children develop logical thinking. His main findings suggested that formal reasoning appears at the stage of formal operations around the age of 11 or 12 (Inhelder and Piaget, 1958). This finding has since been criticised as young children seem to be able to reason logically from a much earlier age (eg.: Leevers and Harris, 2000). Those findings will be discussed in more detail later on.

By the age of 3, children have usually experienced many physical causes and their effects (Goswami, 1998). They know, for example, that scissors cut paper, that if they spill water on their jumper it will get wet. They experience a lot of these physical situations as soon as their motor skills allow them to interact physically with their environment. In infancy, causality is tested using mechanical tasks: one agent object moves and has physical contact with a recipient object which then moves (Koslowski & Masnik, 2002). Research seems to show that young children understand causality in terms of spatial and temporal continuity, and there is no evidence of abstraction of the notion of causation being already in place (see Koslowski & Masnik, 2002 for a review). Older children can reason about non-mechanical causation, such as electrical phenomena. They then use formal rules and content interdependently to inform their judgement (see Koslowski & Masnik, 2002 for a review). Koslowski and Masnik (2002) claimed that being able to reason causally is linked to learning empirical relations in the world. Children need to evaluate the likeliness of a causal agent. For example, they need to evaluate whether car colour is linked to reliability. They also need to integrate some background information about the situation of the problem. These

results lay in the fact that most of the time, causal reasoning in children is a guess about plausibility.

Logical reasoning includes different types of reasoning such as deductive logic and reasoning by analogy that children acquire at an early age, and transitive and scientific reasoning acquired at a later age (Goswami, 1998). Transitive reasoning is close to what Piaget called the 'concrete operational stage' that children reach around the age of 6, and scientific reasoning is similar to Piaget's 'formal operation stage' reached at the age of 11 to 12, although the processes might not be as described by Piaget.

Inductive and deductive reasoning are both, to some extent, linked to the child's knowledge of the world. Inductive reasoning requires the child to include knowledge of the world to make inferences, whereas deductive reasoning problems can be solved without any other knowledge but can help to fill the gaps in knowledge (Goswami, 2002). Children are more likely to successfully solve an inductive reasoning problem if the premise and the conclusion are similar in content and if the premise categories are very typical (eg: a robin; a very typical bird) (See Goswami, 2002, for a review). Analogical reasoning is a particular type of inductive reasoning. It requires the present novel situation to be related to a past familiar one (Goswami, 1998). It is the basis of a lot of everyday problem solving. Faced with a novel situation, human beings, and children in particular, try to identify the similarities between this situation and a previous one. In order to reason by analogy the child needs to be able to match up situations. It seems that from the age of 4 children can reason by analogy when the context is familiar to them (see Goswami, 1998 for a review). The results are again influenced by the similarity of the premise and conclusion categories. The more the context is familiar to the child, the more successful the child will be (Goswami, 2002).

In order to solve deductive reasoning problems the child will have to reason; that is follow a mental process, using the information provided in the problem (Goswami, 2002). One of the simplest forms of deductive reasoning is the syllogism. Syllogisms include two premises and a conclusion. The premises and the conclusion contain some quantifiers such as 'all', 'some', or 'no' (Evans, Newstead and Byrne, 1993). The participants are then asked to evaluate whether the conclusion is valid. One example would be: 'All cows are mammals, all mammals are animals, therefore all cows are animals'. Sometimes the conclusion can be replaced by a question, and the syllogism therefore is: 'All cows are mammals, all mammals are animals. Are all cows animals?' This format is used more frequently with children. Adults and children tend to find syllogistic reasoning difficult. Although no reference to the real world is necessary to solve a syllogism, research has shown that adults and children perform better when they are familiar with the content of the syllogism is presented in. If the content is unfamiliar, participants tend to make empirical mistakes (see Goswami, 2002 for a review). The issues relating to syllogisms in an unfamiliar context will be developed further in the next section.

### **3.2 Syllogisms and Contrary-to-fact reasoning**

Children's reasoning on contrary-to-fact problems has been one of the main foci in studies of reasoning development. As mentioned in the previous section, children and adults find reasoning more difficult when performed with unfamiliar content; therefore reasoning with contrary-to-fact premises is an interesting way of studying the impact of content on performance. It also allows us to study how children can reason independently of the content of the problem. When solving a false belief task children are required to reason about beliefs, and about false beliefs in particular. They need to disengage from reality and then reason.

One of the first research papers on contrary-to-fact reasoning was published by Hawkins, Pea, Glick and Scribner in 1984. They presented children with syllogistic problems and varied the content from fantasy, to congruent or incongruent with practical knowledge. Their findings confirmed that young children can reason deductively and that the content of the problem affects the reasoning performance of the children. When children were asked to solve contrary-to-fact problems their performance was lower than when they tried to solve problems congruent with their practical knowledge of the world.

Several researchers have worked on contrary-to-fact problems with children, and investigated what type of information or context could help children improve their performance. There is a general consensus in the literature that a fantasy, or play, context helps children solve more contrary-to-fact syllogisms. (eg: Markovits and Vachon, 1989; Leever and Harris, 2000).

Markovits and Vachon (1989) asked children between the ages of 10 and 18 to solve contrary-to-fact modus ponens problems such as: 'If it rains, then the street will become dry. It is raining. The street will become wet.' Modus ponens are conditional inferences following the model: If p then q, p is true, therefore the conclusion is q (Evans et al., 1993). Half of the problems were set in an ordinary context and the other half in a fantasy context. In the fantasy context the children were told that the stories were set in an imaginary country where curious things happened. The results showed that the children's performance was improved in the fantasy context. Markovits and Vachon (1989) then argued that the fantasy context encourages children to accept the premises as if they were true.

However, the effect of fantasy context on contrary-to-fact problems seems to be different according to the type of reasoning problem. Markovits, Venet, Janveau-Brennan and Vadeboncoeur (1996) asked children questions equivalent to modus ponens, affirming the consequent, modus tollens, and denying the



antecedent (*MP*, *AC*, *MT*, *DA*). In all these types of conditional inferences the first premise is always 'If p then q' but the second premise varies as therefore does the conclusion. Modus ponens was explained above. The second premise of modus tollens (*MT*) is 'not-q' therefore the conclusion is 'not-p'. The second premise for affirmation of the consequence (*AC*) is 'q' and the conclusion 'p'. Finally the second premise for denial of the antecedent (*DA*) 'not-p' and the conclusion 'not-q'. One example was given by Evans et al. (1997) where the first premise was: 'If the key is turned then the engine will run'. The second premise and the conclusion of the four possible inferences were as follows: *MP*: the key is turned, therefore the engine is running; *DA*: the key is not turned, therefore the engine is not running; *AC*: the engine is running, therefore the key is turned; *MT*: the engine is not running, therefore the key is not turned. *MP* and *MT* are valid inferences and *AC* and *DA* are uncertain logical forms and they cannot always lead to a certain conclusion.

Half of the problems in Markovits et al.'s (1996) experiment were presented in an everyday context and the other half in a fantasy context. The results showed that when true premises were presented in a fantasy context the performance was reduced for the uncertain logical forms *AC* and *DA*. Markovits et al. (1996) explained these results by saying that the fantasy context reduces the probability of retrieving appropriate alternative relations. They argued that fantasy context acts as a cognitive filter reducing access to real world knowledge, in other words to long term memory (*LTM*). This explains why it improves performance on contrary-to-fact problems. The children can no longer access their real world knowledge so they reason with the premises only, therefore they cannot produce empirical responses. It allows them to make direct inferences but reduces their ability to make more complex inferences.

Another series of studies on contrary-to-fact reasoning has concentrated on syllogistic reasoning in the form of 'All cows quack, Susie is a cow, does Susie quack?' (e.g. Dias and Harris, 1988). The first premise of these syllogisms has a contrary-to-fact content. A variety of research papers have studied the types of contexts that would improve the children's performance by reducing the empirical bias they tend to make (e.g. Dias and Harris, 1990). The reasoning bias occurs when children reason according to their own belief about the world rather than the premises; for example when they answer 'no' to the example given above, as in the real world cows do not quack. A range of contexts has been shown to reduce this empirical bias. When the children were asked to engage in a pretend world through play (Dias and Harris, 1988) their performance improved. Giving them some make believe clues such as using imagery, using a different intonation or setting (Dias and Harris, 1990; Leevers and Harris, 1999) or even just asking them to think about the premises (Leevers and Harris, 1999) has the same effect. This effect is very robust as it can carry over a week (Leevers and Harris, 1999) and even two to three weeks (Leevers and Harris, 2000). According to Leevers and Harris (2000), the fact that this effect can persist for a few weeks goes against the theory of the cognitive filter suggested by Markovits et al. (1996). Leevers and Harris (2000) believe that these make believe clues only give a pragmatic clarification to the children. The make believe clues make clear to the children that they have to behave as if the premises were true. They help clarify the experimenter's intention. The children understand they are tested on their reasoning skills rather than their knowledge of the real world.

A closely related type of reasoning to contrary-to-fact syllogism is counterfactual reasoning. Counterfactual reasoning also requires children to suppress some aspect of their knowledge. This type of reasoning is going to be discussed in the following section.

### 3.3 Counterfactual reasoning

#### 3.3.1 Definition

Roese (1997) defines counterfactuals as alternative versions of the past. He explains that 'counterfactual does not refer to future prospects but only to negations of established facts.' (p. 133-134). They are contrary to the actual facts and events that have happened (Roese, 2003). Counterfactuals are characterised by a false antecedent. An example of a counterfactual could be 'If I had left early I wouldn't have missed my plane.' The antecedent, 'leave early', is false. So you infer that in reality the speaker did not leave early and did miss their plane.

Both contrary-to-fact and counterfactual reasoning require the suppression of some knowledge but they differ in their content. The content of a contrary-to-fact premise could not happen in the real world. An example is: 'all cows quack'. In the real world cows never quack. This type of premise is different from a counterfactual premise. The content of a counterfactual premise is contrary to what really happened but it could still have happened in the real world. In the example above ('If I had left early I wouldn't have missed the plane'), 'had left early' is the counterfactual premise. This premise implies that the speaker did not leave early but they could have left early, it could have been possible in the real world. In the literature, the term 'counterfactual' is sometimes used instead of 'contrary-to-fact'. Throughout this thesis the terms 'contrary-to-fact' and 'counterfactual' will be used as just defined.

A counterfactual reasoning problem implies that the antecedent is falsified but the antecedent can be falsified in two different ways: by the negation of the event or the affirmation of an alternative event. When thinking counterfactually we suppress a past event or replace it by another one and imagine what the present would then be like. For example the starting situation could be: Peter missed his train. On the way to the station he stopped at bakery A where there was a queue.

After he missed his train he could think counterfactually. He could try to cancel the antecedent (stopping at the bakery): 'If I had not stopped at the bakery I wouldn't have missed my train'. Or he could replace the antecedent by another one such as 'If I'd stopped at bakery B (where there was no queue) I wouldn't have missed the train.' After suppressing or modifying the antecedent he can imagine a different present situation. Counterfactual thinking is linked to imagination. You need to be able to imagine alternative versions of the past and the present.

When people reason counterfactually and falsify an antecedent it is usually with the aim of modifying the present situation. Counterfactuals can be classified according to two main ways: direction and structure (Roese, 2003). The direction of a counterfactual can be 'upward' or 'downward'. In upward counterfactuals the alternative circumstances are better than the reality and in downward counterfactuals where the alternative circumstances are worse than reality (Markman, Gavanski, Sherman and McMullen, 1993; McMullen, Markman and Gavanski, 1995). Upward counterfactuals are an attempt to modify an unpleasant present situation; for example 'If I had left early I wouldn't have missed the plane'. Downward counterfactuals are generally generated when a negative effect was narrowly avoided (Roese, 1997). People reflect on a lucky escape. An example could be as follows: Mary's alarm clock did not work, so she arrived late at the airport, and missed her plane that later on crashed. All the passengers were killed. The downward counterfactual generated could be: 'If my alarm clock had worked I would have died.'

### **3.3.2 Why do we need counterfactual thinking?**

Counterfactual thinking is omnipresent in everyday life (Byrne, 2002) but sometimes it tends to be stimulated in particular types of situation. Counterfactuals are usually activated in response to a negative outcome (Roese, 1997) for

example missing a train, getting poorly, having an accident or being in a near miss situation. However, some negative outcomes trigger more counterfactual statement than others. For example, the evaluation of the distance between the reality and some once possible but unrealised situation is an important factor when generating counterfactuals (Kahneman and Tversky, 1982). Failing an exam by 1 point will lead to more counterfactual thinking than failing the same exam by 20 points. In the first case it seems easier to undo the bad outcome than in the second case. The antecedent needs less modifying. For example the participant could think: 'If only I'd studied one more evening, if I hadn't gone out one evening... I would have passed the exam.' It gives them more hope of doing better next time. When we have missed a good outcome by a long shot, it is more difficult to think of alternatives to change the outcome. The antecedent would have to change radically. This closeness could be temporal, geographical or numerical (Roese, 1997). A temporal closeness could be missing a plane by 10 minutes (versus 1 hour), a geographical closeness missing a train by going to the wrong platform in the same station (versus going to the wrong station) and a numerical closeness failing an exam by one point (versus 20 points).

However, Roese (1997) stated that the main determinant of counterfactual generation is the affect: counterfactuals are mainly generated after a bad outcome. Roese (1997) explained that negative affect acts as a trigger telling us that a problem needs to be solved but it is also often link to emotion such as regret in that particular case (Roese, 2003). In conclusion, counterfactual thinking should have a beneficial effect. Counterfactual thinking gives some insight into the problem individuals are facing and gives them potential solutions (Roese, 1997). It helps them think about an alternative antecedent that could be chosen, so that the next time the same situation occurs the bad outcome will hopefully be avoided, and a bad outcome that was closely avoided will again be avoided.

### 3.3.3 Counterfactual thinking and developmental work

Counterfactual thinking in children has not been studied as a separate entity on its own to any great extent. Most of the work done in this area has studied the relationships between counterfactual thinking and other abilities, such as theory of mind in particular (eg: Grant, Riggs, Boucher, 2004; Perner, Sprung, Steinkogler, 2004; Riggs et al. 1998) or causal reasoning (e.g.: Harris et al., 1996).

Harris, German and Mills (1996) studied counterfactual thinking in children in relation to causality. They argued that in order to solve a counterfactual problem children need to use their existing knowledge of causal relationships in the world. Harris et al. (1996) used short scenarios they acted out with puppets and props to test the children. One of the scenarios involves a doll called Carol. The children were told the following short story:

*'One day, the floor is nice and clean like that (experimenter points to a square of white plastic). But guess what? Carol comes home and she doesn't take her shoes off. (A doll is brought to the edge of the surface). She comes inside and makes the floors all dirty with her shoes. (The doll is walked across the floor, leaving dirty footprints).' (p.238)*

The experimenter then asked two control questions to the children. The *Now* control question referring to the final situation; 'Is the floor dirty now?'; and the *Before* control question referring to the initial situation; 'Was the floor dirty before?' Finally the experimenter asked them the counterfactual test question: 'What if Carol had taken her shoes off – would the floor be dirty?' In order to answer that question the children have to imagine what the reality could be now if a particular event had happened (Carol took her shoes off).

Harris et al. (1996) used a variety of scenarios requiring some yes/no answers or some pointing. The findings showed that children were accurate when answering the counterfactual questions, even from the age of 3 for some of them.

They also showed that young children can understand what caused a particular consequence and how to prevent it. In the example above they demonstrated that they understood that if Carol had taken her shoes off the floor would not be dirty. In the second experiment included in their paper they found that children could differentiate between antecedents that could have prevented a particular outcome. In this second experiment they used a slightly different type of scenario. For example the children were told:

*One day the floor was nice and clean like that [experimenter points to a square of white plastic]. But guess what? Naughty Teddy comes along and paints the floor like this [experimenter makes Teddy, a hand puppet, paint the floor red].*

The children were then asked the *Before* control questions (Was the floor clean before) and the *Now* control question (Is the floor clean now?). Finally they were asked the two test questions requiring imagining another situation. They were asked the *preserve* test question (If Teddy hadn't painted the floor with his brush, would the floor be clean now?) with an antecedent that would preserve the initial state. And they were asked the *change* test question (If Teddy had painted the floor with his finger instead, would the floor be clean now?) with an antecedent that would change the initial state into the final state. The children, from the age of three, could distinguish between the antecedents the one that could block the observed outcome.

Following this work, Harris (2000) explained that in order to solve counterfactual scenarios children need to use their knowledge of causal relationships (for example: they need to know that if you paint the floor then it is dirty) but they also need to use their imagination to consider possible alternative outcomes if the antecedent had been different.

Guajardo and Turley-Ames (2004) also conducted some research on counterfactual thinking with young children. They studied the generation of counterfactual statements by children. They told short stories to the children as follows:

*'Imagine that you are playing outside in the muddy yard. You are thirsty so you go inside to the kitchen to get a drink of juice. You walk through the mud, you step over the doormat, and you keep your shoes on. Because your shoes are muddy, you get dirt all over the floor.'* (p.59)

After being told this story the children were asked a question prompting them to imagine some different antecedents: 'What could you have done so that the kitchen floor would not have gotten dirty?' And after each response they were asked 'Can you think of anything else?' to prompt them to provide as many different answers as possible.

A different type of story (similar to the Post Office scenario designed by Riggs et al., 1998) was used to help the children to think about a different set of consequences. One of the stories was as follows:

*'Peter is in his house, but Peter is not feeling very well. So he goes to bed. The phone rings and the man from the Post Office asks Peter to come and help put out the fire. Peter gets out of bed and goes to the Post Office.'*

The children were then asked a question to prompt them to give some alternative consequence; 'If there had not been a fire, where would Peter be?'

In their studies Guajardo and Turley-Ames (2004) showed that children as young as 3 could generate counterfactual statements when prompted to do so which was slightly earlier than what had been shown in previous research as, for example, Kuczaj and Day (1979) found that children could not produce counterfactual statements until the age of 4 or 5. In Guajardo and Turley-Ames'



(2004) study children could generate upward and downward counterfactual statements equally well, just as adults do. They produced additive counterfactual statements but very few subtractive statements which was also consistent with adults' performance. For an additive counterfactual an element was added to the antecedent and for a subtractive counterfactual an element was removed. An additive counterfactual could be 'If I had studied more I would have passed my exam', 'studying more' being the added element. A subtractive counterfactual would be 'If I hadn't gone on holiday before the exam, I would have passed my exam', 'going on holiday' being the element removed from the antecedent.

Guajardo and Turley-Ames (2004) found that performance on counterfactual reasoning tasks was related to language comprehension. They found a significant positive correlation between these two measures. This result is not surprising as all the scenarios they used involved listening to the short story. The children needed to be able to understand the story in the first place before they could reason counterfactually correctly.

As with adults, research has shown that the nature of the outcome influenced when children use counterfactual statements. Children are more likely to use counterfactuals when the outcome is negative or unpleasant. Harris (2000) reported one of these previous experiments where the children, aged 3 and 4, were tested. The children were presented with a character called Sally. Sally had a choice between two options: drawing with a pen or drawing with a pencil. There were two possible outcomes, a negative one, where Sally had inky fingers, and a positive one where Sally had clean fingers. Harris (2000) reported that children generated more counterfactuals in the first situation. These findings were also confirmed by German (1999). German (1999) presented 5-year-old children with a series of 'decisive choice' scenarios which had either a positive or a negative outcome. In one of the scenarios, for example, a little girl called Sally was going on

a trip. Her mother gave her the choice between white chocolate or a cheese sandwich for her lunch. In the negative outcome scenario Sally chose the white chocolate and ended up being hungry whereas in the positive outcome scenario Sally chose the cheese sandwich and ended up being full. A series of scenarios including 'irrelevant choices' were also presented. In one of those, for example, Sally was given two choices that both lead to either a negative or a positive outcome. In the first case, she was given the choice between brown or white chocolate and always ended up hungry, in the second case, she was given the choice between cheese or ham sandwich and always ended up being full. The children were then asked to explain the outcome of the stories. The results showed that children were more likely to generate counterfactual statement with the negative outcome of the decisive choice scenario.

In conclusion, it seems that children are more likely to generate counterfactual statements with negative outcomes and they will try to think of ways to avoid this negative outcome from happening in the future. The research has also shown that they have this ability from an early age.

### **3.4 Counterfactual thinking and the false belief task**

As previously discussed, children under the age of 4 do not seem to be able to pass the false belief task (see Wellman, Cross and Watson, 2000, for a review). When tested with a standard false belief task involving an unexpected transfer of an object they tend to make realist errors by pointing to the actual object location rather than where it was before, where the character should think it is.

Some research has looked at potential reasons why children could not pass this task, especially as at the same age they show signs of having a theory of mind, of being able to understand other people's mental states. The domain general view of theory of mind would suggest that general abilities are involved

and therefore a link might be found between theory of mind and executive functions, such as for example, reasoning abilities therefore some researcher studied the link between counterfactual reasoning and theory of mind. This link can be further supported by the representational mind view (Perner, 1991) as theory of mind and counterfactual reasoning require children to compare different models of the world.

Riggs et al. (1998) argued that in order to pass the false belief task children need to disengage from their knowledge of the situation and infer or recall a counterfactual situation. The children need to ignore the situation in front of them and engage in some counterfactual thinking. With the classic 'Sally and Anne' false belief task described earlier, the children need to ignore the fact that the marble is in the box now, and need to think of the following implicit counterfactual question: 'If Anne had not moved the marble where would it be now?' The link between counterfactual thinking and false belief tasks was initially investigated by Riggs et al. (1998) through three different experiments. In the first two experiments, they presented the children with some scenarios acted out with puppets and props. Each scenario was followed by either a theory of mind question or a counterfactual question. One of them was called the Post Office story (as described above). The results showed a significant positive correlation between these tasks, even when controlling for verbal abilities. For their third experiment, they created new scenarios not requiring the children to follow a long narration, and still found this correlation. One of these tasks was called the Shape Task. The children were presented with a puppet called Donald. They were told that Donald brought a few things to nursery, including some dough. They were shown that the dough was in the shape of a ball. The experimenter told them that Donald was going to play with it later but that he was going first into his/her bag to go to sleep. The experimenter carried on saying that that they could play with the

dough and roll it into a sausage shape. At this point the child was asked a prompt question: 'Does Donald know that we've played with the dough?', and was corrected if he/she gave the wrong answer. Finally the child was asked one of the test questions. The counterfactual question was: 'If I had not played with the dough, what shape would it be?' and the false belief question was: 'What shape does Donald think the dough is?'.

Riggs et al. (1998) argued that this consistent correlation between the false belief task and counterfactual reasoning suggested that these tasks required similar abilities. They suggested that the false belief task required children to solve a counterfactual situation. The implicit counterfactual question would be for example, 'If there had been no fire, where would Peter be?'.

This finding seems to be robust as it has been confirmed by other researchers with a variety of groups of children. Peterson and Bowler (2000) tested children with autism and two other groups as a control: children with severe learning disability and typically developing children. Both groups of children with special needs were matched on chronological age and verbal mental age. They used the cake scenario and the fire scenario previously designed by Riggs and al. (1998). Grant, Riggs and Boucher (2004) also tested children with autism and found the same results. These results confirmed Riggs' et al. (1998), as they found a correlation between performance on the false belief task and counterfactual reasoning tasks for the three groups.

As well as finding a correlation between performance on these tasks the pattern of acquisition seems to be regular: children seem to be able to reason counterfactually before they can pass the false belief task, and this seems to be the case with typically developing children (Riggs et al., 1998; Grant et al. 2004); children with severe learning disability (Peterson and Bowler, 2004); children with autism (Grant et al., 2004; Peterson and Bowler, 2004). However, Peterson and

Bowler did not find the same pattern with the group of typically developing children they used as a control in their study. Typically developing children performed the same on the false belief task and the counterfactual tasks. Children with autism and children with severe learning disability performed worse on the false belief task than the counterfactual task. Children with autism and children with severe learning disability found the false belief task more difficult than the counterfactual task, but this was not the case for the typically developing children. Peterson and Bowler (2000) therefore argued that counterfactual tasks make fewer demands than false belief tasks. They explained that the false belief task is driven by at least two components. The first component is necessary to pass the false belief task but not the counterfactual task, the second component is the ability to reason counterfactually and does not develop until the age of 4. Only the first component, allowing children to understand false beliefs, is deficient in autism. Children with autism can pass counterfactual tasks because the counterfactual proposition is explicitly provided but they cannot pass the false belief task because the counterfactual element is implicit and needs to be generated. This finding could suggest that children with autism find this task difficult because of impaired spontaneous generativity. These results seem to suggest that being able to reason counterfactually is necessary but not sufficient to pass the false belief task

The studies mentioned so far tested counterfactual thinking by presenting short scenarios followed by questions to the children. Guajardo and Turley-Ames (2004) studied the relation between the false belief task and counterfactuals by looking at the generation of counterfactual statements. Their study included upward and downward counterfactual statements (as defined by Roese above) and also made the distinction between antecedent counterfactuals, where the children were asked to modify the antecedent, and consequent counterfactuals, where the children are asked to modify the consequent. In the antecedent task the

children were told a story where they came into the house with dirty shoes on and consequently made the floor dirty. They were then asked the test question: 'What could you have done so that the kitchen floor would not have got dirty?' The consequent counterfactual tasks were based on Riggs et al.'s (1998) work. For example in one of the scenarios the children were told that Peter was not very well and went to bed. The man from the post office rang him to ask him to come to put out a fire. So Peter got out of bed and went to the post office. The test question was: 'If there had been no fire, where would Peter be?' The results from this study confirmed Riggs et al.'s (1998) finding by reporting a correlation between theory of mind tasks and consequence counterfactual thinking, but extended it by finding this same relationship with antecedent counterfactual thinking, with antecedent counterfactuals accounting for 16% of the variance in the scores obtained for the theory of mind tasks.

Most of the research on counterfactual thinking and theory of mind tasks has found a relationship between these tasks. Peterson and Riggs (1999) studied these correlations in the light of the adaptative modelling theory. Adaptative modelling allows us to obtain knowledge about inaccessible information by temporarily modifying our database regarding a particular object. The process of modified derivation does not require an up-date of our database of knowledge but some modification in the process of answer-derivation we use to solve a problem (Peterson and Riggs, 1999). Through this process we can ignore facts already present in our database, or add some new facts to it. This modified derivation process supports counterfactual reasoning (Peterson and Riggs, 1999). Peterson and Riggs (1999) argued that this same process is necessary to solve a false belief task, making reference in particular to the Maxi task first designed by Wimmer and Perner (1983). They explained that in order to correctly answer the false belief task, children need to identify the element that Maxi does not know,

implement a modification-instruction that is to ignore this fact (because Maxi ignores it too), ask themselves the test question and attribute the answer to Maxi. They argued that the false belief task requires two types of abilities: having a theory of mind and a mechanism of simulation (i.e. the modified derivation process). The difficulty that children have when they try to apply the adaptive modelling process is that they do not have sufficient cognitive control to deal with conflictive cases. Conflictive cases are present in counterfactual scenarios when the premise conflicts with the real world, and in the false belief task where the protagonist's beliefs are not consistent with the real world and with the child's own beliefs. Peterson and Riggs (1999) concluded that this difficulty with modified derivation is what underlies difficulties with counterfactual reasoning and false belief tasks, especially when theory of mind has already been acquired by the children. They argued that this conclusion is applicable to standard false belief tasks, deceptive box tasks (e.g. : Smarties tasks), and other simpler reasoning problems.

Although the correlation between the false belief task and counterfactual thinking seems often present it is not always the case. Perner, Sprung and Steingogler (2004) conducted experiments looking at this relationship. They used similar scenarios followed by either a counterfactual or a theory of mind question, and varied some parameters, especially the complexity of the scenario. In both the experiments they reported a correlation between the false belief task and counterfactual thinking in their first data analysis but this correlation became non-significant when partialling out age and verbal mental age. In the first analysis it appeared that performance on the false belief task correlated differently with the simple counterfactual scenario and the complex counterfactual scenario. Following these findings Perner et al. (2004) argued for a dissociation between the false belief task and counterfactual thinking.

Perner explained why their results differed from Riggs's regarding the correlation between counterfactuals and false belief tasks. All the counterfactual questions Perner et al. (2004) used had false antecedents that would therefore require a modification of the known database and the application of real world knowledge to deduce the consequence of that modification (*modified derivation*). In the simple problem derivation could proceed without recourse to the actual sequence of events, whereas for the difficult problems the derivation needed to take into account at least some part of the actual sequence of events. With the complex scenarios the children needed to relate the derivation to actual events which is a similarity these scenarios have with the false belief tasks. They therefore rejected the simulation view to explain their results.

## **4 Conclusion**

Having a theory of mind is the ability to understand and attribute mental states to others. It is essential for efficient communication in everyday life. Theory of mind is often assessed using the false belief task that requires imputing false beliefs to a character in a short story or scenario. Research has shown that this task is very reliable: in most cases children can pass the task from the age of 4. This result is very robust as it does not change with the type of task or the country where the research was conducted. The false belief task has been studied in the light of other domain general cognitive abilities it might require, such as inhibition, memory and language, and in particular in relation to counterfactual reasoning. A variety of research papers have mainly shown that there is a consistent strong correlation between the false belief tasks and counterfactual reasoning. Some research suggests that counterfactual thinking is a prerequisite to pass the false belief task. In order to pass each task the children need to be able to suppress their knowledge of current reality, and imagine what reality could be if a past event



had been modified, or what reality is thought to be for somebody who did not witness a particular past event.

To some extent contrary-to-fact syllogisms also require the suppression of knowledge of current reality. Children need to reason in a pretend world and ignore their empirical knowledge.

This thesis will investigate the relationship between these three types of task (false belief tasks, counterfactual scenarios and contrary-to-fact syllogisms) as they all involve suppressing knowledge of current reality, they all require children to manipulate their representations of the world.

Before studying the relationship between these tasks, it is important to better understand the nature of counterfactual thinking, which is going to be the focus of the second chapter. The first experiment will investigate whether there is something special about counterfactual thinking, especially in comparison to future hypothetical thinking.

## **CHAPTER 2**

### **EXPERIMENT 1: AFFIRMATIVE AND NEGATIVE COUNTERFACTUAL QUESTIONS**

The aim of the first experiment of this thesis was to investigate whether there is a difference in difficulty between future hypothetical reasoning and counterfactual reasoning as claimed by Riggs et al. (1998). This chapter will review the literature related to this area, and will then present the study and its findings.

#### **1 Introduction**

Counterfactual thinking has been defined and discussed in the previous chapter. As a reminder, a counterfactual is an alternative to the past (Roese, 1997). In order to reason counterfactually children need to negate or change a past event and imagine that, as a consequence, a present situation would be different. Counterfactual thinking has been studied extensively with adults (see chapter 1) whereas the research with children is more limited. As we have seen counterfactual thinking in developmental studies has mainly been investigated in relation to the false belief task which is a task used to evaluate theory of mind in children (e.g: Riggs et al. 1998; Grant et al., 2004; Guajardo and Turley-Ames, 2004). Children are said to have a theory of mind when they are able to understand other people's mind and mental states (Premack and Woodruff, 1978). The false belief task has been described in depth in the first chapter. In order to pass the false belief task children need to acknowledge that another character did not witness the change of state of affairs they themselves witnessed. They also need to recognise that people will act according to their own beliefs even though they are false.

The components of the false belief task have been investigated, especially regarding a possible embedded counterfactual element. Riggs et al. (1998) tested the hypothesis that performances on counterfactual tasks and false belief tasks are correlated. They ran a series of experiments and found a consistent correlation between these two types of tasks. More details regarding these tasks were given in chapter one. The correlation they found was independent of verbal ability, chronological age, mental age and narrative aspect of the tasks. They therefore argued that there is an implicit counterfactual element embedded in the false belief task, and that children need to be able to reason counterfactually in order to pass the false belief task. They claimed that passing the false belief task requires reporting a situation that is counter to the current reality, as in a counterfactual task. The main difference between both types of tasks is that in the false belief task the counterfactual situation is held as a belief but not in the strictly counterfactual scenario where no mental states are involved.

Peterson and Riggs (1999) investigated the underlying mental simulations of counterfactual thinking and the false belief task. They described very similar processes both requiring children to ignore instructions. For example in the false belief task, in this case the 'Maxi task', the children had to ignore the fact that the chocolate had been moved. However, it can be questioned whether this correlation found between performance on counterfactual tasks and performance on false belief tasks is due to the counterfactual element of the counterfactual task itself (i.e.: changing something that had already happened) or to the hypothetical nature of this task. Is there something special about counterfactual thinking? Riggs et al. (1998) continued their work by investigating the difference between counterfactual and future hypothetical thinking to ensure that the correlation between performance on counterfactual tasks and false belief tasks was due to the counterfactual element rather than just the conditional element included in the

counterfactual tasks. Both counterfactual and future hypothetical thinking involve reasoning conditionally but they are not completely identical. Counterfactual thinking requires children to suppress their knowledge of current reality and leads to a situation, different from reality, and meant to stand in for reality, whereas future hypothetical thinking requires children to imagine the world in the future and leads to a situation different from reality but not meant to stand in for reality (Riggs et al., 1998).

Riggs et al. (1998) designed an experiment to investigate whether children make the same number of realist errors in both of these tasks. They designed two tasks, the modelling clay and the drawing tasks. Both were followed by either a counterfactual question or a future hypothetical question. The children were presented successively with both scenarios and were asked to answer one counterfactual question and one future hypothetical question. In the drawing task, for example, the experimenter showed the children a first box containing items with a picture on and a second box containing items with no picture on. Tidy Teddy was introduced to the children, and then he put a blank piece of paper in the second box before going to sleep in the experimenter's bag. In the counterfactual version, the experimenter took the piece of paper out of the second box and drew a picture on it. The children were then asked to classify the piece of paper by physically putting it in the correct box. After the children had put the piece of paper in the correct box they were asked the following counterfactual test question: 'If I had not drawn on this piece of paper, which box would it be in now?' In the future hypothetical version the experimenter still removed the blank piece of paper from the second box, but then asked the following future hypothetical test question straight away: 'If I draw on this piece of paper, which box will it go into?' The modelling clay involved changing the shape of the clay and was very similar. The modelling clay could be rolled into a ball shape or a sausage shape and was

classified into boxes according to that shape. The counterfactual test question was: 'If I had not played with the dough which box would it be in?', and the future hypothetical test question: 'If I roll the dough into a sausage shape, which box will it go into?' In both scenarios the test question was followed by a memory question asking the children to recall where the piece of paper or the piece of clay was at the beginning.

When comparing these two types of reasoning Riggs et al. (1998) found that future hypothetical questions were easier than counterfactual questions, and they argued that children more easily avoid making realist errors when they are asked to reason conditionally and report a future situation than a counterfactual situation. These results suggested that these two types of tasks are intrinsically different.

Evans and Over (2004) have claimed that although some counterfactuals can be very different to future hypothetical conditionals, most of them are very similar. They explained that, for example, the counterfactual 'if I had taken the pawn, then I would have won the game' is very similar to the equivalent future hypothetical 'If I take that pawn, then I will win the game.' They argued that people process those two conditionals in the same way and make the same judgement of probability. That is to say that they imagine the antecedent and evaluate the likelihood of the outcome in the context of this antecedent. They argued that modifying the past or modifying the future requires the same cognitive processes. This claim was supported by Over, Hadjichristidis, Evans, Handley and Sloman (unpublished). In their study Over et al. (unpublished) presented adults with four conditional statements including some negated ones (e.g.: The cost of petrol increases and traffic congestion does not improve). The participants were asked to rate the probability that the conditional statements were true or false. The results showed that counterfactuals were judged in a very similar way to hypotheticals. Following these claims from Evans and Over (2004) and Over et al. (unpublished) the

question as to why there was a difference between counterfactuals and future hypotheticals in Riggs and al.'s (1998) experiment arises.

An alternative interpretation for Riggs et al's (1998) results can be suggested. In their experiment all the counterfactual questions asked were negative questions (e.g.: 'If I had not played with the dough which box would it be in?') and all the future hypothetical were affirmative questions (e.g.: 'If I roll the dough into a sausage shape, which box will it go into?'). This systematic confound could have influenced the results.

The effect of negation has been studied in relation to language development. The literature on children's language development shows that children produce affirmative sentences before they can produce negative ones, and on the comprehension side negative forms are usually more difficult to understand than affirmative forms (Tager-Flusberg, 1985). Therefore it can be expected that the negative component of the counterfactual question will on its own increase the difficulty of this question. Children might find counterfactual questions in Riggs et al.'s (1998) experiment more difficult just because they are negated.

The effect of negation on reasoning with adults has been studied experimentally. Oaksford and Stenning (1992) used the negation paradigm (if p then q, if p then not q, if not p then q, if not p then not q) to systematically study the effect of negation on reasoning. It appears that the use of negations changes the way participants reason. Three biases have been identified: the negative conclusion bias, the affirmative premise bias and the matching bias (Evans and Handley, 1999). The negative conclusion bias is the tendency to accept more conclusions that are formulated with a negative sentence rather than an affirmative one (Evans and Handley, 1999). Pollard and Evans (1980) explain this as the caution heuristic: accepting a negative answer leaves more options open than accepting an affirmative answer. The affirmative premise bias is the tendency to

accept more inferences when the minor premise is affirmative rather than negative (Evans and Handley, 1999). Finally the matching bias, the one that we are going to discuss in more depth, is the tendency to choose cases when the lexical content matches the initial conditional proposition; it seems that mismatching cases are seen as logically irrelevant by reasoners (Evans, 1998). In other words, the matching bias is the tendency to focus on the proposition that is denied. Matching bias is very robust. It has been shown with three different types of conditional sentence when applying the negation paradigms to them (if (not) p then (not) q; (not) p only if (not) q; (not) q if (not) p (See Evans 1998 for a review). Matching bias is robust with abstract materials but tends to disappear with realistic material (Evans, 1998). Evans, Clibbens and Rood (1996) studied the difference between implicit and explicit negations in relation to matching bias using three types of tasks: a truth table evaluation task, the Wason selection task and another selection task. They found that the matching bias was present when implicit negations were used and disappeared when explicit negations were used.

The matching bias suggests, therefore, that when children are presented with a negative counterfactual question they are going to concentrate on the affirmative lexical content of the question. For example, in the questions 'what if I hadn't drawn on this piece of paper, where would it be' from Riggs' et al. (1998) the children are more likely to understand it as 'what if I had drawn on this piece of paper, where would it be.' This second question is not a counterfactual question as in this scenario the experimenter did draw on the piece of paper just before asking the question.

These biases demonstrate that negative forms influence reasoning performance. Introducing a negation in a premise or in a conclusion is not without consequences. When presented with negated premises reasoners tend to make more mistakes than when presented with affirmative premises.

All of these findings show that processing negated premises is more difficult than processing affirmative premises. This suggests that the difference Riggs et al. (1998) found between counterfactual and future hypothetical questions could be due to the language used in the question rather than the different requirements of these two types of reasoning.

The first aim of the present experiment was to test this hypothesis by assessing whether counterfactual and future hypothetical questions are of the same difficulty. This was done by running experiments similar to the ones run by Riggs et al. (1998) but controlling for the form of the question, i.e. controlling for positive and negative questions.

Riggs et al. (1998) also explained the correlation between the false belief task and counterfactual thinking by saying that children make realist errors because they cannot detach themselves from the current reality in front of them. Answering a counterfactual question correctly and passing the false belief task implies that children are able to suspend their knowledge of current reality. A realist error, in a false belief task or in a counterfactual scenario, occurs when the child chooses the answer that is consistent with reality rather than the correct answer. For example in the Post Office story (as described in chapter one) in the counterfactual and the false belief version the child should point to the house when in reality Peter is at the post office. Riggs et al. (1998) claimed that children made realist errors when they pointed to the post office because Peter really was at the post office. However, the children were offered only two possible locations, therefore two possible answers, by the scenario: Peter could be either in the house or at the post office. The correct answer being 'at home', 'at the post office' is the wrong answer but also a realist error. It is impossible to judge whether an error occurred because children were unable to dissociate from reality or for some other reason. It does not seem appropriate to say that children made realist errors because they



could not ignore reality when no other alternative was offered in the scenario. The second aim of the present experiment was therefore to analyse the mistakes children make and assess whether these mistakes are always realist mistakes or not. In the present experiment the children were presented with three locations to point to, the correct one and two incorrect ones, one being the reality location and the other one an irrelevant location. This design should allow us to understand whether the difficulty is due to a difficulty with suppressing reality or other factors such as lack of attention, or lack of understanding for example.

## **2 Method**

### **2.1 Participants**

Seventy-nine children were tested for this experiment. The sample included 40 girls and 39 boys. The children were aged between 3;4 and 5;5 (Mean = 4;6; Standard Deviation = 7 months).

The participants were recruited in two different schools in Plymouth UK, with each school having a nursery class and two reception classes. Thirty two children were from the nursery classes and 47 from reception classes.

### **2.2 Material**

The experiment consisted in classifying objects according to their particular properties and answering a variety of conditional questions. The children were given three similar boxes to classify the objects. The objects to classify were some play dough with different shapes (ball, sausage, squashed), sheets of paper with either a square, a circle or nothing drawn on them, cards with either an orange patch, a green patch or nothing drawn on them, and toy cars with either a round sticker, a square sticker or no sticker on the roof.

Reminder cards were placed in front of each box so the children could see how the objects were classified without having to remember the classification rule.

A recording sheet was used to note the responses given by the children and to give a set procedure to the experimenter (Appendix A).

As an example of one of the scenarios the 'Cars and stickers' scenario was as follows. Toy cars had either a round sticker, a square sticker or did not have a sticker stuck on the roof. Three empty identical boxes were placed in front of the child. The child was then told that all the cars with a round sticker would go in the first box, cars with a square sticker in the second box and cars with no sticker in the third box. Reminder cards were placed in front of each box so the child could see how to sort out the cars rather than remember the rule.

The child was asked to classify cars in the correct boxes according to the stickers, and was asked test questions such as 'What is on this car?' (a round sticker) to check the required vocabulary.

In the counterfactual version, the experimenter took a car with no sticker on, put a round sticker on it and asked the child: 'What's on this car now?' followed by 'Where shall we put it now?' The child then put the toy car in the correct box and the experimenter asked one of the following critical questions; either the affirmative counterfactual question: 'If I had put a square sticker on this car, which box would it be in now?' or the negative counterfactual question: 'If I had not put a sticker on this car, which box would it be in now?'

In the future hypothetical version, the experiment took a car with no sticker, did not change the properties of the car and asked the child to put the car in the correct box. The child was then asked one of the following critical questions; either the affirmative future hypothetical question 'If I put a round sticker on this car,

which box will it go into?' or the negative hypothetical question 'If I do not put a sticker on this car which box will it go into?'

The other scenarios are reported in more details in Appendix B.

The experimenter recorded which box the child pointed to, and coded it as a right or wrong answer.

### **2.3 Design**

Each child was presented with four scenarios: one involving toy cars and stickers, one cards and colours, another one playdough and finally one including drawings. Each scenario was followed by an affirmative counterfactual question, a negative counterfactual question, an affirmative future hypothetical question or a negative future hypothetical question. It was a within participant design. The children had to answer one of each of the four different types of questions.

The order of the scenarios and the questions was counterbalanced between children using a Latin-square.

### **2.4 Procedure**

The experiment was run individually over one session lasting about 10 minutes. The experiment took place in the children's school during school hours.

A pilot study was run early on to evaluate the feasibility of the task and the time required to run the task. A girl, age 4; 1, took part in the pilot study. She was presented with the four scenarios, and she was asked the four test questions and the control questions. The participant answered all the control questions correctly and always classified the object in the appropriate box. She correctly answered three of the control questions and incorrectly only one of them. The tasks took less than 10 minutes to run. The participant engaged in the task and seemed to enjoy

it. Following the pilot study it was decided that the difficulty level and the length of time of the tasks were suitable, therefore no modification was made to the task.

### 3 Results

Altogether 18 children failed to cope with the task demands: they could not do the pre-requested classifications, answer at least one control question, or understand the questions. These children were excluded from the sample. As a consequence the final sample consisted of 61 children (29 boys and 32 girls) aged between 3;7 and 5;5 ( $M = 4;8$ ,  $SD = 6.2$  m). It is also worth noting that all the children who were less than 3 years and 6 months in age were excluded from the study because they could not cope with its requirements.

Out of the 61 participants 25 answered four questions correctly, 21 three questions, 10 two questions and 5 only one question. No participant answered all the four questions incorrectly.

#### 3.1 Comparison of the different types of questions

Table 1: Number of right answers for each type of question (n=61).

	Affirmative	Negative
Counterfactual	48	36
Future Hypothetical	48	57

The overall effects of negation/affirmation, types of reasoning and the interaction between those two factors were analysed. In order to study the overall effect of negation/affirmation the total of correct answers for affirmative counterfactual questions and for affirmative future hypothetical questions was computed as well as the total of correct answers for negative counterfactual questions and negative future hypothetical questions. The results are shown in the following table.

Table 2: Total number and percentage of correct answers (n=122).

	Affirmative Total	Negative Total	CF Total	FH Total
Number of correct answers	96	93	84	105
Percentage of correct answers	78.69	76.23	68.85	86.05

A series of sign tests were run and the first one showed that the difference between the total number of correct answers to the affirmative questions and the negative questions was not statistically significant ( $z=-0.4$ ,  $p>.05$ , two tailed).

The second one showed that the difference between the total number of correct answers to the counterfactual questions and the future hypothetical questions was statistically significant ( $z= -2.97$ ,  $p<.05$ , two tailed).

Finally the third sign test showed that the interaction between affirmative/negative and the type of reasoning (counterfactual/future hypothetical) was significant ( $z=-2.91$ ,  $p<.05$ , two tailed).

In their experiment Riggs et al. (1998) compared a negative counterfactual question and an affirmative future hypothetical question. The same comparison in this experiment showed that children's performance on the affirmative future hypothetical was better than their performance on the negative counterfactual. A sign test showed this difference to be statistically significant ( $z = -3.13$ ,  $p<.05$ ).

In order to control for the effect of negation on the different types of questions, affirmative counterfactuals and affirmative future hypotheticals should be compared, as well as negative counterfactuals and negative future hypotheticals. The results showed no difference in performance between the affirmative

counterfactual and the affirmative future hypothetical as the total of correct answers was identical.

The negative future hypothetical showed a ceiling effect as it was answered correctly by most participants. It appeared that this question was actually confounded. This question was, for example in the car scenario, 'If I do not put a sticker on this car which box will it go into?' The child was asked not to change anything of the properties so the car should go in the same box as it would go into if it was just classified. Children can answer this question correctly even if they do not process the question and just classify the car.

### 3.2 Analysis of mistakes

Having three boxes allows the analysis of the mistakes the children make as they can either point to the correct box, or one of the two incorrect boxes. One of these boxes will relate to a realist error (the child points to the box where the object is in reality) the other is classified as 'other type of errors.'

Table 3: Number and percentage of types of errors for each type of question.

Type of question	Realist errors	Other errors	Total
Counterfactual affirmative	9 (69%)	4 (30%)	13
Counterfactual negative	16 (64%)	9 (36%)	25
Future hypothetical affirmative	9 (69%)	4 (30%)	13
Total	34 (67%)	17 (33%)	51

Overall about one third of the mistakes made by the children are not realist errors. When children made an error, it was not systematically a realist error. The

pattern of errors was similar for the affirmative counterfactual questions and the affirmative future hypothetical questions.

For the affirmative counterfactual question, out of the four children who made 'other errors' two pointed to the wrong box, but not the reality box, one gave an irrelevant verbal answer, and one said he did not know the answer.

For the negative counterfactual nine children made 'other errors'. Five of them pointed to the wrong box (but not the reality box), the other four gave a wrong verbal answer: 'yes', 'I don't know', 'on the table', 'out of the box'.

For the affirmative future hypothetical the four children giving an incorrect answer pointed to the wrong box, but did not make any other sort of errors.

For the negative future hypothetical all the mistakes made were 'other errors'; three pointed to the wrong box and one to the playdough box. It should be noted that it is not possible to make a realist error as the 'reality box' is the right answer.

The overall results show that most of the errors, but not all of them, are realist errors. Children make the same sort of errors with future hypotheticals and counterfactual suggesting the two types of questions involve similar mental processes.

## **4 Discussion**

This experiment had two main aims. The first aim was to assess whether counterfactual questions and future hypothetical questions are of the same difficulty, and the second aim was to analyse the mistakes children make when answering those types of questions.

Following their experimental work Riggs et al. (1998) suggested that difficulties with counterfactuals were not just a general difficulty with conditionals, but that there was something special about counterfactuals. They suggested that

counterfactual questions were more difficult to answer than future hypothetical questions as children's performance was higher on the latter one. When analysing Riggs et al.'s (1998) questions in more depth it can be identified that their counterfactual questions, in that particular paper, were systematically negative and their future hypothetical questions affirmative. In effect their experiment showed that negative counterfactual questions are more difficult for children than affirmative future hypothetical questions. The results of the present experiment confirmed that finding: children made more mistakes when answering negative counterfactual questions than when answering affirmative future hypothetical questions. However, these results do not allow us to compare counterfactual thinking and future hypothetical thinking as such. In order to do that, more control is necessary.

When controlling for negation the children's performance followed a different pattern. This experiment showed no difference between performance on the affirmative counterfactual tasks and the affirmative future hypothetical tasks, but it showed that negative counterfactual tasks seem more difficult than affirmative counterfactual tasks. It therefore seems that there is no difference between counterfactual and future hypothetical thinking, and that negation increases the difficulty of the question. Consequently this suggests that the difference found by Riggs et al. (1998) was probably due to the formulation of the question, the negative question being more difficult to answer for the children. This first result is confirmed by the fact that, when looking only at the counterfactual questions, the children performed better on the affirmative counterfactual questions than the negative counterfactual questions. This result tends to confirm the reasoning research conducted with adults which states that introducing a negation in a reasoning problem influences the performance (eg: Evans and Handley, 1999).



Riggs et al. (1998) argued that the main difference between counterfactual and future hypothetical reasoning is that in the first case the new world children need to imagine stands for reality but not in the second case, and for that reason counterfactual thinking is more difficult than future hypothetical thinking. In a counterfactual situation the children are required to imagine that a past event was altered and as a consequence the present situation would be different. With a future hypothetical situation children just need to imagine that a future situation could be different, the present reality in front of them does not need to be altered. Whether this new world is meant, or not, to stand in for reality does not make a difference to the children's performance. This probably means that there is a general difficulty intrinsic to reasoning conditionally.

Looking more closely at the difference between the affirmative and the negative question fundamental differences can be identified. The two following questions can be compared: 'If I had put a square sticker on this car, which box would it be in now?' vs 'If I had not put a sticker on this car, which box would it be in now?' The first proposition (affirmative) asserts an action: putting a square sticker on the toy car. The situation is made clear to the children: they know exactly what could have been done. The negative proposition on the other hand eliminates only one of the actions the experimenter could have carried out. The experimenter did not put a sticker on the car, but he, or she, could have carried out another action on that car. He, or she, could have put no sticker at all, or put a round one instead, or maybe done something else. The negated questions can be more confusing for the children as they make several alternative possibilities available.

One interesting effect to emerge was that the polarity (i.e.: affirmative versus negative) variable influenced future and counterfactual problems in different ways. Let's consider why that might be the case. The negative future hypothetical

question presents some interesting features as the children performed close to ceiling effect on that question, with 57 out of 61 giving the correct answer. Its validity can be questioned as it seems that children do not need to process the question completely to give the right answer. For example, in the car situation, the experimenter took a car with no sticker, did not change the properties of this car and asked the child: 'If I don't put a sticker on this car which box will it go into?' The child could give the right answer just by classifying the car in the box, without processing the question. This question does not seem to fit with the definition of a future hypothetical given by Riggs et al. (1998) where they say that future hypothetical thinking requires children to imagine a world in the future that leads to a different situation from reality. Therefore this question cannot discriminate between children who answer correctly for the right reasons (i.e.: process the conditional) and those who base their response on current reality. It seems extremely difficult in the present experimental condition to design a negative future hypothetical question that in effect says that something will not be done to this car but would still change the future.

Another part of this experiment aimed to study the type of mistakes children made when incorrectly answering the questions. Overall 61% of children who made a mistake made a realist mistake, i.e. they pointed to the location where the target object really was. This result was quite constant across tasks (excluding the future hypothetical questions for the reasons explained above). This tends to confirm Riggs et al.'s (1998) findings that children make mistakes regarding counterfactuals because they have difficulties suppressing their current knowledge of reality but not completely. Riggs et al. (1998) explained that children made mistakes because they could not suppress current reality but the present study showed this was not the only reason why children make mistakes. Nearly 40% of the children made other types of errors. The results do not allow us to analyse

precisely what these errors were but it can be hypothesised that they could be due to a lack of understanding of the questions or a lack of attention for example.

In conclusion, it appears that when controlling for polarity there seems to be no difference in performance between counterfactual and future hypothetical tasks, but introducing a negation increases the task difficulty. There does not seem to be something special about counterfactual thinking in comparison to future hypothetical. The difficulty children have when answering counterfactual or future hypothetical questions seems to be partly due to the fact that children need to suppress their current knowledge of reality, but not only this. Other factors influence the children's responses. The conditional element itself is probably a source of difficulty.

The next experiment described in chapter 3 aimed to replicate these results to verify their reliability. Affirmative counterfactual questions, negative counterfactual questions and affirmative future hypothetical questions were studied. The aim of studying counterfactual reasoning in isolation was to get a better understanding of how children manipulate their representations of the world and then link it with the false belief task. The following experiment studied the relationship between counterfactual thinking and false belief task.

# **CHAPTER 3**

## **EXPERIMENT 2: AFFIRMATIVE AND NEGATIVE COUNTERFACTUAL QUESTIONS AND THE FALSE BELIEF TASK**

### **1 Introduction**

The previous chapter looked at the difference between counterfactual thinking and future hypothetical thinking, and at the influence of negation on these two types of reasoning. The results showed that when controlling for negation counterfactual thinking was not more difficult than future hypothetical reasoning. It seems that the difference between counterfactual and future hypothetical reasoning described by Riggs et al. (1998) was probably due to the presence of the negation in the counterfactual questions, and the general difficulty that children have with counterfactuals may be a general problem linked to hypothetical thinking. This distinction between affirmative and negative counterfactual questions has not been studied in depth in previous literature. Therefore to strengthen the findings of the previous experiment the difference between negative and affirmative counterfactual questions will be investigated further. If the findings can be replicated it will give more reliability to the previous experiment.

Counterfactual thinking has been studied in relation to theory of mind, and to the false belief task in particular (e.g.: Grant et al., 2004; Riggs et al, 1998; Guajardo and Turley-Ames, 2004). Having a theory of mind allows us to understand other people's minds and mental states, and predict their behaviour accordingly. The false belief task, as described in chapter 1, is used to assess whether a child has a theory of mind or not. This task is usually presented to children using toys and props. It requires children to understand that a protagonist did not witness a change in the state of affairs, and to ascribe mental states to this protagonist. When asked the false belief question (eg: Where will Sally look for her

marble? – See chapter 1 for a description of the Sally and Ann task) they have to understand that their knowledge of the world is different from Sally's, they need to be able to hold those two conflicting representations and decide which one is right for them and which one is right for Sally, but children tend to find coping with multiple conflicting representations difficult (Scholl and Leslie, 2001). They also need to have the knowledge that a person is going to act according to their own personal beliefs, even if their beliefs are wrong. Children tend to overestimate others' knowledge according to what they themselves know. Birch and Bloom (2004) argued that the 'curse of knowledge', i.e. the tendency to overestimate other's knowledge to match our own beliefs, can explain children's difficulties in passing the false belief task.

The false belief task is one of the most used tasks to assess theory of mind in young children. Through their meta-analysis Wellman et al. (2001) confirmed that there is a shift in ability around the age of 4, which is when children can pass the false belief task, although they show signs of having a theory of mind well before that age. This shift in ability has been studied in depth in the past few years. Many research projects have studied the components of the false belief task and seem to attribute this shift to a component other than theory of mind itself. Peterson and Riggs (1999) suggested that this component is the ability to deal with conflictive cases.

Bloom and German (2000) argued that children need more than just a theory of mind to pass the false belief task and also that having a theory of mind is more complex than just passing the false belief task. To them the false belief task should be used to investigate how children reason about representations: how they cope with two different representations, one corresponding to the real situation and one corresponding to a character's false belief.

Riggs et al. (1998) ran a series of experiments exploring the relationship between counterfactual tasks and false belief tasks. They found a consistent correlation between these two types of task, even when partialling out language skills and narrative factors. They therefore argued that in order to pass the false belief task children need to be able to reason counterfactually. This finding seems to be robust as it has been confirmed by other researchers (see chapter 1 for more details). However, when looking at Riggs et al.'s (1998) experiment in more depth it appears that they more precisely found a link between negative counterfactual questions and the false belief task as they only included negative counterfactual questions in their tasks. Therefore the issue that this experiment will address is the extent to which the relation between counterfactual thinking and the false belief task is the result of a difficulty with processing negatives.

Grant et al. (2004) also investigated the relationship between theory of mind and counterfactual reasoning with children with autism as they tend to have a delayed acquisition of theory of mind (Baron-Cohen, 2000). They used a variety of tasks including some standard false belief tasks, non-standard false belief tasks and some counterfactual tasks. The eight standard false belief tasks they ran were based on Riggs et al.'s (1998) study and included four narrative tasks (e.g.: the Post Office story) and four non-narrative stories (e.g.: The Shape Task). They explained that these tasks require children to engage in some conditional counterfactual reasoning, to have an understanding of belief, and to have an ability to infer counterfactual states of affairs from implicit information. The same scenarios were also presented with a counterfactual question. The non-standard counterfactual tasks' requirements were different from the previous tasks as the false belief was explicitly stated in the scenario, or the event underpinning this event was explicitly described, and they did not require any conditional counterfactual reasoning. Grant et al. (2004) used the task designed by Wellman

and Bartsch (1988). Each scenario was illustrated by a comic strip. In the Explicit False-Belief task, for example, the children were told that Mary wanted to find her kitten. Mary's kitten was really in the bedroom but Mary thought that it was in the kitchen. They were then asked the test question: 'Where will Mary look for her kitten?' For their study Grant et al. (2004) recruited children with autism and for the two control groups children with moderate intellectual disability and typically developing children. They found that the children with autism performed worse than the control groups on the false belief tasks and the counterfactual tasks, and they found a significant correlation between these tasks which was due neither to narration nor memory. They found that the children with autism performed as well as controls on the non standard false belief tasks, and that they performed worse on standard false belief tasks than counterfactual tasks. They concluded that children with autism did not pass the standard false belief task not because of a difficulty to understand false beliefs but because of the cognitive requirements of the task. When the false beliefs were explicit children with autism performed better. This suggests that children with autism have a difficulty with implicit false belief maybe because of their impairment in imagination as mentioned in Wing's triad (Happé, 1994a).

Peterson and Bowler (2000) asked whether children with autism fail the false belief task for the same reason as typically developing children, i.e. because of the implicit counterfactual element embedded in the tasks. They tested children with autism and two other groups as a control: children with severe learning disability and typically developing children. Both groups of children with special needs were matched on chronological age and verbal mental age. The children were presented with tasks previously used by Riggs et al. (1998), the cake scenario and the fire scenario. Each scenario was followed either by a counterfactual question (note: all the counterfactual questions were negative questions) or a false belief

question. They found no significant difference in performance on the counterfactual task between the three groups. They found a strong correlation between counterfactual tasks and false belief tasks for the typically developing children. Children with autism, but not children with severe learning disability, performed worse than typically developing children on false belief tasks

All these research papers looked at the relationship between passing the false belief task and solving a counterfactual problem. Guajardo and Turley-Ames (2004) studied the relationship between theory of mind and the ability to generate counterfactuals. They identified two types of counterfactuals: the 'consequent tasks' where the children are asked to alter a consequent and the 'antecedent tasks' where children are asked to alter the antecedent. Examples of their scenarios were described in chapter 1. They found a relationship between the performance on each type of counterfactual task and theory of mind tasks. Theory of mind correlated highly with antecedent counterfactual thinking and consequent counterfactual thinking, and all these tasks also showed some similar patterns of development.

It seems that a recurrent link between the false belief task and counterfactual thinking has been found in the literature, however, Perner et al. (2004) argued for a developmental dissociation between the false belief task and counterfactual reasoning. In their study, they presented children with false belief questions and counterfactual questions embedded in simple and complex scenarios. The tasks they used were the 'travel scenarios'. The complex (Feichtinger 1999 as cited in Perner et al. 2004) and the simple (Sprung 1999 as cited in Perner et al. 2004) scenarios both involved Peter who wanted to travel; different points of departure and arrival were available as well as different means of transport. In the simple scenario Peter could go to the boat house, the barn or the school. At the boat house he could take the boat to go to the lake, at the barn the coach to go to the



pasture and at the school the tricycle to go to the playground. In this scenario each departure was linked to only one means of transport and one destination. In the complex scenario Peter could go to the green station or the blue station. At the green station he could take the bus to go to the lake or the train to go to the mountain. At the blue station he could take the train to go to the lake or the bus to go to the mountain. In this scenario the correspondence departure/means of transport/destination was not simple. Taking all these variables into account increases the complexity of the task and requires a greater processing and memory capacity. The results showed that children found counterfactual tasks a lot easier to answer than false belief tasks, especially with the simple task. Perner et al. (2004) did not find any difference between the simple false belief tasks and the complex false belief tasks, but children performed better on simple counterfactual tasks and the complex false belief tasks. They explained that this dissociation was an argument against Peterson and Riggs' (1999) modified derivation argument described in chapter 1. Perner et al. (2004) claimed that if the common ability shared by the two tasks was the ability for modified derivation, as argued by Peterson and Riggs (1999), children should be able to solve the false belief task earlier.

Clearly the relationship between counterfactual reasoning and false belief tasks needs to be investigated further to better understand its nature. Riggs et al. (1998), Peterson and Bowler (2000) and Grant et al. (2004) studied the relation between theory of mind and false belief tasks using only negative counterfactual questions. None of these experiments included any affirmative counterfactual questions. Following the results of the previous experiment showing that counterfactual affirmatives were easier than counterfactual negatives for children, and the evidence from the literature studying how adults process negative premises it is necessary to investigate the relationship between counterfactual

thinking and the false belief task controlling for affirmative and negative questions. The power of the studies conducted by Riggs et al. (1998) and Peterson and Bowler (2000) was also not very high as they collected only between two and four data points per type of question for each participant. Grant et al.'s (2004) study was more powerful as they collected eight data points per question per participant. Perner et al. (2004) argued for a developmental dissociation while studying this relationship in the context of simple versus complex scenarios but in their paper they did not give the full detail of all the experiments they conducted. They gave two examples out of the four counterfactual questions used, and both those examples were affirmative counterfactual questions. It was the first time that this relationship between counterfactuals and the false belief task had been investigated in the context of affirmative counterfactual question. However, Perner et al. (2004) did not seem to counterbalance the affirmative counterfactual questions with some negative counterfactual questions. Nor did they highlight this difference in their study when relating it to other research papers in the area. It is therefore difficult to decide whether this difference is due to the polarity of the counterfactual questions.

Another explanation regarding the correlation between counterfactual thinking and the false belief task found in the literature could be suggested. This correlation has mainly been studied with negative counterfactual questions, and in these questions the change of state of affairs is implicit. For example, in the question 'If there had been no fire where would Peter be now?', the change is implicit and leaves many possibilities available. In the Post Office scenario Peter got out of bed to help the man from the post office put out a fire. If there had not been a fire a wide variety of events could have happened. It is not definite whether these events would have led to Peter staying in bed or not. On the other hand with an affirmative counterfactual question such as 'If I had drawn a circle on this piece of

paper, where would this piece of paper be now?' the change is explicit, it does not leave any other possibilities open. If a circle had been drawn on the piece of paper, the piece of paper would be in the correct box. In the standard false belief task the false belief the character holds is implicit. The children need to understand and imagine this false belief. Grant et al. (2004) used some non-standard false belief tasks where the false beliefs were explicit but they did not report any correlation between these tasks and counterfactual thinking. It could be argued that the correlation between negative counterfactual questions and standard false belief task could be related to the ability to comprehend an implicit concept.

The two aims of the present study were as follows. The first was to replicate the finding that including a negation in a counterfactual question increases the difficulty for the children. The second was to investigate the relationship between counterfactual thinking and the false belief task and study whether the wording of the counterfactual questions (affirmative versus negative) influences the results. The false belief questions were presented either using the word 'think' or 'look'. The literature does not report any difference between these two types of question (Wellman et al., 2000) but they have not been studied in combination in relation to counterfactual thinking. The experimenter wanted to investigate whether those two types of questions would correlate differently with the counterfactual scenarios. Using different types of false belief questions would increase the sensitivity of the study.

As the previous study showed that children under the age of 3;6 could not cope with the task requirements it was decided to test children who were older than that age.

## **2 Method**

### **2.1 Participants**

Forty one children (18 boys and 23 girls) between the age of 3; 6 (years; months) and 5; 8 took part in this experiment. The mean age of the sample was of 4; 8 years (s.d. = 9 months).

The children were recruited from two reception classes (24 children including 12 boys and 12 girls) and a nursery (17 children including 6 boys and 11 girls) in mainstream schools in Plymouth, UK.

All the children were typically developing children.

### **2.2 Material**

A recording sheet was used throughout the experiment to note the responses given by the children and to provide strict guidelines to the experimenter (Appendix C)

Each child was presented with eight scenarios using toys, pens and papers. All the scenarios were told from the perspective of a teddy bear called Tidy Teddy. Each child was asked to answer two affirmative counterfactual questions, two negative counterfactual questions, and four theory of mind questions. In order to investigate different possible associations between counterfactual thinking and false belief tasks two types of false belief questions were included: the 'think' questions (e.g.: Where does Tidy Teddy think his piece of paper is?) and the 'look' questions (e.g.: Where will Tidy Teddy look for his piece of paper?).

The scenarios were very similar to the ones used in the first experiment. The first part of the scenario required the children to classify some objects (play dough with different shapes, pieces of paper, cards, bricks and a variety of pictures) in the correct box out of the three in front of them. Reminder cards were placed in front of each box so the children could see how the objects were to be classified

without having to remember the classification rule. This classification was then either followed by a counterfactual question or two theory of mind questions. The false belief tasks were systematically followed by a control question: the memory question (e.g.: Where was the brick at the beginning?)

#### Example of a scenario: Pieces of paper and shapes

The child was presented with 3 boxes, and some pieces of paper with either a circle, a square or no shape drawn on them. The child was asked to classify the pieces of paper into the correct box according to the shape on them. All the pieces of paper with a circle on went in the first box, those with a square in the second one and the blank pieces of paper in the third one. The experimenter was helped by Tidy Teddy. Tidy Teddy added a blank piece of paper in the third box and then went to sleep in the experimenter's bag. The experimenter removed Tidy Teddy's piece of paper, drew a circle on it and then asked the following questions to the child: 'What is on this piece of paper now?', 'Where shall we put this piece of paper now?' In the counterfactual situation the children were then asked either an affirmative question; 'What if I had drawn a square on this piece of paper, which box would it be in now?', or a negative question: 'What if I hadn't drawn a shape on this piece of paper, which box would it be in now?' In the theory of mind condition the experimenter still removed Tidy Teddy's piece of paper, drew a shape on it and asked the child to classify it in the correct box. At this stage Tidy Teddy woke up and wanted to play with his piece of paper. The child was then asked the two following questions, the order being counterbalanced across scenarios: 'Where will Tidy Teddy look for his piece of paper?', 'Where does Tidy Teddy think his piece of paper is?' In every condition the child was then asked a memory control question: 'Which box was Tidy Teddy's piece of paper in at the beginning?'

The other scenarios involved cards and colours, playdough rolled in different shapes, bricks and stickers, faces and glasses, cartoon characters and balls, teddies and bow ties and cartoon characters and hats (Appendix D).

### **2.3 Design**

This task has a within participant design. Each child was presented with eight scenarios and had to answer eight questions: two affirmative counterfactual questions, two negative counterfactual questions, and four theory of mind questions (two including the verb 'think' and two 'look').

### **2.4 Procedure**

The children were tested at school, during school time. The scenarios were presented over two sessions lasting between 10 and 15 minutes each, one week apart. In each session the children were presented with four scenarios and had to answer one affirmative counterfactual question, one negative counterfactual question, two pairs of false belief questions once with 'think' first followed by 'look', and once with 'look' first then followed by 'think'. The order of task was randomised within each session using a graeco-latin square.

## **3 Results**

The data collected were the numbers of correct answers per participant for each type of question.

### **3.1 Counterfactual Questions**

Firstly, the overall results for the counterfactual questions were analysed. The total number of correct answers for each participant was computed. Very few children answered none or only one question correctly. The vast majority of children (75.6%) answered at least three out of four questions correctly.

The performances between affirmative counterfactual questions and negative counterfactual questions were compared. The total number of correct answers for each type of question was computed and the mean and the standard deviation calculated. The results are presented in Table 4.

Table 4: Means and standard deviations for affirmative counterfactual questions and negative counterfactual questions (n=41) – Maximum = 2

Type of question	Mean	Standard Deviation
Affirmative CF	1.63	.58
Negative CF	1.39	.83

The results showed that the children's performance on the affirmative counterfactual questions was better than on the negative counterfactual questions. A t-test showed that this difference was significant on a one tailed test,  $t(40)=1.919$ ,  $p<.05$ .

### 3.2 Analysis of mistakes for the counterfactual tasks:

When asked a counterfactual question, children had a choice of three possible answers: the correct answer (the correct box), the realistic answer (the box where the object really was), and the other mistakes (third box).

The data showed that children made three types of mistakes: the 'realist mistakes' when they chose the box where the target object really was, the 'irrelevant box mistakes' when they pointed to the irrelevant box, and 'other mistakes' such as a lack of answer, or saying 'I don't know'. The total number of mistakes was computed for each of those categories and percentages were calculated. The interesting aspect of these data is the difference between Realist Errors and any other type of error, so the 'irrelevant box' mistakes and the 'other mistakes' were computed together and labelled 'Non Realist Errors'. The results are presented in Table 5.

Table 5: Total number and percentages of each type of errors for counterfactual affirmative and counterfactual negative questions (n=41).

	Realist Errors	Non Realist Errors	Total
C.F. Aff.	8 (57%)	6 (43%)	14
C.F. Neg.	14 (58%)	10 (42%)	24
Total	22 (58%)	16 (42%)	38

In both cases, affirmative counterfactuals and negative counterfactuals, children tended to make more realist errors than non realist errors, and the percentage of realist and non-realist errors are very similar.

### 3.3 False belief tasks for the whole sample

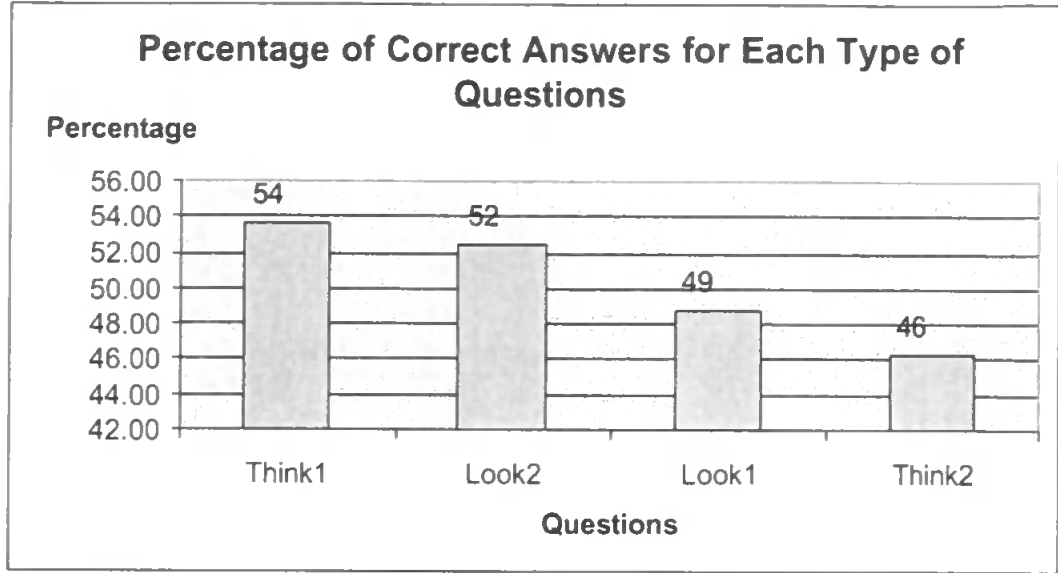
All the false belief tasks included a control question: the memory question. The data were analysed in two different ways. Firstly the whole data set was analysed, ignoring the answers to the control questions. Secondly the data were analysed taking into account the responses to the memory question. When children did not correctly answer the memory questions their data on the test question were not included in the data analysis. The results of the second analysis did not differ significantly from the first so they are presented in Appendix E.

The percentage of correct answers for 'look total' and 'think total' was calculated. Fifty percent of all the 'think' questions and 50.6 percent of the 'look' questions were answered correctly. A t-test showed that this difference was not significant,  $t(40) = -.121$ ,  $p = .905$ .

The influence of using 'think' or 'look' in the false belief questions as well as the order of presentation was explored. The percentage of correct answers was calculated for each type of questions in relation to the verb used and their order. The results are presented in the graph below.



Figure 1:Percentage of correct answers for each type of questions.



Differences between all the false belief questions can be identified in but an ANOVA showed that there was no significant difference in performance between the four questions,  $F(3, 120) = .59, p=.62$ . Therefore, there was no difference in performance between 'think' and 'look', and there was no order effect.

3.4 Analysis of mistakes for the false belief task

As for the counterfactual questions, when asked the theory of mind questions the children had a choice of three possible answers: the correct answer (the correct box), the realistic answer (the box where the object really was), and the other mistakes (the third box, or other types of mistakes such as irrelevant verbal answer or lack of answer altogether). The total number of realist errors and the total number of non-realist mistakes were computed and are reported in Table 6.

Table 6: Total number of each type of errors for theory of mind questions (n=41).

	Realist Errors	Non-Realist Errors	Total
Number of Errors	107 (66%)	56 (34%)	163

The results showed that children made more realist errors than non-realist errors but the proportion of the latter is still important.

**3.5 Comparison between performance on counterfactual tasks and false belief tasks**

The performances on counterfactual tasks and false belief tasks were compared. The total number of correct answers on counterfactual tasks was computed and the total number of correct answers on false belief questions was calculated and then divided by two to obtain a score out of four for each set of data. The means and the standard deviations were calculated for each type of question. The results are presented in Table 7.

Table 7: Means and standard deviations for the total number of correct answers for counterfactual questions and false belief questions (n=41).

Type of Question	Mean	Standard Deviation
Counterfactual questions	3.04	1.11
False Belief Questions	2.01	1.23

The results showed that children's performance on counterfactual questions was higher than on the false belief questions and a t-test showed that this difference was highly significant,  $t(40)=4.38$ ,  $p<.001$ .

**3.6 Correlations between counterfactual thinking and false belief task for the whole sample**

The link between counterfactual thinking and false belief task was investigated by running correlations. The results are presented in Table 8.

Table 8:Correlations between counterfactual thinking and false belief task (n=41).

	T.o.M. Total	C.F. Total	T.o.M Think	T.o.M. Look	C.F. Aff	C.F. Neg
T.o.M Total	1					
C.F. Total	.172	1				
T.o.M Think	.883**	.146	1			
T.o.M. Look	.888**	.159	.568**	1		
C.F. Aff	.059	.688**	.095	.011	1	
C.F. Neg	.190	.866**	.131	.205	.233	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

The correlation matrix showed that there was no significant correlation between any of the counterfactual tasks and any of the theory of mind questions. There were some positive, but not significant, correlations between false belief tasks and counterfactual tasks, and in particular between false belief task tasks and negative counterfactual tasks.

As Riggs et al. (1998) argued that the children make mistakes on counterfactual and false belief task tasks because they cannot suppress their knowledge of current reality the relationship between the number of realist errors and counterfactual thinking was investigated. The correlations are presented in Table 9.

Table 9: Correlations between the number of realist errors and performance of false belief questions (n=41).

	Realist errors on counterfactual	Think 1	Look 2	Look 1	Think 2
Realist errors on counterfactual	1				
Think 1	.116	1			
Look 2	-.174	.508(**)	1		
Look 1	-.115	.338(*)	.616(**)	1	
Think 2	-.010	.535(**)	.401(**)	.538(**)	1
Think total	.057	.865(**)	.516(**)	.505(**)	.887(**)
Look total	-.160	.470(**)	.897(**)	.901(**)	.523(**)

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

These results showed that none of the correlations between realist errors on counterfactuals and false belief questions were significant.

The correlations between the number of realist errors (on counterfactual and theory of mind task) and performance on counterfactual and theory of mind tasks have also been studied. The results are presented in Table 10.

Table 10: Correlations between the number of realist errors and performance on counterfactual and theory of mind tasks.

	Realist errors (counterfactual)	Realist errors (theory of mind)	Theory of mind	Counterfactual
Realist errors (counterfactual)	1			
Realist errors (theory of mind)	-.147	1		
Theory of mind	-.059	-.699**	1	
Counterfactual	-.757**	-.014	.172	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

The results did not show any correlation between the number of realist errors and performance on theory of mind and counterfactual tasks.

As mentioned earlier the results were also analysed in a different way. The data of the participants who failed a control question on the false belief task were discarded from the analysis. The results were extremely similar to the ones presented in the present results section. This second analysis is available in Appendix F.

## **4 Discussion**

This experiment aimed to study the influence of negative versus affirmative questions on counterfactual thinking and the link between counterfactual thinking and the false belief task.

The children were asked to answer four counterfactual questions, two of which were affirmative and the other two negative. The majority of the children (75%) could answer at least three out of four counterfactual questions correctly. These results are very similar to the ones reported by Harris et al. (1998) who found that 76.92% of their participants, aged between 3 and 5, correctly answered three or four test questions out of four. Harris et al. (1998) also included two affirmative and two negative questions in their design. It seems that these children were able to suppress their current knowledge of reality and imagine how the world would be if an event had not happened or a different event had happened. The present tasks were different from Harris et al.'s (1998) tasks as they involved less narration. Harris et al. (1998) acted out narrative scenarios and then asked counterfactual questions. In the present study children only needed to be able to classify objects in different boxes according to one of their properties and then process a counterfactual question, they did not have to listen to a narrative story. This study required less language but the results were still very similar.

In order to study the influence of negation on counterfactual thinking half of the questions included in the present experiment were negative and the other half affirmative. The results showed that the language used to ask counterfactual questions to children was important: the negative questions were more difficult to answer than the affirmative questions. This confirmed the findings reported in the previous chapter of this thesis. Although Harris et al. (1998) counterbalanced the affirmative and the negative counterfactual questions in their study, they did not analyse the results accordingly, so we cannot be sure that their results were similar to the present ones.

Some explanations can be suggested regarding the difficulty introduced by the negation of the question. Firstly, it could be argued that negative questions are more ambiguous than affirmative questions, especially in the present context. An affirmative counterfactual premise such as 'What if I had drawn a square on this piece of paper, which box would it be in now' requires the children to imagine only one situation 'drawing a square'. The change is explicit. However, the negative premise 'What if I hadn't drawn a shape on this piece of paper, which box would it be in now?' requires the children to imagine that one thing did not happen, i.e. the experimenter did not draw a shape, but it could be any type of shape. The question does not explicitly state that nothing else was drawn, therefore the children might be more confused by the negative than the positive premise. The change is in this case implicit.

The negative question in the experiment included the word 'shape' on which the children might focus, and they might therefore point to one of the boxes where pieces of paper with shapes on are classified. If they do so, they then point to one of the incorrect boxes. This effect is described as the 'matching bias' (Evans, 1998), which is the tendency to choose as correct a conclusion that matches the lexical content of the premise. The matching bias could be a second explanation

for this difficulty with negation. Children concentrate on the semantics of the key words in the sentence but ignore the negation itself.

Riggs et al. (1998) found that counterfactual reasoning was more difficult than future hypothetical reasoning, but as explained in the previous chapter, they compared negative counterfactual questions and affirmative future hypothetical questions. The results from the present experiment tend to support their findings, as it appeared that negation makes a question more difficult to answer. However, in their paper, Riggs et al. (1998) offered a different explanation for their findings based on the nature itself of those two types of reasoning. They argued that children find counterfactual reasoning more difficult than future hypothetical reasoning because it requires them to suppress their current knowledge of reality and imagine a different world which is meant to stand in for reality. According to Riggs et al. (1998) this difficulty leads the children to make realist errors. This argument regarding realist errors cannot be fully supported by their data as the children were only offered two choices for their answer: the correct location or the realist location. Children could therefore either give the correct answer or make a realist error. It is then not possible to argue that children make errors because they cannot suppress reality. In order to distinguish realist errors and other types of errors in the present study the children were offered a third choice being an 'irrelevant location'. When they were asked a counterfactual question such as 'If I had drawn a square on that piece of paper which box would it be in now?' they had a choice of three boxes: the correct box, the reality box, which is where the piece of paper is at the time the question is asked, and another box which is not relevant. The errors children made in this study were analysed as being either 'realist errors' or 'non realist errors'. The results showed that children both made realist and non realist errors. Children tended to make more realist errors, which tends to support Riggs et al.'s (1998) claim that children find counterfactual

reasoning difficult because they cannot suppress their current knowledge of realist but an important proportion of errors children make are still non-realist. There seem to be several factors leading children to make mistakes. Some make mistakes because they cannot suppress current reality but others make mistakes because of other difficulties of the tasks such as cognitive or linguistic difficulties.

This study also investigated the false belief task and its relationship to counterfactual tasks. The children were asked to answer eight false belief questions that included two types of test questions: the 'think' questions and the 'look' questions. Using the word 'think' in the test question, such as 'Where does Tidy Teddy think his piece of paper is?' is asking the children whether they know what Tidy Teddy's mental state is. On the other hand in order to pass the 'look' question, 'Where will Tidy Teddy look for his piece of paper?', children need to acknowledge Tidy Teddy's mental state, especially his false belief, then they have to understand that Tidy Teddy will act according to his own beliefs, even though they are false. The results of the experiment did not show any significant difference between both types of questions, which is consistent with previous research (see Wellman et al., 2001, for a review). It does not seem that just acknowledging a mental state is easier than predicting the protagonist's action according to the mental state. The order of presentation had no effect either. Asking one before the other neither improves, nor worsens, the children's performance on the second question.

A number of studies have shown a link between counterfactual thinking and the false belief task (eg: Riggs et al, 1998; Grant et al., 2004; Peterson and Bowler, 2000) so this relation was investigated in the present experiment. When comparing performance on the counterfactual tasks and the false belief tasks it appeared that children found counterfactual tasks easier than false belief tasks. This same difference was reported by Riggs et al. (1998) and Peterson and



Bowler (2000), but only for clinical samples (children with autism and children with severe learning disability) in that latter publication. These findings confirm Riggs et al.'s (1998) suggestion that counterfactual thinking may be necessary but not sufficient to pass the false belief task. Riggs et al. (1998) and Peterson and Bowler (2000) explained that difference by saying that the counterfactual questions include an explicit counterfactual proposition and that the counterfactual proposition embedded in the false belief task is implicit and needs to be inferred making that task more difficult.

In order to investigate this relationship between those two types of tasks we examined the correlations between them. Having established that affirmative and negative counterfactual questions were not of the same difficulty it is interesting to investigate how they independently correlated with the false belief task.

The first important result is that the false belief task questions positively correlated with each other very strongly indicating that they share a high degree of common variance. However, this experiment failed to show any significant correlation between any of the false belief task questions and any of the counterfactual questions. There are slight positive correlations between the overall score for the counterfactual questions and the overall false belief tasks, the 'think' questions and the 'look' question but they are not significant. If we look at the affirmative and the negative counterfactual questions separately it appears that the counterfactual affirmative did not correlate with any of the false belief task measures, whereas the counterfactual negatives correlate slightly positively with them, although this correlation is not significant. These results, although not significant, tend to confirm the finding reported in the literature on this topic. The studies which found a significant systematic positive correlation between the false belief task and counterfactuals used negative counterfactual questions (e.g.: Grant et al., 2004; Peterson and Bowler, 2000; Riggs et al., 1998). This pattern of

findings needs to be investigated further. The non-significant correlations in this study could be due to a lack of statistical power because only two data points were collected respectively for negative and affirmative counterfactual questions.

On the other hand, Perner et al. (2004), who did not find a systematic correlation, used some affirmative counterfactual questions in their experimental design. This dissociation again suggests that affirmative and negative counterfactual questions may differ. The false belief task might share a common ability with the negative counterfactual questions but not with the affirmative counterfactual question. The fact that the negative counterfactual is more difficult might be one of the reasons.

Riggs et al. (1998) suggested that the correlation between counterfactual thinking and the false belief task could be due to the fact that children cannot suppress their knowledge of current reality leading them to make realist mistakes. However, the correlation matrix did not show any correlation between the number of realist errors made in the counterfactual tasks and performance on the false belief task. The number of realist errors between the counterfactual and the false belief tasks do not seem to correlate either. The findings of this experiment seem to show a divide between these two types of tasks.

The positive correlation found in the literature and here (although not significant) between negative counterfactual questions and false belief tasks could be due the fact that they both include an implicit component. As mentioned before, in a negative counterfactual question the change is implicit and in traditional false belief tasks the false belief is implicit. When the change becomes explicit (in the affirmative counterfactual question) the correlation decreases. This relationship needs to be explored further.

A last interesting point to discuss is the usefulness of the control questions in the false belief task. The present false belief tasks included one control question: the 'memory question' (e.g.: Where was this piece of paper at the beginning?) that checks that the child can remember where the target object was at the beginning of the task. It is argued (Wellman et al., 2001) that the results of the test questions need to be validated by correct answers on the control questions. The other side of the argument would be to say that by the time children answer a control question, they have had to listen to a rather complex scenario, and to answer at least one test question. The child could be confused by the number of questions asked. In this study when comparing the data for the whole sample (including the data of all the children who took part in the experiment) and the reduced sample (excluding the children who failed the control question) it can be seen that the results are very similar. The comparison of the 'think' and 'look' questions leads to the same conclusion in both cases, and the correlation matrices are very similar. Removing the children who fail the control question can be seen as increasing the internal validity of the study, but then the number of participants diminishes, from 41 to 25 in the present case, therefore decreasing the power of the statistical analysis. It does not seem that these control questions are particularly necessary, at least with typically developing children. They might still be necessary when testing children from clinical populations as their cognitive skills and language skills tend to vary a lot.

In conclusion, this study confirms that affirmative and negative counterfactual questions are not equivalent, they are not of the same difficulty and they probably require different abilities. This experiment helped us understand better the relation between false belief tasks and counterfactual reasoning, especially showing that the wording of the counterfactual questions is important. It

confirmed that introducing a negation in a counterfactual question increases its difficulty.

This experiment also showed that only a proportion of the mistakes children make when solving counterfactual questions are realist mistakes, so suppressing reality is probably not the only difficulty children have. The results also showed that counterfactual tasks were easier than false belief tasks but the correlations between those tasks were not conclusive and need to be further investigated. This experiment finally suggested that the relevance of the control questions is very limited, at least with typically developing children. The next chapter will investigate the relationship between the false belief task and counterfactual thinking, as well as contrary-to-fact reasoning.

# **CHAPTER 4**

## **EXPERIMENT 3: FALSE BELIEF TASKS, COUNTERFACTUAL THINKING AND CONTRARY TO FACT SYLLOGISMS**

### **1 Introduction**

The previous chapter of this thesis replicated the findings from experiment 1 that strongly suggested that using a negation in a counterfactual question makes that question more difficult. It also investigated the relationship between counterfactual thinking and the false belief task but contrary to previous research failed to show a significant association. The present chapter will study this link using a different set of tasks, including false belief to self tasks, as well as a potential link between false belief and contrary-to-fact syllogisms.

The relationship between counterfactual thinking and the false belief task as reported in the literature has been discussed in depth in chapter 1 and chapter 3. Riggs et al. (1998) initiated the work on the relation between counterfactual reasoning and the false belief task. They found a correlation between performances on these tasks and explained that in order to pass the false belief task children need to acknowledge a counterfactual physical state, they need to acknowledge how the state of affairs would be if a previous event had not happened. In the Sally and Anne task for example the children need to think about what the situation would be like now if Anne had not transferred the marble to the box. These findings were replicated by Peterson & Bowler (2000) and Grant et al. (2004) with children with autism and Guajardo & Turley-Ames (2004) with typically developing children.

This work on counterfactual thinking and the false belief task has always been conducted with traditional false belief tasks to others such as the 'Sally and Anne task' or the 'Maxi task' (described in Chapter 1) that require the children to

attribute a false belief to another person or character. In the example mentioned above the children need to attribute the false belief to Sally, or Maxi. In the previous chapter of this thesis no link was found between counterfactual and false belief tasks, but different types of tasks were used. In the present experiment more traditional tasks will be used, and the power of the experiment will be increased by using more tasks and testing more children.

A different type of false belief task, the false belief to self task, sometimes referred to as 'the unexpected content task' or 'the representational change task', has been reported in the literature. The main difference between this task and the traditional false belief to others task is that children need to attribute a false belief to themselves rather than somebody else. Perner et al. (1987) tried to help children understand another's person mistaken belief by giving them the opportunity to experience false beliefs themselves first, and to that end designed the Smartie task. In their experiment the participant came with a friend and both were told that the experimenter would show them what was in her box. The friend was told to wait outside. Inside the room the participant was shown a Smartie box and was asked what they thought was inside. Perner et al. (1987) reported that they all said 'Smarties'. The participant was then shown that he, or she, was wrong and that the box did not contain Smarties but a pencil. The pencil was put back inside and the box closed. The participant was then asked a control question about the actual content of the box ('Can you remember what's inside here?') and the first test question about their previous false belief regarding the content of the box ('But what did you think was in here?'). The participant was then told that it was their friend's turn and that their friend would be shown the closed box and asked what was in the box. The participant was finally asked the second test question ('What will [name of the friend] think is in here?'). The first part of the Smartie task required children to attribute false belief to themselves and the

second part to somebody else, their friend. The aim of the present experiment was to investigate false belief to self only so just the first part of this experiment was run. No friend or other character was involved in the running of this task. Perner et al. (1987) found that children could pass this task from the age of 4. Regarding the results of the test questions requiring the participants to attribute false belief to their friend, Perner et al. (1987) found that having experienced false belief themselves first did not help the children attribute false beliefs to their friend but their results also showed that children found the first task easier than the second one. They suggested that the false belief to self question was easier because children could only remember their previous wrong answer without having to understand the conflicting values represented by the reality and the false belief. Lewis and Osborne (1990) replicated this study and found the same discrepancy.

Gopnik and Astington (1988) ran an experiment looking at the difference between false belief to self, false belief to others and the appearance-reality distinction using the Smartie task. Their results contradicted the ones previously cited and they found that children performed better on the false belief to others task than the false belief to self task. Two explanations for these results could be suggested. First of all Gopnik and Astington (1988) did not ask the children what they thought was inside the box when it was first presented to them to avoid the children just reporting their previous answer when asked the control question. Secondly Gopnik and Astington (1988) argued that this result could mean that there is a need for children to understand the difference between their own beliefs and others' beliefs before they can learn to understand changes in their own mental states.

In their meta-analysis Wellman et al. (2001) included 117 self-conditions and compared them to 118 other-conditions. They found no significant difference in difficulty between the two types of tasks. Their analysis also showed that both

tasks developed similarly across ages. It seems that the difference between performances on these tasks in the studies cited above disappeared when a large number of papers were compared.

False belief to self can also be tested with a slightly different type of task. The first part of the Band Aid task designed by Bartsch and Wellman (1989) can be used for this purpose. In order to run that task the contents of a plaster box was moved to a plain box before the experiment started. The participants were shown both boxes and were asked to pick up the box they thought had plasters in. The participants tended to always pick the plaster box which was clearly marked. They were then asked to check the contents of the box and discovered that the plaster box was empty and that the plasters were in the plain box. At this stage of the scenario Bartsch and Wellman (1989) introduced a hand puppet and the scenario was presented either in prediction condition or in explanation condition. In the prediction condition the participants were required to predict the hand puppet's behaviour and were asked the following questions: 'Look, here's Sam. Sam has a cut, see? And she wants a band-aid. Where do you think she'll look for band aids?' In the explanation condition the participants were asked to explain the hand puppet's behaviours. The participants watched the puppet starting looking in the empty plaster box without revealing the content and they were asked the following question: 'Look, here's Bill. Bill has a cut, see? And he wants a Band Aid. Why do you think he's looking in there?'

The Band Aid task can be used to test false belief to self attribution if just before the point where the hand puppet is introduced both boxes are closed and the participants are asked the following test question: 'Where did you think the plasters were before you opened the boxes?' and a control question: 'Where are the plasters really?'



Riggs et al. (1998) suggested that, in the false belief to self task, children do not need to reason counterfactually but need to report a counterfactual situation. The false belief to self task will be included in the present experiment but according to Riggs et al.'s (1998) study no correlation should be expected with the counterfactual task. However, it could be argued that there is a counterfactual element embedded in this task. In the example of the Smartie task (as described above) the implicit counterfactual question could be: 'What if I hadn't swapped the content of the box, what would be in it now?' Riggs et al. (1998) made the previous claim regarding the false belief to self task but without precise supporting evidence. Therefore this potential link between the false belief to self task and counterfactual reasoning needs to be investigated further.

A study conducted by Peskin and Ardino (2003) included these two types of theory of mind tasks. This study aimed at exploring the relationship between the development of theory of mind and the ability to engage in two social behaviours that are hide-and-seek games and keeping a secret. The participants who took part in this study were children from the age of 3 to 5. Peskin and Ardino (2003) included three types of theory of mind questions in their design: a traditional false belief to others tasks (that they called a location task) and a traditional Smartie type tasks that included the false belief to self question and the second false belief to others question. They found a strong relationship between all these measures indicating that they evaluated the same type of general abilities; however, the details of the results showed some differences. They found that the false belief to others task significantly positively correlated with hide-and-seek but not with keeping a secret and that the false belief to self task did not correlate to either of these social measures. These results suggested that the theory of mind to self task and the theory of mind to others task, although both being theory of mind tasks, do not exactly tap into exactly the same ability. These findings suggested

that included these two types of theory of mind tasks in the present study would enrich the variety of tasks and would allow to study the relationship between counterfactual reasoning and contrary-to-fact syllogisms with theory of mind in more depth.

The strong link found in the literature between counterfactual thinking and the false belief task has mainly been attributed to the counterfactual element itself. However, it could be argued that the correlation could be linked to more general reasoning abilities. In order to test this hypothesis the present experiment will examine the relationship between the false belief task and another type of reasoning task: contrary-to-fact syllogisms.

Developmental work on syllogisms has mainly been conducted using contrary-to-fact syllogisms (Dias & Harris, 1990; Leavers & Harris, 2000; Leavers & Harris, 1999; Scott et al., 1999). The principal focus of this work was to study the impact of content on reasoning (see chapter 1 for more details). In all these experiments the quantifier was always 'all' which is defined as universal as it refers to all the members of a set (Evans et al., 1993) and all the syllogisms were valid and followed by a strong conclusion as there was only one possible correct answer. The structure of the contrary-to-fact syllogisms was as follows. The first premise was contrary-to-fact, i.e. the statement was contrary to the real world, could not happen in the real world (e.g.: all cows quack, from Leavers and Harris, 2000). This premise could take the form of All X are/feel/have/do Y. In order to check that the first premise was contrary-to-fact for the children, participants were asked a factual question (e.g.: What noise do cows make?). The children performed extremely well on the factual questions consistently across experiments (e.g.: Scott et al., 1999; Dias and Harris, 1990). The second premise gave more information about the middle term 'cows'. It defined a particular member of set X (e.g.: Suzie is a cow). The final statement was a question summarising the first

two premises. The questions required either a yes-answer (Does Suzie quack?) or a no-answer (Does Suzie moo?). In both cases children need to draw a conclusion that contradicts their knowledge of the world. These tasks were deductive and could be solved without any knowledge of the world.

Although the validity of these arguments is a function of the logical structure rather than the empirical content children seem to have difficulties giving an answer that contradicts their knowledge of the world (Dias and Harris, 1988). Contrary-to-fact syllogisms require children to suppress their knowledge of the real world in order to reason with the new piece of information. They need to set reality aside to avoid empirical bias (Harris, 2001). Contrary-to-fact syllogisms require children to represent a counterfactual situation so to some extent they share some similarities with counterfactual thinking tasks.

Ways of facilitating contrary-to-fact reasoning have been studied and the literature has shown that some particular contexts helped children solve contrary-to-fact syllogisms, in particular a fantasy context where children are asked to imagine a fantasy world that follows different rules than the real world (Scott et al., 1999; Leever et al., 2000; Leever et al., 1999). Markovits (1985) argued that a fantasy context acts as a cognitive filter preventing children from accessing long term memory. This cognitive filter helps inhibition of real world knowledge then the children find reasoning about with contrary-to-fact proposition easier.

Even if we accept that counterfactual thinking is a pre-requisite to passing the false belief task, it does not explain the full variance of the latter (Guajardo et al., 2004), therefore some other variables have to be taken into account. Other cognitive abilities such as the central executive have been studied in relation to the false belief task. Carlson and Moses (2001) reported a strong correlation between theory of mind and inhibition even after controlling for age, gender, verbal ability, motor sequencing, family size, and performance on pretend-action and mental

state control tasks. To pass the test children need to inhibit response to the wrong answer which represents reality. Therefore a correlation between false belief to others tasks and contrary-to-fact syllogisms could be expected.

In order to pass the false belief task children need to listen to a narrative and reason about the main elements of the story, before they could do that they need to hold these main elements in their memory. It can then be hypothesised that verbal memory could play an important part in solving the task, therefore a measure of verbal memory was included in the experiment. The measure used was the CNRep (Gathercole, Willis, Baddeley, & Emslie, 1994a). In this test children are asked to repeat non-words presented orally to them.

The aim of this study was to study the relationship between contrary-to-fact and counterfactual reasoning and performance on the false belief task. The present study will further investigate the link between counterfactual thinking and the false belief to others task. In order to test whether this potential link is only due to the counterfactual element rather than general reasoning abilities we will compare false belief to others task with contrary-to-fact syllogisms. If a link is then found it would mean that the link between counterfactual and false belief tasks previously found in the literature was mainly due to general abilities. The false belief to self task will also be included in this study to investigate its relationship, or lack of it, with counterfactual thinking.

## **2 Method**

### **2.1 Participants**

Seventy one participants, including 27 boys and 44 girls, were recruited in two schools and one playgroup in Plymouth, UK.

Forty three children attended nursery classes and 28 reception classes.

The mean age of the participants was 4 years 5 months and the standard deviation 6 months (range 3;3 to 5;2).

## 2.2 Materials

The children were presented with four counterfactual tasks, four false belief to others tasks, four false belief to self tasks, eight contrary-to-fact syllogisms and the CNRep.

The CNRep test was administered using the recording sheet and the tape provided with the test. Toys were used to act out all the theory of mind and counterfactual scenarios.

### 2.2.1 Counterfactual tasks

The four counterfactual tasks were the same as the ones designed by Harris et al. 1996, (experiment 1), except some minor variations due to physical constraints when the specific materials were not available. Each of the scenarios was acted out in front of the children with the help of props.

The first scenario was as follows:

*'One day the floor was nice and clean like this (experimenter points to a white sheet of paper). But guess what? Carol comes home (a doll is put on the table next to the sheet of paper) and she doesn't take her shoes off. She comes inside and makes the floor all dirty with her shoes (the doll walks across the floor leaving dirty footprints made of paint).'*

At this stage the children were asked two control questions: the Now Control Question (e.g.: 'Is the floor clean now?') in relation to the final situation and the Before Control Question ('Was the clean before?') in relation to the initial situation.

Finally the children were asked the test question. For example: 'What if Carol had taken her shoes off – would the floor be dirty?'

Three other similar scenarios using different props were used. The first two episodes requested a yes/no answers and the last two a selective pointing answer.

One test question required a no answer and the next one a yes answer in order to avoid any bias. For the same reason, the control questions required either a yes or a no answer. Each scenario required a mixture of yes and no answers.

Two of the final test questions were formulated in the positive ('What if...?') whereas the other two were formulated in the negative ('What if...not...?').

The other scenarios are included in Appendix F.

### **2.2.2 False belief to others tasks**

The scenarios used were similar to the classical tasks such as the Sally and Anne task ((Baron-Cohen et al., 1985) and the Maxi scenario (Wimmer & Perner, 1983).

One of the scenarios was as follows:

The Sally and Anne task was acted out with two dolls. The children were first introduced to each of the doll and the experimenter checked that they knew which one was which. Sally placed a marble in her basket covered with a handkerchief, and then left the scene (Sally was placed in the experimenter's bag). While she was away Anne moved the marble from the basket to her box closed with a lid. Sally then returned to the scene and the children were asked the critical question: 'Where will Sally look for her marble?' If the children pointed to the basket, then they passed the false belief task because they showed they could acknowledge Sally's false belief.

They were then asked two control questions: the Reality question ('Where is the marble really?') to check that the children knew what the situation was, and the

Memory question ('Where was the marble in the beginning?') to check that the children remembered the first state of affairs.

The Pooh Bear task was the same as the Sally Anne tasks. The characters, the containers and the objects to be moved were slightly different.

The Maxi task was inspired from Wimmer and Perner (1983), but very close to the Sally and Anne tasks described above.

The last false belief task was the 'Post office story', inspired from (Riggs et al., 1998). This story was told as follows while being acted out with toys:

*'One day Sally and Peter were in their house but Peter wasn't feeling very well, so he went to bed. Sally then went to the shops to get some medicine (Sally was put in the experimenter's bag). Whilst Sally was at the shops the phone rang and the man from the Post Office asked Peter to come to help to put out a fire. Peter got out of bed and went to the Post Office. Sally had finished shopping (Sally was brought back to the scene).'*

The children were then asked the false belief question: 'Where does Sally think Peter is?' As in the other scenario they were then asked the Reality question: 'Where is Peter really?' and the Memory question: 'Where was Peter before Sally left?'

For the four scenarios the children could either give a verbal answer or point to the location. The other scenarios are included in Appendix G.

### **2.2.3 False belief to self task**

The Smartie task was similar to the first part of the one designed by Perner, Leekham and Wimmer (1987). The children were shown a Smartie box and asked: 'What's inside the box?' The children should say Smarties, sweets or chocolate. They were then asked to open the box to check the content. When they did so, they found a pencil, instead of Smarties, in the container. After the lid was then put back on the box the children were asked the critical question: 'What did you think was inside the box before you opened it?' Finally they were asked the Reality

control question: 'What is really inside the box?' to check that they knew what the real situation was.

The Crayon task was similar to the Smartie task, but it used a crayon box which content was replaced by a rubber.

The Band Aid task was similar to the one designed by (Bartsch & Wellman, 1989) but used to test the children's own beliefs. For this task the children were shown two boxes: one marked with a familiar and obvious picture (i.e.: a plaster box) and another plain unmarked one of similar size. The children were asked a question to set their belief, such as 'Where are the plasters?' They should then point to the plaster box. They were then asked to open both boxes and check their content. At this point they could note that the content of the boxes had been switched, for example the plasters were in the unmarked box and the plaster box was empty. Both boxes were then closed and the children asked the critical questions: 'Where did you think the plasters were before you opened the boxes?' They were finally asked a Reality control question: 'Where are the plasters really?'

The Cereal task was exactly the same as the Band Aid task, only the boxes and their content differed. A small breakfast cereal box and a plain green box were used. The content of the breakfast cereal box was move to the other box before the beginning of the experiment.

The Crayon task and the Cereal task are included in Appendix H.

#### **2.2.4 Contrary to Fact syllogisms**

The contrary to fact syllogisms used in the experiment were similar to the ones designed by Leever and Harris (2000). As Leever and Harris (2000) found that children answer extremely well the factual questions testing their knowledge of the real world it was decided not to include them in the present experiment.



The children were asked to solve eight contrary to fact syllogisms. Four of them required a 'yes' answer and four a 'no' answer. The order of presentation of the syllogisms was randomised ([www.randomizer.org](http://www.randomizer.org)). The children were tested individually with the syllogisms read aloud to them.

The following instructions were first read to the children:

*'Let's try this one together. Now, I'm going to tell you something that may sound funny, but I want you to think about it. I want you to imagine a picture in your head. It could be; All fish live in trees. So, the fish you are thinking about, in the picture in your head, do they live in trees? Yes? Good. So now I'm going to tell you some stories and every time I tell you a story I want you to think about it. I want you to make another picture inside your head of it.'*

And then the first syllogism was read, for example:

*'All cows quack  
Susie is a cow  
Does Susie quack?'*

The first two premises were read twice to the children before asking them the final question. The full set of syllogisms is presented in Appendix I.

### 2.2.5 CNRep

In order to test the children's verbal memory the Children's Test of Nonword Repetition (Gathercole, Willis, Baddeley, & Emslie, 1994b) was used. This test shows a close developmental links with vocabulary, reading and comprehension skills. The test was administered strictly according to the manual. The children were asked to repeat non-words recorded on a tape and their answers were recorded on another tape. The recordings were scored later. Each of word was scored as a right or wrong answer. The total number of right answers was then computed.

## **2.3 Design**

A correlational design was used with each participant being asked to complete four counterfactual tasks, four theory of mind to others tasks, four theory of mind to self tasks, eight contrary-to-fact syllogisms and the CNRep test.

## **2.4 Procedure**

Children were tested individually at their school or playgroup over two sessions one week apart. Each session took 10 to 15 minutes per child.

In the first week the children were presented with two counterfactual scenarios, two theory of mind tasks and all the eight contrary-to-fact syllogisms. In the second week they were presented with the remaining two counterfactual scenarios and two theory of mind tasks and the CNRep. The order of the tasks was randomised within each session. The tasks were randomised with the help of an online software ([www.randomizer.com](http://www.randomizer.com)).

## **3 Results**

### **3.1 Control questions and biases**

The counterfactual and the false belief scenarios included two control questions each. In the literature, researchers tend to eliminate the participants who fail at least one control question, as they believe that control questions are a check that the children understand the task. The present research included four types of tasks including a total of 12 counterfactual and false belief scenarios, each including 2 control questions; altogether a total of 24 control questions per child. In the present sample only 11 children answered all the control questions correctly and did not present a yes or no bias on the syllogisms. It therefore did not seem appropriate to run correlations on such a small sample. In the previous chapter of this thesis the data analysis on false belief task and counterfactual tasks was conducted first of all by including all the participants and secondly by including

only the participants who answered all the control questions correctly and the results did not show any difference between these two types of analysis. Therefore, it was decided to include all the data collected for the false belief tasks and the counterfactual tasks in the present analysis.

On the other hand the contrary-to-fact syllogisms did not include any control questions but the participants showing a yes or a no bias (i.e.: when the children systematically answered yes or no to all the questions) were excluded from the data as we cannot be sure that they were processing the questions. These children could have replied automatically to the questions without processing them. Two participants showed a no-bias and 12 a yes-bias. Two other participants refused to do the task. The data from 55 participants was therefore finally analysed.

**3.2 Counterfactual task, false belief to others tasks and false belief to self tasks**

First of all the internal reliability of each task was computed using a Cronbach's Alpha. The results are reported in Table 11.

Table 11 . Cronbach's alpha for each task.

Task	Cronbach's Alpha
Counterfactual	.6251
False belief to others task	.6891
False belief to self task	.3822
Contrary-to-fact syllogism (all)	.4263
Contrary-to-fact syllogism (yes)	.6834
Contrary-to-fact syllogism (no)	.31

As a rule, it is usually accepted that a Cronbach's alpha higher than .7 demonstrates a high degree of reliability (Kline, 2000). Therefore in the present study the results showed that counterfactual tasks and false belief to others task had a good internal reliability as both Cronbach's alpha were close to .7 whereas false belief to self task had a poor internal reliability as they were less than .6.

The results also showed that contrary-to-fact syllogisms requiring a yes answer had a good internal reliability as the Cronbach's alpha was higher than .6, but that the contrary-to-fact syllogisms requiring a no answer and the overall syllogisms had a poor internal reliability as it was lower than .6.

The score for counterfactual tasks, false belief to others tasks and false belief to self tasks was the number of correct answers. The maximum score for each of these task was four. The mean and the standard deviation for these tasks were computed and are presented in Table 12.

Table 12. Mean and standard deviation for each task (n=71).

Task	Mean	Standard Deviation
Counterfactual thinking	1.83	1.35
False belief to others task	1.69	1.36
False belief to self task	1.32	1.27

A repeated measures ANOVA was conducted to explore whether these differences were significant. First of all, the assumption of sphericity was checked. The Mauchly's test of sphericity was not significant ( $w = .943$ ,  $p = .134$ ) therefore it was concluded that the variances of differences were not significantly different and that the F value of the ANOVA could be trusted. The main ANOVA suggested that there was a significant difference  $F(2) = 4.302$ ,  $p < .05$  between at least two of these tasks. A post hoc analysis showed that the only significant difference was between the false belief to self task and the counterfactual task. The children performed significantly better on counterfactual tasks than false belief to self tasks. There was no difference in performance between the two types of false belief tasks, or between false belief to others tasks and counterfactual tasks.

The counterfactual tasks could be analysed further by comparing the affirmative and the negative counterfactual questions. The scores were computed for both these types of questions. The mean and standard deviation for the affirmative counterfactual questions ( $M = .93$ ,  $S.D. = .78$ ) and the negative counterfactual questions ( $M = .90$ ,  $S.D. = .79$ ) were very similar. A t-test showed that this difference was not significant,  $t(70) = .293$ ,  $p = .770$ .

### **3.3 Contrary to Fact Syllogisms**

As mentioned above 55 participants were included in this analysis. The score on the contrary-to-fact syllogisms was computed. Altogether the mean score of correct answers for the children was 2.44 (maximum possible score = 8) with a standard deviation of 1.64, and no participant answered more than 5 syllogisms correctly.

The performances on the two types of syllogisms (maximum score was four for each of them) were compared and the results showed that children performed better on the syllogisms requiring a yes answer ( $M = 1.47$ ,  $SD = 1.39$ ) than on those requiring a no answer ( $M = .93$ ,  $SD = .97$ ). Statistical analysis revealed that this difference was highly significant;  $t(57) = 8$ ,  $p < .001$ , two-tailed.

### **3.4 CNRep**

The CNRep was presented to 71 participants. One participant could not complete it because of language difficulties and 8 children refused to do the task. The final sample therefore included 62 children.

As not all the children fell into the standardised age range for this test, only the raw score was computed, not the standard score or the percentile range. Each correct item counted for one point and all the points were added to give a raw score. The maximum possible raw score was 40. The range of performance varied from 3 to 40. The mean performance was 16.17 and the standard deviation 5.83.

### 3.5 Correlations

Correlations were run to explore the relation between all the tasks. The correlation table is presented in Table 13.

Table 13:Correlations between all the tasks separating yes and no syllogisms.

	C.F.	C.F. Aff.	C.F. Neg.	F.B.T. Others	F.B.T. Self	C.to.F	C.to.F (yes)	C.to.F (no)
C.F.	1							
C.F. Aff	.855** (n=71)	1						
C.F. Neg	.861** (n=71)	.472** (n=71)	1					
F.B.T. Others	.527** (n=71)	.420** (n=71)	.483** (n=71)	1				
F.B.T. Self	.314** (n=71)	.181 (n=71)	.356** (n=71)	.239* (n=71)	1			
C.to.F	-.181 (n=55)	-.175 (n=55)	-.133 (n=55)	-.180 (n=55)	.143 (n=55)	1		
C.to.F (yes)	-.204 (n=55)	-.207 (n=55)	-.141 (n=55)	-.236 (n=55)	-.182 (n=55)	.804** (n=55)	1	
C.to.F (no)	-.011 (n=55)	.003 (n=55)	-.021 (n=55)	.035 (n=55)	.499** (n=55)	.529** (n=55)	-.080 (n=55)	1
CNRep	.376** (n=62)	.375** (n=55)	.280* (n=55)	.306* (n=62)	.129 (n=62)	-.078 (n=49)	-.080 (n=49)	-.024 (n=49)

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

The two types of theory of mind tasks positively correlated with each other ( $r = .239, p < .05$ ). Both these types of tasks also correlated with the overall counterfactual tasks. However, false belief to others tasks correlated more strongly

with the counterfactual ( $r=.527$ ,  $p<.01$ ) than the false belief to self task did ( $r=.314$ ,  $p<.01$ ). False belief to others task correlated with affirmative ( $r = .420$ ,  $p<.01$ ) and negative ( $r= .483$ ,  $p<.01$ ) counterfactuals, whereas false belief to self task correlated with negative counterfactuals ( $r= .356$ ,  $p<.01$ ) but was not affirmative counterfactuals ( $r=.181$ , n.s.).

The false belief to others task also positively correlated with the CNRep ( $r=.306$ ,  $p<.05$ ) but did not correlate with any of the measures of contrary-to-fact reasoning.

The yes syllogisms and the no syllogisms both correlated with the overall contrary-to-fact syllogisms ( $r=.804$ ,  $p<.01$  and  $r=.529$ ,  $p<.01$ ) but they did not correlate with each other ( $r=-.080$ , not significant).

Finally the false belief to self task positively correlated with the contrary-to-fact syllogisms requiring a no answer ( $r=.499$ ,  $p<.01$ ) but not with contrary-to-fact syllogisms requiring a yes answer ( $r=-.183$ , n.s.) nor with the overall contrary-to-fact syllogisms( $r=.143$ , n.s.).

The CNRep only correlated with two other measures; the counterfactual task ( $r=.376$ ,  $p<.01$ ), including affirmative ( $r=.375$ ,  $p<.01$ ) and negative ( $r=.280$ ,  $p<.01$ ) questions, and the false belief to others task ( $r=.306$ ,  $p<.05$ ).

There was no significant correlation between counterfactual tasks and contrary-to-fact syllogisms.

In order to determine whether counterfactuals predicted theory of mind independently of verbal abilities a series of multiple regression analysis were conducted.

Firstly the potential predictors of false belief to others task were examined. A regression showed that counterfactual thinking predicted 27.3% of the performance on false belief to others independently of verbal memory (Beta= .522,



$t=4.461$ ,  $p<.05$ ,  $R^2 = .273$ ). When included contrary-to-fact in the model it showed that the two predictors combined only predicted 28.1% of the performance on false belief task to others showed that contrary-to-fact were a weak predictor (Beta= .507,  $t<.05$ ,  $R^2 = .281$ ).

Secondly the predictors of false belief to self were studied. A first model showed that counterfactuals predicted 10.3% of the performance of false belief to self task independently of verbal memory (Beta= .317,  $t=2.382$ ,  $p<.05$ ,  $R^2 = .103$ ). A second model studying the impact of counterfactual thinking and no-syllogisms showed that counterfactuals predicted only 8.9% of the performance of false belief to self task (Beta= .304,  $t=2.706$ ,  $p<.05$ ,  $R^2 = .089$ ) and no-syllogisms 24.5% (Beta= .499,  $t=4.192$ ,  $p<.001$ ,  $R^2 = .245$ ).

The pattern of correlations suggested that there might be two underlying factors, therefore an exploratory factor analysis was conducted to explore further the relationship between the different measures. As shown in

Table 14, the rotated component measure identified two factors. These results showed that the false belief to others task and false belief to self task have different predictors. The false belief to others, affirmative counterfactual and negative counterfactual measures loaded highly on factor 1, whereas the false belief to self and no-syllogisms measures loaded highly on factor 2.

Table 14: Rotated factor matrix in factor analysis

	Component 1	Component 2
CtoF (yes)	-.539	.124
CtoF (no)	.152	.874
FBT Others	.755	-.053
FBT Self	.531	.668
CF aff	.751	-.176
CF neg	.743	-.137
CNRep	.518	-.284

The correlation, the regression and the factor analysis gave very similar outcomes. They all suggested a link between false belief to others and counterfactuals (affirmative and negative) on one side and a different link between false belief to self, counterfactuals (to lesser extent) and no-syllogisms on the other side. They also suggested a lack of link between counterfactuals and contrary-to-fact syllogisms.

## 4 Discussion

The aim of this experiment was to investigate the relationship between theory of mind to others tasks, theory of mind to self tasks, counterfactual tasks, contrary-to-fact syllogisms and verbal memory. Correlations were run to investigate these links between the tasks, and regressions and a factor analysis were also computed to investigate those links further.

The results showed that counterfactual tasks correlated both with the false belief to others task and false belief to self task and that counterfactuals predicted the performance on both these tasks. The correlation found between the false belief to others tasks and counterfactual tasks is consistent with findings previously reported in the literature (Grant et al., 2004; Guajardo and Turley-Ames, 2004; Riggs et al., 1998; Peterson and Bowler, 2000). It supports the explanation

presented by Riggs et al. (1998) that there is a counterfactual element embedded in the false belief to others task. In the Sally and Anne task, for example, the counterfactual embedded element is implicit and would be 'what if Anne had not transferred the marble, where would the marble be?' In both tasks children need to ignore the real situation in front of them and reason counterfactually to answer the test question correctly. The false belief to others task correlated similarly with affirmative and negative counterfactual questions. Most research papers (e.g. Grant et al. (2004), Peterson and Bowler (2000), Riggs et al (1998)) unintentionally only explored the link between theory of mind and negative counterfactuals. The present experiment seems to indicate that there is a link between the false belief to others task and both affirmative and negative counterfactual questions. When comparing both types of counterfactual questions it also appeared that there was no difference in performance between the two. This contradicts the results reported in the first two experiments of this thesis. This difference in results could be due to the fact that the tasks used were different. In experiments 1 and 2 the task involved very little language and relied on the children classifying objects and pictures according to one of their physical properties (e.g. colour, shape, pattern...) and then answering a counterfactual question regarding a change of property and therefore of classification. These tasks did not involve a lot of narrative. It should also be noted that the false belief tasks and the counterfactual tasks were very similar, only the test question varied. In the present experiment, the counterfactual tasks designed by Harris et al. (1998) were used. These tasks involved following a short narrative story and then answering some questions. A bias might have been introduced in this series of tasks. The children had to answer two affirmative and two negative counterfactual questions. One of the affirmative counterfactual questions and one of the negative counterfactual questions required the children to point to a particular location and had a very

similar structure; however, the other two questions differed slightly. The affirmative question required a no answer and the negative question a yes answer. This difference could have influenced the results by increasing the difficulty of the affirmative counterfactual question, as usually children find giving a no answer more difficult than a yes answer.

In the present experiment, it seems that children performed at the same level on false belief to others tasks and counterfactual tasks. A possible explanation could be that the ability to solve both these tasks develops at the same time in parallel. This contradicts the previous findings reported in the literature. Peterson and Bowler (2000) found that children with autism and children with severe learning disability performed better on the counterfactual task than the false belief task, Grant et al. (2004) reported the same finding for children with autism and typically developing children.

The results also showed that counterfactuals correlated with the false belief to self task and that the former predicted 32% of the performance of the latter. This result does not support Riggs et al.'s (1998) claim that there is no counterfactual thinking involved in the false belief to self task. On the contrary, this result suggests that there might be a counterfactual element embedded in the false belief to self task. In the Smartie task, for example, the content of the box is swapped with a pencil before the start of the experiment. The children are asked what they think is in the box. They are then shown the actual content of the box and finally they are asked to recall their previous false belief. The children have to report a counterfactual situation: their false belief is different from the actual content of the box. It could be argued that a possible embedded counterfactual question could be: 'What if I had not swapped the content of the box, what would be inside?' The results also showed that children performed better on the counterfactual task than the false belief to self task. This could suggest that

counterfactual thinking might be a pre-requisite for the false belief to self task. A similar pattern was not found in the present experiment in relation to the false belief to others task. It is also interesting to note that when looking at the counterfactual task in more detail, it appears that false belief to self task correlated with the negative counterfactual questions and not the affirmative counterfactual questions.

Both false belief tasks positively correlated with each other. This correlation is probably due to the theory of mind element of the task. The results did not show any difference in performance between the two. This is consistent with the meta-analysis conducted by Wellman et al. 2001) that did not find any difference either.

The correlation between the false belief to others task, false belief to self task and counterfactual task could be explained by the theory of modified derivation reported by Peterson and Riggs (1999). They described strategies allowing us to make predictions about inaccessible systems. Solving the three tasks mentioned above involves a system that we have no real knowledge of, as for example Sally's false beliefs, our previous false belief or what would happen if Carol had taken her shoes off. This information is not readily available to us, therefore we need to adapt and modify the system available to us (in this case current reality) to infer what could happen. Peterson and Riggs (1999) reported a system called modified derivation (MD) as a process requiring modifications of the process of answer-derivation. Through this system we can ignore or impose facts. This system usually requires overriding conflicting facts in the database. They described this system of modification for the false belief to others task, false belief to self task and counterfactual task and they found some similarities. The tables they presented could be adapted as follows for the present data.

**Table 15 Subtractive MD over the false belief to others task**

Instruction (Sally)	Ignore (The marble was moved from the basket to the box at t2)
Query	Where does Sally think the marble is?
Base-fact	The marble was placed in the basket at t1.
Norm	Things generally stay where they are put
Defeating fact	The marble was moved to the box at t2
Answer (Sally)	In the basket

**Table 16. Subtractive MD over the counterfactual task**

Instruction	Ignore (Carol has her shoes on)
Query	Would the floor be dirty?
Base-fact	Carol took her shoes off.
Norm	When we keep our shoes on the floor generally gets dirty.
Defeating fact	Carol has her shoes on.
Default answer	Floor is not dirty

**Table 17. Subtractive MD over false belief to self task.**

Instruction	Ignore (the tube contains pencils)
Query	What is in the tube?
Base –fact	It's a Smarties tube
Norm	Containers generally contain the thing indicated on the outside
Defeating fact	The tube contains pencils
Answer	Smarties

In these three cases the reasoning processes are very similar: the children are instructed to ignore a fact embedded in reality. Children need to deal with conflicting representations and ignore the defeating fact. Peterson and Riggs (1999) argued that in order to solve a false belief task children need to understand other people's mental states but they also need to be able to perform modified

derivation. They argued that children under four years of age are not able to perform modified derivation and that therefore the lack of success on the false belief task before this age is due to that fact.

The false belief to others tasks in this experiment did not correlate with contrary-to-fact syllogisms. This strongly suggested that the correlation found between the false belief to others task and counterfactual tasks was probably due to the counterfactual element of the task rather than more general reasoning abilities.

For the contrary-to-fact syllogisms the twelve participants who showed a yes bias and the two who showed a no bias were not included in the analysis. The experimenter believed that children presenting a bias should be removed from the data set as there was no evidence that they were actually processing the questions. It was also believed that children who did not understand what was asked from them tended to answer yes, usually by nodding, to every question they were asked. These children would therefore have presented a yes bias. It was thought that the participants who presented a no bias might have answered no to every syllogism as a rejection of the first contrary-to-fact premise. More children showed a yes bias than a no bias which is consistent with what previous studies have found (Leevers and Harris, 1999; Leevers and Harris, 2000). These papers also reported excluding the participants showing a yes or a no bias from the analysis.

The correlation matrix showed a positive correlation between the no contrary-to-fact syllogisms (No C.to.F) and the false belief to self task, and the regression showed that no C.to.F predicted 50% of the performance of the false belief to self task. This was the only significant correlation for the contrary-to-fact syllogisms. The overall score for contrary-to-fact syllogisms, and for the yes syllogisms did not correlate with any of the other tasks. This suggests that this task is different from



the other tasks included in this study. This lack of correlations could be explained by looking at what needs to be ignored by the children. The modified derivation process (Peterson and Riggs, 1999) described previously suggested that for both false belief tasks and counterfactual tasks children need to ignore only one particular fact. The children, however, need to hold their beliefs of the real world to solve these tasks. With the contrary-to-fact syllogisms the children need to ignore the real world altogether. The first contrary-to-fact premise requires children to ignore what they know about the real world. This claim was supported by Markovits (1995) who explained that when the children were presented with contrary-to-fact premises in a fantasy context their performance increased, as the fantasy context acted as a cognitive filter preventing access to long term memory, i.e. knowledge of the real world.

Other analyses have shown that yes and no syllogisms differed significantly. The analyses showed that the children performed better on the syllogisms requiring a yes answer than the ones requiring a no answer. Although Leever and Harris (1999, 2000) used both types of contrary-to-fact syllogisms they did not report whether there was a difference between the two types so no comparison can be made with their findings. A detailed analysis of the syllogisms could give an explanation as to why the no-syllogisms were more difficult than the yes-syllogisms. Here are the two examples that are going to be examined. A yes-syllogism could be: 'All cows quack, Suzie is a cow. Does Suzie quack?' In this case the children need to accept the contrary-to-fact premise as a new rule for this fantasy world for the time it takes to solve the syllogism. The question follows from the first two premises and is still contrary-to-fact. The child could also just substitute 'cows' in the first premise with 'Suzie' to come to the correct answer. An example of a no-syllogism could be: 'All cats bark, George is a cat. Does George meow? In this case the first and the second premises have the same role as in the

yes-syllogisms but this time the question does not follow. This question is not contrary-to-fact; it refers to the real world the children know: cats do meow in reality. It is therefore possible that the question introduces some confusion in that particular situation; the children do not know whether they should reason in a real world or a contrary-to-fact situation. Although syllogisms are deductive tasks, people tend to make empirical biases (Evans et al., 1993). This point could explain the correlation with the false belief to self task. It could be argued that in order to answer the no syllogisms children require a stronger cognitive filter as reality is stated in the final question. It could be suggested that children who can answer the no-syllogisms correctly have a better ability to deal with contradictory representations and perform better on the false belief to self task. To some extent the false belief to self task can be seen as more embedded in the real world and real world knowledge. In the real world a Smartie container contains Smarties and not pencils. It is different in some respect from the false belief to others task and the counterfactual tasks. In the Sally and Ann task, for example, the ball can potentially be in either the box or the basket. Both possibilities have same probability of happening in the real world. The no-syllogisms and the false belief to self might have higher inhibitory demands as they ask the children in the first case to inhibit their knowledge of the real world while a question consistent with reality has just been asked to them (eg: Does George [cat] meow?), and in the second case they also need to inhibit their knowledge of the real world.

In order to check this hypothesis regarding the no-syllogisms and the necessity for a stronger cognitive filter the no-syllogisms could be worded differently. The first two premises could be identical and the question could be 'Does George moo?' It would still require a no-answer but it is now contrary-to-fact. The nature of the no-questions can also be questionable in a different way: Does 'bark' mean 'not meow'? The experiments are designed having in mind that

'bark' is exclusive of 'meow'. For example in the syllogism 'All snow is black. Len is a snowman made of snow. Is Len white?' the question implies that white is the equivalent of not-black. However, some syllogisms are more ambiguous. For example the question 'Does Susie moo?' uses the verb 'moo' as exclusive of 'bark', but it could be possible that a cow could quack and moo as they are different actions that could happen at different times. It would therefore be less confusing to use syllogisms where the two possibilities are mutually exclusive.

Looking at the relationship between other types of tasks it is interesting to note that no correlation was found between counterfactual reasoning and contrary-to-fact syllogisms. Although these two tasks involve suppressing knowledge of current reality and imagining a different situation they involve two different sets of cognitive abilities. It seems that counterfactual reasoning and syllogistic reasoning are different types of reasoning.

This study also included a memory measure, the CNRep. The CNRep positively correlated with counterfactual tasks and false belief to others tasks, but not very strongly. For both these tasks memory seems to be an important factor as the children need to remember what the situation was like at the beginning and it is now or what it could be now. A very small and non significant correlation was found with the false belief to self task. No correlation was found with contrary-to-fact syllogisms. This could be explained by the fact that the main ability contrary-to-fact syllogisms require is not so much remembering information but being able to suppress knowledge of the real world.

In this experiment theory of mind was assessed using two different types of task: the false belief to self task and false belief to others task. These types of task could be thought of as being equivalent as they are both supposed to evaluate whether an individual can attribute false beliefs and as they seem equally difficult (as reported by Wellman, 2001). They are to some extent related as they did

correlate with each other in our study. They both involve attributing false belief and suppressing one's knowledge of the current state of affairs. In the false belief to others task, the children need to suppress this knowledge, for example, that the marble really is in the box, rather than in the basket. In the false belief to self task children need to suppress their knowledge that there is a pencil a Smartie box rather than Smarties. Both also correlated with the counterfactual task, suggesting that they include a counterfactual element, but they are also different in some respects as they correlated in different ways with the contrary-to-fact syllogisms.

In conclusion, this experiment showed an interesting pattern. It confirmed a possible association between false belief to other tasks, false belief to self tasks and counterfactual thinking but suggested a dissociation between these tasks and contrary-to-fact syllogistic reasoning. This experiment showed that the correlation between false belief tasks and counterfactual tasks was probably due to the counterfactual element of the tasks rather than general reasoning abilities. The results also showed a link between contrary-to-fact syllogisms (especially the ones requesting a no answer) and the false belief to self task. This study also highlighted the difference between counterfactual thinking and contrary-to-fact thinking. This difference will be studied further in the next chapter.

## **CHAPTER 5**

### **EXPERIMENT 4: THE EFFECT OF CONTEXT ON COUNTERFACTUALS AND CONTRARY-TO-FACT SYLLOGISMS**

#### **1 Introduction**

The previous chapter investigated the relationship between several cognitive tasks: false belief to self tasks, false belief to others tasks, counterfactual tasks, contrary-to-fact syllogisms and verbal memory. The results from the experiment showed a link between false belief to others tasks, false belief to self tasks and counterfactual thinking on one side, and another link between false belief to self tasks and contrary-to-fact no-syllogisms on the other side. A link was also found between false belief to self tasks and contrary-to-fact syllogisms (no) which suggested that both tasks involved similar cognitive processes; both these tasks could require a high inhibition. On the other hand, the findings also showed that counterfactual thinking and contrary-to-fact syllogisms did not correlate which suggested that different cognitive processes underlie each type of reasoning. The purpose of the present chapter was therefore to investigate this difference further.

These two tasks differ in two respects: their structure and their content. In the particular case of the previous chapter and the present one the syllogisms have a contrary-to-fact content whereas the counterfactual tasks have a factually true content. Contrary-to-fact syllogisms, as defined in chapter 1 and chapter 4, include premises whose content is factually false (Markovits and Vachon, 1989). The content of these contrary-to-fact premises runs against the participants' beliefs and their knowledge of the real world. These premises can be statements such as 'all cats bark' (Dias and Harris, 1998) or 'if an object is placed in boiling water then it will become cold' (Simoneau and Markovits, 2003). According to Markovits (1989) reasoning with contrary-to-fact premises implies two abilities.

First of all, the children need to be able to accept the contrary-to-fact premises as a basis for reasoning without interference from their empirical knowledge, and secondly children need to be able to reason formally. He explained that most mistakes children made in their study were inversion mistakes where they used their real world knowledge rather than their reasoning skills to solve the task.

Contrary-to-fact syllogisms have been used in many studies with children to assess the impact of content on reasoning and it was found that children had difficulty correctly solving these tasks because of the unfamiliar content (see chapter 1). This difficulty with contrary-to-fact content was found with preschool children (eg: Leever and Harris, 2000), but also with teenagers up to the age of 18 (Markovits and Vachon, 1989). One of the first research papers on contrary-to-fact premises was published by Hawkins, Pea, Glick and Scribner (1984). They presented the children with syllogistic problems with three different types of content: fantasy, congruent or incongruent with practical knowledge. Their findings confirmed that young children can reason deductively and also showed that problem content has an effect on the reasoning performance of the children. When children were asked to solve contrary-to-fact problems their performance was lower than when they tried to solve problems congruent with their practical knowledge of the world. Subsequently the effect of contrary-to-fact premises has been studied in two main different reasoning paradigms: with syllogisms (e.g.: Dias and Harris, 1988) and with conditional problems (e.g.: Markovits and Vachon, 1989); with many of these studies focusing mainly on the potential factors that could facilitate this type of reasoning. The main factor that has been studied is the effect of the context in which the contrary-to-fact premises are presented on performance (e.g.: Dias and Harris, 1990). It has been found that a variety of different contexts, especially fantasy context, could improve performance, and two main hypotheses offering an explanation for this effect have been presented in the

literature. Leever and Harris (1999; 2000) argued that a fantasy context, as other non standard contexts, would give useful pragmatic clues to the children encouraging them to accept the contrary-to-fact premise as a basis for reasoning, whereas Markovits (e.g.: Markovits, 1995, Markovits et al., 1996) offered a more cognitively orientated explanation. He suggested that a fantasy context would act as a cognitive filter reducing access to long-term memory and therefore to empirical knowledge then preventing inversion mistakes. The data from which both these theories were constructed is now going to be reviewed.

Dias and Harris (1988) started this line of work on instructions, and pragmatic clues in particular, by giving children syllogisms including facts agreeing with their knowledge or facts contrary to their knowledge of the world. They used two different conditions to present the tasks: the verbal condition or the play condition. In the verbal group the children were read aloud standard instructions introducing the syllogisms, whereas in the play group condition the contrary-to-fact premise was acted out with toys. The children were then asked to answer the syllogisms and justify their answer. They found that children in the play group gave more accurate answers than in the verbal group for the contrary-to-fact syllogisms, and that they also provided more theoretical justifications and less empirical justifications for their answers. The play condition was a good help for the children when answering those questions.

Intonation was studied as a make-believe clue by Dias and Harris (1990). They presented contrary-to-fact syllogisms to the children in either a make-believe intonation where the experimenter introduced dramatic emphasis into her voice, or a normal matter-of-fact intonation. The results showed that using a make-believe intonation increased the number of correct answers and theoretical justifications. In the same experiment Dias and Harris (1990) asked the children to make a picture in their head of the contrary-to-fact premises, for example 'a pig that flies'.

They found as previously that using imagery increased the number of correct answers and the number of theoretical justifications. The effect of imagery has been further studied by several authors.

Leevers and Harris (1999) investigated whether the influence of pragmatic clues would carry over in time. They presented contrary-to-fact syllogisms in four different contexts: basic, thinking, imagery, or imagery and cognition. The basic group was a standard context without any further instructions. In the thinking group the children were encouraged to think about the premises. In the imagery group the children were asked to form a mental image of the premises. Finally in the imagery and metacognition group the children were encouraged to use imagery and were told that using imagery would be beneficial. The children in the three instructed groups performed better than children in the basic group and the effect carried over for a week. This showed again the positive effect of instructions in general and imagery in particular.

Contrary-to-fact syllogisms have also been studied with clinical populations. Leevers and Harris (2000) replicated their previous experiment with children with autism, children with learning disabilities and typically developing children matched for verbal receptive abilities by using the TROG. The children were tested under two conditions: the basic group and the instruction group where children were asked to use imagery. The results confirmed that the use of imagery improved performance for all the three groups, and that this effect carried over for 2 or 3 weeks. However, the children with autism still performed at chance level and, instead of giving consistent logically or empirical justifications they used an unsystematic mixture of justification. The children with autism did not give more logical responses, in contrast to the other groups, and they tended to display a yes bias. Scott et al. (1999) studied the effect of context on contrary-to-fact syllogisms with autistic children as well and compared them to two control groups matched for



verbal mental age: typically developing children and children with moderate learning difficulties. The syllogisms were presented either in a standard context or in a context where the children were asked to imagine the contrary-to-fact premise in their head and were encouraged to use their imagination. Both control groups performed as expected and improved their performance in the imagery condition. However, the children with autism performed more poorly in that condition. The authors suggested that these results could be due to the fact that children with autism have difficulty pretending. This could also be due to the fact children with autism tend to show a 'yes' bias, as shown by Leavers and Harris (2000), and that in this experiment all the syllogisms used required a 'yes' answer.

The last type of instructions that will be discussed is fantasy context. Dias and Harris (1990) presented the contrary-to-fact syllogisms to children in two different contexts: 'the planet setting' and 'the non-planet setting'. In the planet setting the children were told that the experimenter was from a different planet where everything is different and he is going to tell them some stories about this planet. This context could be assimilated to a fantasy context. The 'non-planet' setting is just a standard context. The results showed that setting the story on a different planet increased the number of correct answers and theoretical justifications.

All the studies reported so far consistently showed that non standard instructions such as play (Dias and Harris, 1988), make belief intonation (Dias and Harris, 1990), imagery (Dias and Harris, 1990; Leavers and Harris, 1999; Leavers and Harris, 2000; Scott et al., 1999), and fantasy (Dias and Harris, 1990) improved children's performance with typically developing children as well as children with learning disabilities. Leavers and Harris (1999; 2000) argued that this effect occurred because these instructions provided a pragmatic clue regarding the study. They signalled to the children that they needed to reason with the contrary-to-fact premises. The children were no longer confused by the situation; they no

longer wondered whether they were tested on their reasoning skills or their knowledge of the world. Harris (2000) claimed that non-standard instructions act as a cue to encourage the children to focus on the initial contrary-to-fact premise leading them to accept this premise as a basis for reasoning. He supported his argument by saying that the effect of instructions was shown on contrary-to-fact premises as well as unfamiliar premises (Dias and Harris, 1988) as well as made-up premises (Leevers and Harris, 1999) where no interference of real world knowledge could have interfered.

Markovits and Vachon (1989) started another trend of studies with reasoning tasks including contrary-to-fact premises. They studied fantasy context in relation to contrary-to-fact premises included in conditional problems with children older than the ones tested in all the previous experiments discussed. They found that children had difficulties with these tasks and tended to make many errors. They described the errors children made with contrary-to-fact reasoning problems as 'inversion errors'. These errors are made when children have difficulty in accepting empirically false premises without interference from their real world knowledge. Their hypothesis was then that if the contrary-to-fact problems were presented in a fantasy context then there would be less interference from empirical knowledge as the children would be able to construct a more coherent representation of the premises. Children should be able to accept contrary-to-fact premises more easily if they are set in a fantasy context. They tested children ages 5 and 7 using conditional problems including a contrary-to-fact premise; the problems were either set up in a simple or a fantasy context. In the fantasy context the children were told about an imaginary country where curious things happened. The results showed that again fantasy context improved performance for both age groups. They concluded that presenting contrary-to-fact premises in a fantasy context helps the children accept the premises without interference from the real

world. This supports the hypothesis that accepting contrary-to-fact premises 'involves creating a coherent representation that does not incorporate empirical knowledge' (p.410). In order to explain this effect in more depth Markovits conducted a series of studies including different types of reasoning problems.

Markovits (1995) conducted an experiment with even older children, aged from 12 to 14, and presented them with contrary-to-fact conditional problems in either a realistic or a fantasy context. The children were presented with either modus pollens (MP) problems or affirmation of the consequent (AC) problems. The results showed, as previously, that fantasy context reduced the number of inversion errors, therefore reduced inference from empirical knowledge. Markovits (1995) explained that a fantasy context acts as a cognitive filter reducing access to long term memory. Subsequently the participants are less likely to produce a complete model of the major premises and it therefore reduces the number of correct answers to the invalid form (Affirmation of the consequent).

In order to investigate further the hypothesis that fantasy context acts a cognitive filter Markovits et al. (1996) investigated the effect of fantasy context on factually true premises. They tested 200 schoolchildren from the age of 7 to 13. The children watched some videotapes telling them stories. One story was set in a standard context and the other story in a fantasy context. Four conditional questions were asked of the children (MP, MT, AC, DC). The results showed that context had no effect on MP and MT, and that fantasy context made the task significantly more difficult for AC and DC, the uncertain forms. The results also showed that the pattern of justifications for MP was different depending on the context. In the standard context the children gave more justifications involving knowledge. This is consistent with Markovits' theory that fantasy context acts as a cognitive filter blocking access to long term memory. This study showed that presenting true and familiar premises in a fantasy context reduces the children's

performance on uncertain logical forms (AC and DA). It seems that a fantasy context limits how children use their real-world knowledge when solving reasoning problems (Markovits et al., 1986).

Markovits' hypothesis regarding a cognitive filter could be tested using counterfactual tasks. In order to correctly answer a counterfactual task children need to undo an event and imagine what the present situation could be. In the context of this thesis the counterfactual task included only factually true premises. Solving a counterfactual task requires accessing knowledge of the world. For example in one the scenarios used in the previous chapter the counterfactual questions was 'What if Carol had taken her shoes off, would the floor be dirty?' In order to answer this question the children need to imagine that Carol took her shoes off but they need to access their knowledge of the real world, in that case they need to know that when children do not wear dirty shoes the floor stays clean. If fantasy context acts as a cognitive filter reducing access to long term memory it therefore implies that counterfactual tasks presented in a fantasy context will be more difficult for the children.

The aim of the present experiment was to investigate the difference between contrary-to-fact syllogistic reasoning and counterfactual problems, and to explore the role of fantasy context on each on these tasks in relation to the cognitive demands of each task. According to the literature discussed above we can predict that fantasy context will facilitate contrary-to-fact syllogistic reasoning (e.g.: Dias and Harris, 1989; Leever and Harris, 2000). The fantasy context should act as a cognitive filter and help children reason in isolation of their real world knowledge. For the same reason the effect of fantasy should be different on counterfactual reasoning. In order to solve a counterfactual question children need to ignore one element of the state of affairs but still need to use their real world knowledge, therefore it is predicted that it will increase the difficulty of the task if a cognitive

filter prevents them from accessing real world knowledge. This experiment should provide more knowledge regarding the difference between these two types of tasks. It should help us understand why they did not seem to correlate in the previous experiment and why they correlated with different tasks.

## **2 Method**

### **2.1 Participants**

Seventy three children, including 47 girls and 26 boys, took part in this experiment. Forty one of them attended a nursery and 32 a reception class in Plymouth.

The children were aged from 3;3 to 5;8 (mean: 4;7 – standard deviation: 8 months).

### **2.2 Material**

The children were presented with four counterfactual scenarios and four contrary-to-fact syllogisms either in a standard or a fantasy context. The context was counterbalanced between participants with 37 participants tested in the fantasy context and 36 in the standard context. The counterfactual scenarios and the contrary-to-fact syllogisms were randomised for each participant by using a randomisation software ([www.randomizer.com](http://www.randomizer.com))

#### **2.2.1 Counterfactual scenarios**

The counterfactual scenarios were acted out with puppets and toys and were very similar to the ones designed by Harris et al. (1996). Just a few details were changed because of physical constraints. This is one example of the scenarios acted out in the standard context:

*I'm going to tell you some stories. Listen carefully and then I'll ask you some questions. Are you ready? This is the first one.*

*One day in that forest this tree was standing up like that. (The experimenter puts one small plastic tree standing up on the table). But guess what! There was a storm and the wind blew the tree over like this. (The experimenter knocks the tree down).*

The children were then asked two control questions: the Now Control Question in relation to the final situation (e.g.: 'Is the tree standing up now?') and the Before Control Question in relation to the initial situation (e.g.: 'Was the tree standing up before?'). Finally the children were asked the test question (E.g.: 'What if the wind had not come – would the tree be standing up?').

Three other similar scenarios using different props were used (Appendix J). Altogether two episodes requested yes/no answers and the other two a selective pointing answer. Two of the test questions were negative (e.g.: What if the wind had not come, would the tree be standing up?) and the other two positive (e.g.: What if Lucy had taken her shoes off – would the floor be dirty?)

One test question required a no answer, one a yes answer and the last two a selective pointing (i.e.: the children needed to point to a particular location to answer the question) in order to avoid any bias in the answers.

In the fantasy context the structure stayed the same although the introduction differed and was as follows:

*I now want you to imagine that a wizard have given you a magic carpet. He tells you that the carpet will fly you to lots of mysterious worlds in the sky. He also says that when you reach these strange places there will be someone there to meet you and tell you about the place where they live. So you get on the carpet and it flies off into the clouds.*

The introduction to each scenario was also slightly different. The first sentence reinforced the fantasy context. The example given above was as follows in the fantasy context:

*You fly over the rivers and you arrive in an enchanted wood. You meet a little boy who tells you that in this forest all the trees can walk like this (makes the tree walk). One day in that forest this tree was standing up like that. But guess what! There was a storm and the wind blew the tree over like this.*

The actual scenarios and questions were the same as in the standard context.

## 2.2.2 Contrary-to-fact syllogisms

The contrary-to-fact syllogisms were the same as the ones used by Leevers and Harris (2000). Four of the syllogisms were selected. In order to avoid any systematic yes or no answers two syllogisms were correctly answered by affirmation (yes problems) and the other two by denial (no problems).

The syllogisms were as follows:

- 1) All cows quack. Suzie is a cow. Does Suzie quack? (yes)
- 2) All zebras have spots on their back. Alex is a zebra. Is Alex spotty? (yes)
- 3) All snow is black. Len is a snowman made of snow. Is Len white? (no)
- 4) All hedgehogs feel soft. Harry is a hedgehog. Does Harry feel prickly? (no)

The contrary-to-fact syllogisms were presented either in a standard context or a fantasy context. The standard context just set the scene, using the same introduction as Leevers and Harris (1999) as follows:

*'Now we're going to think about little stories together. Some of the things in the stories may sound a bit funny, but we are going to think about what things would be like if all the things in the stories were true.'*

In the fantasy context the introduction was the same as the one reported above for the counterfactual scenarios.

## 2.3 Design

A mixed design was used to investigate a possible interaction between the two types of tasks (counterfactual and contrary-to-fact) and context (standard/fantasy). The tasks were within participant and the context between participants. The children were presented with four counterfactual task and four contrary-to-fact syllogisms. The tasks were randomised. Half of the participants were presented with these tasks in a standard context and the other half in a fantasy context.

## **2.4 Procedure**

The participants were tested individually at the school or the nursery they attended over one session lasting between 10 to 15 minutes.

## **3 Results**

The data of the counterfactual problems and the contrary-to-fact syllogisms can be analysed in two different ways according to whether the participants who did not answer the control questions of the counterfactual scenarios correctly or showed a bias when answering the contrary-to-fact syllogisms are included or not in the analysis. The first part of this results section, as in the previous chapter, will include all the data for the counterfactual questions regardless of whether children answered the control questions correctly but will exclude the data from the participants who showed a yes or a no bias when answering contrary-to-fact syllogisms.

The score for each participant was the total number of right answers for the counterfactual scenarios or the contrary-to-fact syllogisms. Participants could score a maximum of 4 points for each type of task. The average score and the standard deviation were computed for each type of task in each context: counterfactual scenarios in standard context, counterfactual scenarios in fantasy context, contrary-to-fact syllogisms in standard context, and contrary-to-fact syllogisms in fantasy context. The results are presented in Table 18.



Table 18. Mean and standard deviation of the scores for each type of tasks in the two different contexts.

	Standard context (n= 28)	Fantasy context (n=27)
Counterfactual tasks	2.28 (1.48)	1.77 (1.15)
Contrary-to-fact syllogisms	.85 (.75)	1.26 (1.19)

A mixed design ANOVA 2\*2 (context \* task) was computed and showed a significant main effect of task,  $F(1, 71) = 19.23, p < .01$ .

The test of within-subject contrast showed that this difference was statistically highly significant,  $F(1, 71) = 19.23, p < .01$ . The results suggested that the participants found the counterfactual task easier than the contrary-to-fact syllogisms independently of the context.

The test of between-subjects showed that this effect was not significant,  $F(1, 71) = .054, p = .817$ . There was no effect of context on its own.

The statistical analysis showed a significant interaction between task and context,  $F(1, 71) = 4.20, p < .05$ . This result indicated that the context did not have the same effect on each type of task. It appears that fantasy context facilitates contrary-to-fact reasoning but increases the difficulty of counterfactual reasoning.

As explained at the beginning, the data can be analysed in a different way by excluding the participants who did not answer correctly at least one control question. The results for the interaction did not differ in any meaningful way. These results are presented in appendix K. The interaction was not strictly significant with this reduced sample. This difference might be due to the lack of power when the sample size is reduced.

The tasks were also studied individually and the results were as follows.

The previous chapter showed that children found the yes syllogisms easier than the no syllogisms so the same comparison was examined in the present experiment. The mean and the standard deviation were calculated for each type of syllogism and the results are presented in Table 19.

Table 19.Means and standard deviation for yes and no syllogisms.

	Mean	Standard Deviation
Yes Syllogisms	.56	.65
No Syllogisms	.56	.71

The means were identical and the standard deviations were similar. This experiment did not identify a difference between the yes and the no syllogisms, which was not the case in the previous experiment.

The difference in performance between affirmative and negative counterfactuals was also investigated. The mean and the standard deviation were computed for each type of counterfactual questions and the results are presented in Table 20.

Table 20.Means and standard deviation for each type of counterfactual question.

	Mean	Standard Deviation
Affirmative Counterfactual	.90	.74
Negative Counterfactual	1.01	.69

A paired sample t-test showed that this difference was not significant,  $t(73) = -1.380$ ,  $p = .172$ . This experiment did not identify any difference in performance between negative and affirmative counterfactual questions.

The previous chapter did not identify a correlation between counterfactual tasks and contrary-to-fact syllogism, so in order to compare the results with the present experiment some correlations were computed. The correlation matrix is presented in Table 21.

Table 21. Correlations.

	C.F.	Aff C.F	Neg. C.F	C.to.F	C.to.F (yes)	C.to.F (no)
C.F.	1 (74)					
Aff C.F.	.892** (74)	1 (74)				
Neg. C.F.	.875** (74)	.562** (74)	1 (74)			
C.to.F	-.015 (55)	-.048 (55)	.022 (55)	1 (55)		
C.to.F (yes)	-.083 (55)	-.070 (55)	-.077 (55)	.757** (55)	1 (55)	
C.to.F (no)	.171 (55)	.135 (55)	.170 (55)	.780** (55)	.435** (55)	1 (55)

\*\* Significance value  $<.05$

\*\* Significance value  $<.01$

No correlations were identified between counterfactual reasoning and contrary-to-fact syllogisms. However, on this occasion, there was a significant positive correlation between contrary-to-fact (yes) and contrary-to-fact (no),  $r=.435$ ,  $p<.01$ .

## 4 Discussion

The aim of the present experiment was to investigate the effect of fantasy context on contrary-to-fact syllogistic reasoning and counterfactual reasoning. It was hypothesised that fantasy context would not have the same effect on both types of task. It was expected that it would improve performance on contrary-to-

fact syllogisms and worsen performance on counterfactual scenarios. It was predicted that fantasy context would act as a cognitive filter preventing children from accessing their real world knowledge therefore it would help them solve contrary-to-fact syllogisms as they are a deductive tasks, but it would increase the difficulty of counterfactual tasks and children need to access their knowledge of the world to solve them.

The results showed a significant interaction between task and context suggesting that context influenced performance on these two types of tasks differently: fantasy context improved performance on contrary-to-fact syllogisms but decreased performance on counterfactual scenarios.

The effect of fantasy context on contrary-to-fact syllogisms found in this experiment confirmed the hypothesis that had been made. This result was also consistent with previous findings reported in the literature. The findings were similar to Dias and Harris' (1990). Dias and Harris (1990) presented contrary-to-fact syllogisms to preschool children in a fantasy context. They told them that the stories were taking place on a different planet where everything was different. Their results showed that fantasy context improved performance on contrary-to-fact syllogisms. The children showed fewer empirical biases and gave more logical justifications for their answers. In Dias and Harris' (1990) experiment the participants' mean age was 4 years 9 months, which was very similar to the mean age of the children in the present experiment (mean age: 4 years 7 months). They argued that the planet setting allowed children to reason in another world where the contrary-to-fact premises might be true; that it helped the children accept the contrary-to-fact premises as a basis for reasoning.

Markovits also conducted a series of studies with contrary-to-fact premises but in a different context, as the contrary-to-fact premises were included in causal scenarios rather than syllogisms, and they tested children from a different age

range: older children and teenagers. Markovits consistently found that presenting contrary-to-fact scenarios in a fantasy context helped children improve their performance (e.g.: Markovits, 1995; Markovits and Vachon, 1989; Markovits et al., 1996).

Two theories have been presented in the literature to try to explain the positive effect of fantasy context on contrary-to-fact reasoning tasks. Leever and Harris (2000) provided an explanation relating to instructions in general. They explained that instructions provide a pragmatic clarification regarding the experimenter's intention, with fantasy context being just one particular type of instruction. Instructions would help children understand that the experimenter wanted them to act as if the contrary-to-fact premises were true and then solve the syllogisms within that particular framework. However, Markovits (1995) argued for a different theory suggesting that the fantasy context acts as a cognitive filter that prevents access to long term memory. The children then cannot access their knowledge of the real world as easily so they make fewer empirical mistakes. In order to successfully solve contrary-to-fact syllogisms children need to ignore their knowledge of the real world and reason with only the information presented to them in the syllogisms. Syllogisms are deductive tasks that do not require using knowledge about the world, whether they are contrary-to-fact or true-to-fact. All the information necessary to solve the task is provided in the premises.

The interaction between context and task showed however, that context did not have the same effect on counterfactual tasks as it had on contrary-to-fact syllogisms. It seems that fantasy context decreases performance on counterfactual scenarios. In the present experiment children were presented with true-to-fact counterfactual scenarios, half of them were set in a standard context and the other half in a fantasy context, and they performed more poorly in the second condition. It is interesting to analyse these findings in relation to the two

theories mentioned above about fantasy context. Leevers and Harris (1999; 2000) argued that instructions help clarify the experimenters' intention. The counterfactual scenarios being true-to-fact a fantasy context would not really bring any clarification regarding the experimenter's intention. It could be argued that it might actually confuse the children who are then uncertain about whether the task should be solved using their empirical knowledge or not. This hypothesis could therefore be consistent with the finding that fantasy context make true-to-fact counterfactuals more difficult to solve for children.

The second possible explanation, Markovits' (1995) cognitive filter theory, could also bring a convincing insight into this effect and does seem consistent with the present findings. Markovits (1995) explained that a potential cognitive filter would reduce access to long term memory and therefore to knowledge about the real world. The cognitive filter would help children to avoid making empirical mistakes and would help them reason with only the information presented in the reasoning task. This hypothesis is consistent with the results found for the contrary-to-fact syllogism but how would it apply to the counterfactual tasks? In order to solve a counterfactual scenario, children need to ignore one event and imagine an alternative present to the actual one. In the first scenario for example the test question was 'What if Lucy had taken her shoes off – Would the floor be dirty?'. The children need to suspend their knowledge that Lucy kept her shoes on and imagine the consequence of taking her shoes off. They need to suppress one causal relationship (e.g.: keeping dirty shoes on and making the floor dirty) but hold all the other causal relationships of the real world constant. They need to access real world knowledge to know what would happen if Lucy did not leave her shoes on. The cognitive filter provided by the fantasy context would then prevent children from accessing this necessary knowledge to solve the task, and therefore the children would not perform as well. The data on the counterfactual tasks

presented in a fantasy context seems to support Markovits' (1995) hypothesis regarding the cognitive filter.

In summary in order to solve contrary-to-fact syllogisms children need to suspend their knowledge of the real world whereas in order to solve a counterfactual scenario they only need to suppress one particular event but refer to their knowledge of the world. The amount of knowledge to be suppressed could maybe explain the difference in performance found between these two tasks. The results did show that there was a main effect of task and that children performed better on counterfactual tasks than on contrary-to-fact syllogisms. However, it is difficult in this experiment to know whether this difference in performance was due only to the type of tasks (counterfactual scenarios versus syllogisms) or the contrary-to-fact element of the syllogisms, as all the counterfactual scenarios were true-to-fact and all the syllogisms were contrary-to-fact. For future research it would be interesting to compare performance on counterfactual scenarios and syllogisms controlling for contrary-to-fact and true-to-fact premises. The effect of fantasy could also be studied systematically in this context. Children could be presented with contrary-to-fact syllogisms and true-to-fact syllogisms in a fantasy context and explore the effect of context. The contrary-to-fact syllogisms could be similar to the ones used in this experience and an example of a true-to-fact syllogism could be as follows: 'All cows moo, Suzie is a cow, does Suzie moo.' Counterfactual thinking could be studied in the same way. True-to-fact counterfactual scenarios could be similar to the ones used in the present experiment and an example of a contrary-to-fact counterfactual scenario could be as follows: 'I'm going to tell you a story that may sound a little bit strange. I would like you to imagine that everything in the story is true. When Carol walks in the house with muddy shoes it cleans the floor. One day the floor is really dirty. Carol comes home and she doesn't take her shoes off. She comes inside and makes the

floor all clean with her dirty shoes.' The children could then be asked the before control question (Was the floor clean before?) and the now control question (Is the floor clean now?). Finally a counterfactual question could be asked: What if Carol had taken her shoes off? Would the floor be dirty? The same scenario could also be presented in fantasy context elaborating on the magical world where Carole lives, and telling the children that wizards and fairies live in this magical world and that everything in that world is different. This design would therefore allow us to compare counterfactual thinking and syllogistic reasoning and would also allow us to study the effect of fantasy context on contrary-to-fact premises on two different types of tasks. If Markovits' hypothesis is true some predictions can be made. As syllogisms are deductive tasks in nature not requiring any knowledge of the real world it should not have an impact of the true-to-fact syllogisms. Regarding the counterfactual scenario it can be hypothesised that including a contrary-to-fact element (e.g.: muddy shoes clean the floor) would increase the difficulty of the counterfactual scenario in comparison to a standard one. In order to solve this particular scenario children need to know what the causal relation is between 'muddy shoes' and 'cleanliness of the floor'. In the present case the elements given in the scenario conflict with real world knowledge, therefore to succeed children need to suspend their real world knowledge. Logically it could then be predicted that a fantasy context, which would act as a cognitive filter, would help them solve this task.

The present experiment provided some other results that are worth considering. First, there was still a clear lack of correlation between counterfactual tasks and contrary-to-fact syllogisms which replicated the findings reported in chapter 4. This confirms that contrary-to-fact syllogistic reasoning and counterfactual reasoning require different cognitive abilities. As mentioned above at this stage is it difficult to know if this difference is due to the task itself or the



nature of the content. Further research (as explained above) could help clarify this point.

Secondly, no difference in performance was found between contrary-to-fact syllogisms requiring a yes answer and contrary-to-fact syllogisms requiring a no answer. The previous experiment had shown that children found no-syllogisms more difficult than yes-syllogisms. In the present experiment, children only had to answer two of each whereas in the previous one they answered four of each. It might be necessary to increase the number of data points collected to find this difference again.

Thirdly, no difference in performance between negative counterfactual questions and affirmative counterfactual questions was found. This confirmed the findings reported in chapter four but not the finding reported in chapters 2 and 3. This particular difference between the experiments could be explained the nature of task. This point will be discussed in more detail in the final discussion of this thesis.

In conclusion, it seems that contrary-to-fact syllogism and counterfactual scenarios require different skills, but further research is needed to investigate whether this difference is imputable to the nature of the tasks or their content. This experiment also showed that fantasy context does not influence these tasks in the same way and the finding supported Markovits' claim that fantasy context acts as a cognitive filter.

# **CHAPTER 6**

## **DISCUSSION**

### **1 Introduction**

The following discussion will include three main sections. The first section will review the main goals of this research. The second section will review the four experiments; their hypotheses, their main findings; as well as the links between them. These findings will be discussed in relation to the previous literature. The third section will discuss the applications of the findings and some ideas for further research.

### **2 Main goals**

The main goal of this thesis was to investigate children's reasoning abilities, especially their abilities to solve counterfactual tasks. This study looked at three main aspects of counterfactual thinking: the nature of counterfactual thinking itself, its relation with the false belief tasks and how a fantasy context can influence children's ability to solve the tasks.

The very first goal at the beginning of the thesis was to investigate the influence of the polarity of counterfactual questions, that is to say the difference between affirmative and negative counterfactual questions. The work then concentrated on the investigation of children's reasoning abilities in relation to the false belief task. The false belief to others task, such as the Sally and Ann task (Baron-Cohen et al., 1985) or the Maxi task (Wimmer and Perner, 1983), was designed to evaluate children's understanding of mental states, but its components have been studied in more depth recently. One of the main aims of this thesis was to investigate the components of the false belief to others task in relation to reasoning, especially to counterfactual reasoning, as Riggs et al. (1998) claimed that there was a counterfactual element embedded in the false belief task.

This line of work has extended to the study of the false belief to self task such as the Smartie task (Perner et al., 1987). There is a consensus in the literature that children can pass the false belief task from the age of four, therefore for this thesis it was decided to study typically developing children from the ages of 3 to 5. Finally, this thesis looked at the influence of context, in particular fantasy context, on children's ability to solve a counterfactual task.

### **3 A review of the studies conducted in this thesis**

This thesis consists of four experiments each studying counterfactual reasoning and false belief tasks in relation to each other or in isolation.

The first experimental study followed directly from Riggs et al.'s (1998) work on the relationship between counterfactual thinking and the false belief task. This experiment studied the difference between counterfactual thinking and future hypothetical thinking and investigated the effect of negative versus affirmative polarity on the test questions. Riggs et al. (1998) argued that the correlation they found between counterfactual thinking and the false belief task was due to the counterfactual element rather than the conditional element of the counterfactual tasks. They made this claim after showing that counterfactual thinking differed from future hypothetical thinking, as the children seemed to find the former more difficult than the latter. However, a confound could be identified in their experiment as all the counterfactual questions were negative and all the future hypothetical questions were affirmative. Work previously conducted on the processes involved in counterfactual thinking in adults claimed that counterfactual thinking and hypothetical thinking did not differ (Evans and Over, 1996), and some work using negated conditional questions showed that negation decreased reasoners' performance (Evans, 1998). Therefore, the first experiment of this thesis compared performance on negative counterfactual questions, affirmative

counterfactual questions, affirmative future hypothetical questions and negative future hypothetical questions. Very similar tasks to the ones designed by Riggs et al. (1998) were used. The tasks required children to physically classify objects into three different boxes according to one of their physical properties (e.g.: shape, colour...) and then answer a test question. The results showed that when controlling for negative versus affirmative questions there was no difference in performance between counterfactual thinking and future hypothetical thinking: the children performed at the same level on affirmative counterfactual questions and affirmative future hypothetical questions. The results also confirmed that negation increased the difficulty of the questions: children performed better on affirmative counterfactual questions than on negative counterfactual questions. Children performed close to ceiling level on the negative future hypothetical questions as they did not have to process the question to give a correct answer. Therefore no conclusions could be drawn from the performance on this particular question.

The second experiment aimed to study further the effect of negation on counterfactual questions and the relation between counterfactual thinking and the false belief task. The same tasks as in the first experiment were used. The results confirmed one of the findings of the first experiment which was that negative counterfactuals were more difficult for children than affirmative counterfactuals. The false belief tasks used in this experiment followed the same model as the counterfactual task, just the end of the scenario and, of course, the test questions differed. In this false belief task, Tidy Teddy physically classified some objects in three different boxes according to one of their physical properties. At some point, Tidy Teddy left the scene and while he was away the experimenter moved one object from one box to another one. When finally Tidy Teddy came back, the children were asked where they thought Tidy Teddy thought his object was or where he was going to look for it. In this experiment, no significant correlations

were identified between counterfactual thinking and the false belief task. These results did not confirm Riggs et al.'s (1998) findings. Riggs et al. (1998) claimed that the correlation they found between counterfactual thinking and the false belief task was due to the fact that children could not suppress their knowledge of current reality and this difficulty led them to make realist errors. However, their design did not allow them to analyse these mistakes. The tasks that we used in this second experiment allowed us to distinguish between realist and other types of mistakes. It appeared that about a third of the mistakes children made were not realist errors. This showed that the difficulty with counterfactual thinking and false belief task was not exclusively due to a difficulty with suppressing knowledge of current reality.

The third experiment studied further the relationship between counterfactual thinking and the false belief task. The main rationale behind this experiment was to investigate whether a potential correlation between counterfactual thinking and false belief was due to a counterfactual element or more general reasoning abilities, therefore a set of contrary-to-fact syllogisms was also included in this experiment. Two types of false belief tasks were used: the false belief to others task and the false belief to self task. In this instance, overall, the experimenter used more traditional tasks, such as the Sally and Anne task (Baron-Cohen et al., 1985) and the Smartie task (Wimmer and Perner, 1983) to assess theory of mind, and tasks similar to the ones designed by Harris et al. (1996) to assess counterfactual thinking. The results showed an interesting pattern of correlations. The findings showed that counterfactual thinking correlated with both types of false belief tasks but did not correlate with contrary-to-fact syllogisms. These findings suggested that there might be a counterfactual element embedded in both types of false belief task. The results also highlighted that children found counterfactual tasks easier than the false belief to self task but of the same difficulty as the false

belief to others task. In this experiment, no difference in difficulty was found between the affirmative and the negative counterfactual questions, which contradicted the previous two experiments. This difference could be due to the tasks themselves. This point will be discussed further in the next section. Overall, the contrary-to-fact syllogisms did not correlate with any of the false belief tasks, only contrary-to-fact syllogisms that required a no-answer correlated with the false belief to self task. One possible explanation for this correlation could be the high level of inhibition that both these tasks required. The fact that counterfactual thinking correlated with both types of theory of mind tasks, but contrary-to-fact syllogisms did not, supports Riggs et al.'s (1998) claim that this correlation is due to the counterfactual element itself rather than general reasoning abilities. Another interesting finding was that no link was found between contrary-to-fact and counterfactual reasoning. This lack of relationship and two different patterns of correlations suggests that these two types of reasoning differ from each other in terms of their processes. Both tasks involved suppressing knowledge of current reality but the contrary-to-fact task required children to suppress their knowledge of the relations that hold in the real world whereas the counterfactual tasks required children to undo only one particular fact or action.

The last experiment further investigated the difference between counterfactual and contrary-to-fact reasoning. The children were asked to answer four counterfactual scenarios (two affirmative and two negative ones) and four contrary-to-fact scenarios (two yes-syllogisms and two no-syllogisms). These tasks were presented either in a standard context or in a contrary-to-fact context. According to Markovits (1995) a fantasy context acts as a cognitive filter preventing access to long term memory. Therefore it was expected, as reported in the literature (e.g.: Dias and Harris, 1989), that it would increase performance on contrary-to-fact reasoning, but it was hypothesised that it would increase the

difficulty of counterfactual tasks as children need to access their knowledge about the real world to solve them. In order to solve a counterfactual scenario, children need to refer to the causal relationships that happen in the world to imagine a new outcome after the antecedent has been potentially modified. The results confirmed this prediction. They showed an interaction between task and context. The context did not have the same effect on both types of tasks. The fantasy context facilitated contrary-to-fact reasoning but increased the difficulty of counterfactual reasoning. These results were consistent with Markovits's (1995) claim. In order to solve a contrary-to-fact syllogism children need to ignore their knowledge about the real world and the cognitive filter helps them to do that. On the contrary, in order to answer a counterfactual question children need to ignore only one fact about a particular situation but they need to access their knowledge of the real world. When looking at the performances on each type of task no difference in difficulty was found between affirmative and negative counterfactual questions or between yes and no syllogisms.

Some of the similarities and differences between these experiments and their results will now be discussed.

## **4 Main issues**

### **4.1 Counterfactual thinking**

All of the experiments in this thesis included some counterfactual tasks. The first findings suggested that counterfactual thinking was not different from future hypothetical thinking when controlling for negative and affirmative questions. This difference between affirmative and negative was investigated further. The results gained from comparing affirmative and negative counterfactual questions were not consistent across all the experiments. In experiments 1 and 2 it was found that negative counterfactual questions were more difficult than affirmative

counterfactual questions but this difference was not found in experiments 3 and 4. These differences could be explained by looking at the tasks used. There are good theoretical reasons (see chapter 2) suggesting that introducing a negation in a conditional question increases the difficulty of this question (Evans, 1998). This was confirmed in experiments 1 and 2 but not in experiments 3 and 4. Experiments 1 and 2 used the same types of task whereas experiments 3 and 4 included a different set of tasks. The tasks in the first two experiments were adapted from Riggs et al.'s (1998) experiment and we will refer to them as the 'boxes tasks'. In these tasks the children had to physically classify objects in a particular box according to one of their physical properties. Then this particular physical property, and therefore the classification, was changed and the children were asked a counterfactual question. These tasks did not involve following a narrative scenario and they were highly controlled as all the tasks were very similar to each other, only the material varied between tasks.

In contrast, experiments 3 and 4 used some more traditional tasks similar to the ones designed by Harris et al. (1996). These tasks involved listening to a short narrative scenario and then answering a counterfactual question. The four tasks used each time involved different types of answer. One of the negative and one of the affirmative counterfactual questions required the children to point to a particular location to answer the questions whereas the other affirmative counterfactual question required a no answer and the other negative counterfactual questions a yes answer. It could be argued that in that particular configuration the affirmative questions were made more difficult by requiring a no answer and the negative questions easier by requiring a yes answer. The variations of the type of answer expected could have influenced the results. In the first two experiments the tasks were very similar and the type of expected answer was the same for each question; the children were just required to point to the



right location. Experiments 3 and 4 required different types of answers therefore they might not be the best way of testing the difference between affirmative and negative counterfactual questions. In order to investigate the difference between affirmative and negative counterfactual questions it is necessary to control the task even more. An experiment could be conducted with traditional tasks with, for example, all the tasks requiring the children to point to a location to answer the question. Because of the theoretical background (e.g. Evans, 1998) and the more controlled design in experiments 1 and 2 it does seem that introducing a negation would increase the difficulty of a counterfactual questions.

#### **4.2 Counterfactual thinking and the false belief task**

An association between false belief tasks and counterfactual tasks was not systematically found in all the experiments. No significant correlation between counterfactual thinking and the false belief task was found in experiment 2 but there was a trend in the right direction with a correlation coefficient equal to .172. A significant correlation was found in experiment 3. In experiment 2 the 'boxes tasks' were used for counterfactual questions and the false belief task. In experiment 3 traditional counterfactual and traditional false belief tasks were used. It is difficult to know if the lack of significant correlation found in experiment 2 was an isolated event or if an explanation can be given. As mentioned above the tasks used in experiment 2 and in experiment 3 were very different. One of the main differences between these tasks was the importance of narration. Very little narration was involved in experiment 2 but more narration was used in experiment 3. The correlation between counterfactual thinking and the false belief tasks found could be due to language ability. Most of the research on the link between counterfactual thinking and the false belief tasks has been conducted with traditional tasks (e.g.: Grant et al., 2004) but when verbal abilities were partialled out the correlations remained. The difference between experiment 2 and 3 cannot

be explain in terms of narration and language abilities. There is a strong probably that a significant correlation would be found in experiment 2 if the power on this study was increased.

A correlation was also found, in experiment 3, between counterfactual thinking and false belief to self task (e.g. Smartie task). It therefore seems that counterfactual thinking is linked to false belief tasks in general. Both types of task involved undoing an event and holding two conflicting representations of a situation.

### **4.3 Contrary-to-fact syllogisms**

Traditionally children find contrary-to-fact syllogisms difficult (see chapter 4), but their performance tends to be improved if the syllogisms are presented in a fantasy context. The results of experiment 4 confirmed those claims and supported Markovits' (1995) theory that fantasy context acts as a cognitive filter reducing access to long term memory. Experiment 3 suggested that there was a difference between the syllogisms requiring a yes answer and the ones requiring a no answer as children found the first type easier, but this difference was not found in experiment 4. Experiment 3 included eight contrary-to-fact syllogisms (four yes and four no) whereas experiment 4 only included four contrary-to-fact syllogisms (two yes and two no). This difference in number of data points collected for each type of syllogism could explain why no difference was found between the yes and the no syllogisms in experiment 4. Leever and Harris (2000) found that children with autism gave more logical justifications for the yes syllogisms than for the no syllogisms. Their results supported the previous findings suggesting that no syllogisms are more difficult than the yes. According to Leever and Harris (2000) only the children with autism found the no syllogisms more difficult but the results presented in this thesis suggested that typically developing children also find the

no syllogisms more difficult. A hypothesis regarding the greater difficulty of no-syllogisms could be that they require a higher level of inhibition. An example of a no syllogism could be as follows: All snow is black; Len is a snowman made of snow; is Len white? In this case, the final question restates reality. As children need to inhibit their knowledge of the real world to solve this type of syllogism, inhibition needs to be increased with a no-syllogism as a real world question is asked at the end. This special characteristic of the no-syllogisms could explain their correlations with the false belief to self task. In the false belief to self tasks a strong inhibition is required as, for example in the Smartie task, smartie containers contain smarties. Children need to inhibit this real world knowledge.

#### **4.4 Counterfactual thinking and contrary-to-fact syllogisms**

Experiments 3 and 4 have shown a dissociation between these two types of reasoning tasks. In experiment 3 they correlated with different tasks and did not correlate with each other, and in experiment 4 it was shown that fantasy context did not influence performance in the same way. Both types of task require the children to suspend their knowledge of the real world and imagine a new situation.

In the 'floor' scenario the counterfactual question was 'What if Carol had taken her shoes off? Would the floor be dirty?' Children need to suppress their knowledge that Carol had her shoes on and need to imagine what would have happened if she had taken her shoes off. They then need to hold these two representations in their mind and use then both to solve the task as they need to compare the two.

In a contrary-to-fact syllogism such as 'All cows quack, Suzie is a cow, does Suzie quack? They need to suppress their knowledge that cows moo and now accept as a basis for reasoning that cows quack. They do not need both these

representations to solve the task, they just need to refer to the new state of the world.

This could explain why counterfactual thinking correlated with the false belief task and contrary-to-fact syllogisms did not. In false belief tasks, such as the Sally and Anne task, the children need to build two representations of the real situation where the marble changed location, and the false belief that Sally has got about the situation. Again, as in the counterfactual task, they need both representations, they need to compare them to solve the task.

#### **4.5 Control questions: do they have their place?**

Traditionally the false belief tasks and the counterfactual tasks have included some control questions to check that the children understand the task. They check in particular if the children can remember the situation at the beginning of the task and that they are aware of the final situation. The results of experiments 2 and 4 were analysed in two different ways: with the whole sample or with the sample only including the data of the children who passed all the control questions. The results from both these analyses did not significantly differ. The results were the same; the second analysis was just more powerful. These results raise the question of the usefulness of the control questions, at least for typically developing children. When the children are asked control questions then they have to answer 2 to 3 times as many questions overall. As the answers to these questions do not seem to influence the results, it could be argued that control questions could confuse the children and decrease their level of attention. The results might therefore be more reliable if no control questions were asked when the sample is made up of typically developing children.

Experiment 3 also showed that the link between counterfactual thinking and the false belief task was mainly due to the counterfactual element as the false

belief tasks correlated with the counterfactual tasks and not the other reasoning tasks (the contrary-to-fact syllogisms).

#### **4.6 The role of imagination**

The role of imagination maybe highlighted in relation to this work; it seems to be a common factor in most of the tasks that have been used. Imagination emerges in the second year of life when children start engaging in pretend play (Harris, 2000), and the absence of imagination, even this early in life, is seen as pathological as it is one of characteristics of autism (Baron-Cohen, 1994).

The contrary-to-fact syllogisms require children to suppress their knowledge of the real world and imagine a pretend world. They need to imagine what the world would be like if the first contrary-to-fact premise was true. In a counterfactual scenario children need to use their imagination to consider an alternative sequence of events if the antecedent had been different (Harris, 2000): they need to imagine a new consequence. It can also be argued that imagination is important to pass the false belief task as the children need to imagine what the false belief of the character is.

Harris (2000) studied the importance of children's imagination and claimed that it is usually well adapted to social engagement and interaction, mainly through shared pretend play, and also from a cognitive point of view, it helps children analyse reality. He even claimed that 'children's analysis of reality is infused by their ability to imagine what might have happened' (Harris, 2000, p. 138). Imagination seems to be key to children's social and cognitive development.

### **5 Theoretical background to the findings**

This section will link the findings summarised above to a theoretical perspective. We will argue that the findings of this thesis support the representational theory of mind view as described by Perner (1991). Perner (1991)

argued that children need to develop and manipulate their representations of the world to understand theory of mind. They need to be able understand that mental representations are representations, or models, of the world and they also need to be able to compare 'true' or 'false' representations. The representational view can also be applied to other aspects of children cognition than just theory of mind.

Starting with the end of the present work, experiments 3 and 4 of this thesis showed a clear distinction between counterfactual reasoning and contrary-to-fact reasoning. First of all the impact of fantasy context in experiment 4 was not identical on both types of tasks. Fantasy context acts as a cognitive filter preventing access to long term memory (Markovits, 1995) and it therefore facilitates contrary-to-fact syllogistic reasoning, as children need to suppress their knowledge of the world, but it increases the difficulty of counterfactual reasoning as children need to access their knowledge of the world in this case (see chapter 5 for more details). The different impact of fantasy context on both these types of tasks indicates that these are not underpinned by the same processes. This difference could be explained in representational terms. Contrary-to-fact syllogisms do not require children to represent different situations – they just require them to suppress their knowledge of the world whereas counterfactual scenarios require children to represent an alternative world and compare it to the actual situation. With the contrary-to-fact scenarios they only need to ignore, for example, that cows quack in the real world, but they do not need to represent a new model of the world. In the 'dirty shoes' counterfactual scenario, however, they need to represent two different models of the world, one where Carol walked into the room with dirty shoes on and another one where Carol had taken her shoes off.

The difference between contrary-to-fact syllogisms and counterfactual reasoning was also highlighted in chapter 4 where it was found that they did not

correlate with the false belief tasks in the same way: counterfactual tasks correlated with both false belief to others tasks and false belief to self tasks whereas contrary-to-fact syllogisms did not correlate with any of them. The distinction regarding representational difference between these two types of task can explain this difference in correlation with theory of mind tasks found in experiment 3. With the counterfactual task, children need to compare two representations of the world: the one that is consistent with current reality and the one of what the world could be if the antecedent had not happened. False belief tasks shares similar characteristics: with the traditional false belief to others tasks children need to compare their own representation of the world with the other character's representation of the situation and with the false belief to self tasks the children need to compare their current representation of the situation with their previous representation of the same situation. The representational approach is successful in explaining the correlation between the false belief tasks and counterfactual thinking as well as the lack of correlation with contrary-to-fact syllogisms as contrary-to-fact syllogisms do not require children to build model of the world. However, the representational approach does not seem to be able to explain the correlation found between contrary-to-fact (no) syllogisms with the false belief to self task. This link needs to be explored further.

The representational mind account can also explain the possible dissociation between negative and affirmative counterfactual questions found in experiment 1 and 2. It could be argued that negative counterfactual questions require a higher level of representation than affirmative questions. Going back to experiment 1 an affirmative counterfactual could be 'What if I had put a star sticker on the mouse's brick, which box would it be in?' The children need to represent, in their mind, a brick with a star sticker, instead of a round sticker, and reason counterfactually to answer the question. The situation to represent is clearly stated to the child. The

affirmative counterfactual question in that similar situation could be 'What if I hadn't put sticker on the mouse's brick, which box would it be in?'. In that situation the representation to construct is not as obvious as in the first case. The child need to represent a model where the sticker has not a star on and it opens the door to a number of other possible models. This type of model is therefore more difficult to construct. This might be why negative counterfactual questions tend to correlate more with the false belief task. Most of the literature exploring the link between counterfactual thinking and theory of mind used negative counterfactual questions only finding therefore a link between the counterfactual scenarios requiring more skills to build a model.

The representation view seems to explain most of the findings of this thesis; however some findings are not explained by it. For example the correlation between the contrary-to-fact syllogisms (no answer) with the false belief to self task could be explained by the fact that both tasks require a high level of inhibition. The no syllogisms (e.g.: All cats bark. George is a cat. Does George meow?) state the real world in the last premise therefore soliciting higher inhibition capacity from children. In the false belief to self children are expect to ignore their own previous false belief which could be argued requires a higher level of inhibition. This link between both these tasks with inhibition could be investigated further.

## **6 Future work**

A number of issues have been discussed in the previous section and some of these could be studied further in future research projects.

First of all the difference between yes and no syllogisms could be explored further. It was found in experiment 3 that children found no syllogisms more difficult than the yes syllogisms. It was hypothesised that this was because the question associated with a no syllogism was consistent with the real world,



therefore the children needed to have an even stronger inhibition (e.g.: All cats bark, George is a cat, does George meow?). The no-syllogisms could be rephrased in such a way that the question is not consistent with reality. For example, it could be as follows: All cats bark, George is a cat, Does George moo? The experimental design could include four no-syllogisms and four yes-syllogisms, all presented in a similar context, to typically developing children between the age of 3 to 5. Their performance on both types of task could then be compared. A link was found in experiment 3 between no-syllogisms and the false belief to self task. This new experiment could then be combined with four false belief to self tasks and an inhibition measure, such as a stroop task, to test the hypothesis that this link was due to higher demands on inhibition. This experiment would test whether inhibition is a factor that needs to be taken into account.

Then, the dissociation between contrary-to-fact syllogisms and counterfactual thinking could be explored further. Experiments 3 and 4 suggested that these two types of reasoning abilities differ significantly. Experiment 4 showed that fantasy context did not influence the performance on these tasks in the same way. A fantasy context tends to improve performance on contrary-to-fact syllogisms but increases the difficulty of counterfactual scenarios. However, it could be noted that the counterfactual scenarios in this experiment were true-to-fact. It is therefore difficult to know whether this difference is due to the reasoning element or the contrary-to-fact element.

First of all it would be interesting to investigate the difference between true-to-fact syllogisms and counterfactual reasoning. The children could be given four examples of each and all of them would be presented in a neutral context. A true-to-fact syllogisms could be as follows: All cats like milk, George is a cat, Does George like milk? The same counterfactual scenarios as in experiment 4 could be used. Performance on these tasks would then been compared.

It would be interesting to run a similar experiment with contrary-to-fact syllogisms and contrary-to-fact counterfactual scenarios. One example of a scenario could be as follows:

*You are in a magical land where fairies live. You meet a fairy who is called Lucy. Lucy tells you that in her world houses are cleaned with muddy shoes. One day, in her house the floor was dirty. But guess what! Lucy comes home that she doesn't take her shoes off, she makes the floor all clean with her shoes!*

*Before control question: Was the floor clean before?*

*Now control question: Is the floor clean now?*

*Test question: What if Carol had taken her shoes off? Would the floor be dirty?*

The children could be presented with four contrary-to-fact syllogisms and four contrary-to-fact counterfactual scenarios. There could be two conditions: a standard context and a fantasy context. It would still be predicted that fantasy context would improve performance on contrary-to-fact syllogisms. From experiment 4, it appears that fantasy context decreases performance on counterfactuals but, in this particular case, it might help children to accept the contrary-to-fact situation as a basis for reasoning, as claimed by Harris (2000). The finding could be that two opposite effects cancel each other and that, overall, fantasy context has no effect on contrary-to-fact counterfactual scenarios.

The work conducted in this thesis on theory of mind could be replicated with children with autism. Children with autism tend to show a specific deficit in theory of mind (Baron-Cohen, 1995). Baron-Cohen developed the theory that children with autism lacked a theory of mind and specifically developed the Sally and Ann task to test this hypothesis (Baron-Cohen et al., 1985). The link between the false belief task, counterfactual thinking and contrary-to-fact syllogisms could be explored with this special population. It would be expected that children with autism would find the false belief task difficult, as previously reported in the literature (e.g. Baron-Cohen, 1981). Children with autism should also find counterfactual reasoning more difficult than typically developing children as was

demonstrated by Grant et al. (2004) and Peterson and Bowler (2000). Grant et al. (2004) also found a correlation between counterfactual thinking and the false belief task when testing children with autism. This correlation could be due to the impairment of imagination that children with autism present (Baron-Cohen, 1987). Harris (2000) claimed that children use imagination for three main purposes: firstly to enter a make-believe world where many causal principles are retained, secondly to make comparison between actual outcomes and other possible outcomes and finally to explore impossible and magical events. This claim strongly suggested that imagination is necessary to solve counterfactual tasks and false belief tasks. In order to solve a counterfactual task, children need to compare the actual outcome with other possible outcomes. To solve a false belief task, children need to imagine a different world as well. The imagination of children with autism can be explored further by presenting them with contrary-to-fact syllogisms as they then need to enter a make believe world and test impossible events. It has been previously found that children with autism find contrary-to-fact syllogisms difficult (e.g. Leavers and Harris, 2000; Scott et al., 1999). It can also be predicted that a fantasy context would not help children to solve these syllogisms as they require children to imagine a pretend world. It would be interesting to run experiments 3 and 4 with children with autism to test these hypotheses. Children would have to be matched for intelligence and language in order to make the results comparable.

The influence of early interactions on the development of theory of mind can be studied with deaf children. Several studies have shown that deaf children from hearing families presented a delayed acquisition of theory of mind (e.g.: Deleau, 1996; Peterson and Siegal, 1995). Peterson and Siegal (1998) even showed that these children perform to the same level as children with autism on the false belief task. However, when native signers (i.e. deaf children from deaf families) were presented with the false belief task the delay previously found disappeared: these

children performed at least as well as hearing typically developing children (Courtin and Melot, 1998; Woolfe, Want and Siegal, 2002). The interaction between false belief tasks, counterfactual thinking and contrary-to-fact syllogisms could be studied with deaf children. The tasks would be presented in sign language to prevent language difficulties influencing the results.

## **6.1 Conclusion**

This work of this thesis has mainly focused on the cognitive skills necessary to pass the false belief task and counterfactual tasks. The experimental work has shown a relationship between counterfactual thinking and both types of false belief task: false belief to others and false belief to self. This work has also highlighted the difference between counterfactual thinking and contrary-to-fact syllogistic reasoning.

A large part of the research looked at the influence of negation, in the phrasing of a question but also in the answer expected. The results of experiments 1 and 2 showed that introducing a negation in a counterfactual question increased the difficulty of this question. Experiment 3 showed that syllogisms requiring a no-answer were more difficult than syllogisms requiring a yes-answer.

It also seems that imagination might be a strong common factor in all these tasks, including contrary-to-fact syllogisms. Imagination seems to be important for children's reasoning as it allows them to imagine alternative possibilities (Harris, 2000). Imagination is also a basis for scientific reasoning as children are required to imagine possible outcomes to experiments.

This thesis highlighted that the ability to construct models of the world in order to solve counterfactual tasks and false belief tasks is important, therefore supporting Perner's (1991) theoretical account of theory of mind. However,

another factor, inhibition, needs to be studied too as some results suggested that it might be an important factor too.

Another aspect that must not be forgotten is the way children use their theory of mind in the real world. Being able to reason counterfactually is not enough to pass the false belief task, children still need to take into account people's mental states, and passing a false belief task is not as complex as dealing with people's mental states in a naturally occurring social interaction.

# APPENDICES

## List of tables for all the appendices

Table 1: Number of errors by type.....	203
Table 2: Mean and standard deviation for the total of correct answer for counterfactual questions and false belief questions (n=25).....	203
Table 3 Correlations between counterfactual thinking and false belief task questions.....	204
Table 4: Correlations between the number of realist errors and the false belief questions for the reduced sample (n=25). ....	205
Table 5: Correlations between number of errors and tasks. (n=25).....	205
Table 6: Mean and standard deviation of the scores for each type of tasks in the two different contexts. ....	214

## Figure for all the appendices

Figure 1. Percentage of right answers for each false belief question.....	202
---	-----

## APPENDIX A

### RECORDING SHEET FOR EXPERIMENT 1

Participant			Scenario	Questions + Expected answer	Answers	Right or Wrong
Number	Gender	Age				
1			Cars + CFaff	1. What is on this car now? (circle)		
				2. Where shall we put this car now? 1		
				3. If I had put a square sticker on this car, which box would in be in? 2		
			Colours + CFneg	1. What is on this card now? (Orange)		
				2. Where shall we put this card now? 1		
				3. If I hadn't drawn on this card, which box would it be in? 3		
			Dough + FHaff	If I roll the dough into a ball, which box will it go into? 1		
			Drawing + FHneg	If I don't draw anything on this piece of paper, which box will it go into? 3		
2			Dough + CFneg	1. What shape is the dough now? (ball)		
				2. Where shall we put the dough now? 1		
				3. If I hadn't rolled the dough at all, which box would it be in? 3		
			Drawing + CFaff	1. What is on this p. of p. now? (circle)		
				2. Where shall we put this p.of p. now? 1		
				3. If I had drawn a square on this p.of p., which box would it be into now? 2		
			Cars + FHneg	If I don't put a sticker on this card, which box will it go into? 3		

			Colours + FHaff	If I draw an orange patch on this card, which box will it go into? 1		
3			Drawing + FHaff	If I draw a circle on this p. of p., which box will it go into? 1		
			Dough + FHneg	If I don't roll the dough at all, which box will it go into? 3		
			Colours + CFaff	1. What is on this card now? ( <b>orange</b> )		
				2. Where shall we put this card now? 1		
				3. If I had drawn a green patch on this card, which box would it be in? 2		
			Cars + CFneg	1. What is on this car now? ( <b>Circle</b> )		
				2. Where shall we put this car now? 1		
				3. If I hadn't put a sticker on this car, which box would it be in? 3		
4			Colours + FHneg	If I don't draw on this card, which box will it go into? 3		
			Cars + FHaff	If I put a round sticker on this car, which box will it go into? 1		
			Drawing + CFneg	1. What is on this p. of p. now? ( <b>Circle</b> )		
				2. Where shall we put this p. of p. now? 1		
				3. If I hadn't drawn anything on this p. of p., which box would it be into? 3		
			Dough + CFaff	1. What shape is the dough now? ( <b>ball</b> )		
				2. Where shall we put this dough now? 1		
				3. If I had rolled the dough into a sausage, which box would it be in? 2		



## **APPENDIX B**

### **OTHER COUNTERFACTUAL AND FUTURE HYPOTHETICAL SCENARIOS USED IN EXPERIMENT 1**

#### **Play dough and shapes**

The play dough could have the shape of a ball, a sausage or be squashed.

Counterfactual (affirmative): What if I had rolled the dough as a sausage, which box would it be in?

Counterfactual (negative): If I hadn't rolled the dough at all, which box would it be in?

Future hypothetical (affirmative): If I roll the dough into a ball, which box will it go into?

Future hypothetical (negative): If I don't roll the dough at all, which box will it go into?

#### **Sheets of paper and shapes**

On each sheet of paper there could be a square, a circle or nothing drawn on.

Counterfactual (affirmative): If I had drawn a square on this piece of paper, which box would it be into now?

Counterfactual (negative): If I hadn't drawn anything on this piece of paper, which box would it be into?

Future hypothetical (affirmative): If I draw a circle on this piece of paper, which box will it go into?

Future hypothetical (negative): If I don't draw anything on this piece of paper, which box will it go into?

## **Cards and colours**

On each card there could be a green patch, an orange patch or nothing.

Counterfactual (affirmative): If I had drawn a green patch on this card, which box would it be in?

Counterfactual (negative): If I hadn't drawn on this card, which box would it be in?

Future hypothetical (affirmative): If I draw an orange patch on this card, which box will it go into?

Future hypothetical (negative): If I don't draw on this card, which box will it go into?

## APPENDIX C

### RECORDING SHEET FOR EXPERIMENT 2

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Shapes + CF Aff	1. What's on this p. of paper now? (circle)		
				2. Where shall we put this p. of paper now? (1)		
				3. If I had drawn a square on this p of p which box would it be in now?		
			Colours + CF Neg	1. What's the colour on this p of p now? (red)		
				2. Where shall we put it now? (1)		
				3. If I hadn't drawn on this p of p which box would it be in now? (3)		
			Playdough + ToM (t/l)	1. What shape is the playdough now? (ball)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his playdough is?		
				4. Where will the mouse look for his playdough first? (3)		
				5. Where was the mouse's playdough at the beginning? (3)		
			Bricks + ToM (l/t)	1. What's on this brick now? (circle)		
				2. Where shall we put it now? (1)		
				3. Where will the Mouse look for his brick first?		
				4. Where does the Mouse think his brick is? (3)		
				5. Where was the brick at the beginning? (3)		

Combination 1/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Faces + CF Aff	1. What sort of glasses does the man wear now? (round)		
				2. Where shall we put this man now? (1)		
				3. If I had drawn a square glasses for this man which box would he be in now? (2)		
			Bob + CF Neg	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put Bob now? (1)		
				3. If I hadn't drawn a hat for Bob which box would he be in now? (3)		
			Teletubbies + ToM (I/I)	1. What colour is the ball now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his Teletubby is?		
				4. Where will the mouse look for his Teletubby first? (3)		
				5. Where was the mouse's Teletubby at the beginning? (3)		
			Teddies + ToM (I/t)	1. What colour is the box tie now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where will the Mouse look for his Teddy first? (3)		
				4. Where does the Mouse think his Teddy is? (3)		
				5. Where was the Teddy at the beginning? (3)		

Combination 1/b

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Colours + CF Aff	1. What's on this p. of paper now? (red)		
				2. Where shall we put this p. of paper now? (1)		
				3. If I had drawn a blue patch on this p of p which box would it be in now? (2)		
			Shapes + CF Neg	1. What's on this p of p now? (circle)		
				2. Where shall we put it now? (1)		
				3. If I hadn't drawn on this p of p which box would it be in now? (3)		
			Bricks + ToM1 (l/l)	1. What's on this brick now? (circle)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his brick is?		
				4. Where will the mouse look for his brick first?		
				5. Where was the mouse's brick at the beginning?		
			Playdough + ToM2 (l/t)	1. What shape is the playdough now? (ball)		
				2. Where shall we put it now? (1)		
				3. Where will the Mouse look for his playdough first? (3)		
				4. Where does the Mouse think his playdough is?		
				5. Where was the playdough at the beginning?		

Combination 2/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Bob + CF Aff	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put Bob now? (1)		
				3. If I had drawn a pointy hat on Bob's head which box would he be in now? (2)		
			Faces + CF Neg	1. What sort of glasses does this man wear now? (round)		
				2. Where shall we put this man now? (1)		
				3. If I hadn't drawn glasses for this man which box would he be in now? (3)		
			Teddies + ToM1 (t/l)	1. What colour is the tie bow now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his Teddy is? (3)		
				4. Where will the mouse look for his Teddy first? (3)		
				5. Where was the mouse's Teletubby at the beginning? (3)		
			Tubbies + ToM2 (l/t)	1. What colour is the ball now? (blue)		
				2. Where shall we put the teletubbie now? (1)		
				3. Where will the Mouse look for his Teletubby first?		
				4. Where does the Mouse think his Teletubby is?		
				5. Where was the Teletubby at the beginning?		

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Playdough + CF Neg	1. What shape is the playdough now? (ball)		
				2. Where shall we put the playdough now? (1)		
				3. If I hadn't rolled the playdough which box would he be in now?(3)		
			Bricks + CF Aff	1. What shape is on this brick now? (circle)		
				2. Where shall we put this brick now? (1)		
				3. If I had put a star on this brick which box would he be in now? (2)		
			Shapes + ToM2 (I/t)	1. What's on this p of p now? (circle)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his p of p first? (3)		
				4. Where does the mouse think his p of p is? (3)		
				5. Where was the mouse's p of p at the beginning? (3)		
			Colour + ToM1 (t/l)	1. What colour is on this p of p now? (blue)		
				2. Where shall we put this p of p now? (1)		
				3. Where does the Mouse think his p of p is?		
				4. Where will the Mouse look for his p of p first?		
				5. Where was the p of p at the beginning?		

Combination 3/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Tubbies + CF Neg	1. What colour is the ball now? (red)		
				2. Where shall we put the ball now? (1)		
				3. If I hadn't coloured in the ball which box would he be in now?(3)		
			Teddy + CF Aff	1. What colour is the tie bow now? (red)		
				2. Where shall we put this teddy now? (1)		
				3. If I had coloured the tie bow in blue which box would he be in now? (2)		
			Faces + ToM2 (l/t)	1. What sort of glasses does this man wear now? (round)		
				2. Where shall we put this man now? (1)		
				3. Where will the mouse look for his p of p first? (3)		
				4. Where does the mouse think his p of p is? (3)		
				5. Where was the mouse's p of p at the beginning? (3)		
			Bob + ToM1 (t/l)	1. What sort of hat does Bob wear now? (blue)		
				2. Where shall we put Bob now? (1)		
				3. Where does the Mouse think his Bob is?		
				4. Where will the Mouse look for his Bob first?		
				5. Where was the Mouse's Bob at the beginning?		

Combination 3/b



Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Bricks + CF Neg	1. What shape is this brick now? (circle)		
				2. Where shall we put this brick now? (1)		
				3. If I hadn't put a sticker on this brick which box would he be in now?(3)		
			Playdough + CF Aff	1. What shape is the playdough now? (circle)		
				2. Where shall we put this playdough now? (1)		
				3. If I had rolled the playdough into a sausage which box would he be in now? (3)		
			Colour + ToM2 (l/t)	1. What colour is on this p of p now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his p of p first? (3)		
				4. Where does the mouse think his p of p is? (3)		
				5. Where was the mouse's p of p at the beginning? (3)		
			Shapes + ToM1 (t/l)	1. What's on this p of p now? (circle)		
				2. Where shall we put this p of p now? (1)		
				3. Where does the Mouse think his p of p is?		
				4. Where will the Mouse look for his p of p first?		
				5. Where was the p of p at the beginning?		

Combination 4/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Teddy + CF Neg	1. What colour is the Teddy's tie bow now? (ball)		
				2. Where shall we put the Teddy now? (1)		
				3. If I hadn't coloured in the tie bow which box would it be in now?(3)		
			Tubbies + CF Aff	1. What colour is the ball now? (blue)		
				2. Where shall we put it now? (1)		
				3. If I had coloured the ball in red which box would it be in now? (2)		
			Bob + ToM2 (l/t)	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put Bob now? (1)		
				3. Where will the mouse look for his Bob first? (3)		
				4. Where does the mouse think his Bob is? (3)		
				5. Where was the mouse's Bob at the beginning? (3)		
			Faces + ToM1 (v/l)	1. What sort of glasses does this man wear now? (round)		
				2. Where shall we put this man now? (1)		
				3. Where does the Mouse think his p of p is?		
				4. Where will the Mouse look for his p of p first?		
				5. Where was the p of p at the beginning?		

Combination 4/b

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Faces + ToM1 (v/l)	1. What sort of glasses does the man wear now? (round)		
				2. Where shall we put the man now? (1)		
				3. Where does the mouse think his p of p is?		
				4. Where will the mouse look for his p of p first?		
				5. Where was the mouse's p of p at the beginning?		
			Bob + ToM2 ((l/t)	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put Bob now? (1)		
				3. Where will the mouse look for his p of p first?		
				4. Where does the mouse think his p of p is?		
				5. Where was the mouse's p of p at the beginning?		
			Tubbies + CF aff	1. What colour is the ball now? (blue)		
				2. Where shall we put it now? (1)		
			Teddy + CF neg	1. What colour is the bow tie now? (blue)		
				2. Where shall we put it now?		
				3. If I had not coloured in the bow tie which box would it be in? (3)		

Combination 5/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Shapes + ToM1 (v/l)	1. What's on this p of p now? (round)		
				2. Where shall we put this p of p now? (1)		
				3. Where does the mouse think his p of p is?		
				4. Where will the mouse look for his p of p first?		
				5. Where was the mouse's p of p at the beginning?		
			Colours + ToM2 ((l/t)	1. What colour is on this p of p now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his p of p first?		
				4. Where does the mouse think his p of p is?		
				5. Where was the mouse's p of p at the beginning?		
			Playdough + CF aff	1. What shape is the playdough now? (ball)		
				2. Where shall we put it now? (1)		
				3. If I had rolled the playdough into a sausage which box would it be in now? (2)		
			Bricks + CF neg	1. What's on this brick now? (blue)		
				2. Where shall we put it now?		
				3. If I had not put a sticker on this brick now which box would it be in? (3)		

Combination5/b

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Bob + ToM1 (t/l)	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his p of p is?		
				4. Where will the mouse look for his p of p first?		
				5. Where was the mouse's p of p at the beginning?		
			Faces + ToM2 (l/t)	1. What sort of glasses does this man wear now? (1)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his p of p first?		
				4. Where does the mouse think his p of p is?		
				5. Where was the mouse's p of p at the beginning?		
			Teddy + CF aff	1. What colour is the bow tie now? (blue)		
				2. Where shall we put it now? (1)		
				3. If I had coloured the bow tie in red which box would it be in now? (2)		
			Tubbies + CF neg	1. What colour is the ball now? (blue)		
				2. Where shall we put it now?		
				3. If I hadn't coloured in the ball which box would it be in now? (3)		

Combination 6/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Colours + ToM1 (u/l)	1. What colour is on this p of p now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his p of p is?		
				4. Where will the mouse look for his p of p first?		
				5. Where was the mouse's p of p at the beginning?		
			Shapes + ToM2 (l/t)	1. What's on this p of p now? (circle)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his p of p first?		
				4. Where does the mouse think his p of p is?		
				5. Where was the mouse's p of p at the beginning?		
			Bricks + CF aff	1. What's on this brick now? (circle)		
				2. Where shall we put it now? (1)		
				3. If I had put a star sticker on this brick, which box would it be in now? (2)		
			Playdough + CF neg	1. What shape is the playdough now?		
				2. Where shall we put it now?		
				3. If I hadn't rolled the playdough which box would it be in now?		

Combination 6/b

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Tubbies + ToM2 (l/t)	1. What colour is the ball now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his teletu first?		
				4. Where does the mouse think his teletubby is?		
				5. Where was the mouse's teletubby at the beginning?		
			Teddy + ToM1 (t/l)	1. What colour is the bow tie now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his teddy is?		
				4. Where will the mouse look for his teddy first?		
				5. Where was the mouse's p of p at the beginning?		
			Faces + CF neg	1. What sort of glasses does the man wear now? (round)		
				2. Where shall we put it now? (1)		
				3. If I had not drawn any glasses on this man which box would it be in now? (3)		
			Bob + CF aff	1. What sort of hat does Bob wear now? (blue)		
				2. Where shall we put it now?		
				3. If I had drawn a pointy hat on Bob's head which box would he be in now? (3)		

Combination 7/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Playdough + ToM2(l/t)	1. What shape is the playdough now? (ball)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look to his playdough first?		
				4. Where does the mouse think his playdough is?		
				5. Where was the mouse's playdough at the beginning?		
			Bricks + ToM1 (t/l)	1. What's on this brick now? (round)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his brick is?		
				4. Where will the mouse look for his brick first?		
				5. Where was the mouse's brick at the beginning?		
			Shapes + CF neg	1. What's on this p of p now? (circle)		
				2. Where shall we put it now? (1)		
				3. If I hadn't drawn anything on this p of p which box would it be in now?		
			Colour + CF Aff	1. What colour is on this p of p now? (blue)		
				2. Where shall we put it now?		
				3. If I had drawn a red patch on this p of p which box would it be in now? (2)		

Combination 7/b



Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Teddy + ToM2(l/t)	1. What colour is the bow tie now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his teddy first?		
				4. Where does the mouse think his Teddy is?		
				5. Where was the mouse's playdough at the beginning?		
			Tubbies + ToM1 (t/l)	1. What colour is the ball now? (blue)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his teletubby is?		
				4. Where will the mouse look for his teletubby first?		
				5. Where was the mouse's teletubby at the beginning?		
			Bob + CF neg	1. What sort of hat does Bob wear now? (round)		
				2. Where shall we put him now? (1)		
				3. If I hadn't drawn anything on Bob's head which box would he be in now?		
			Faces + CF Aff	1. What sort of glasses does this man wear now?		
				2. Where shall we put him now?		
				3. If I had drawn square glasses on this man which box would he be in now? (2)		

Combination 8/a

Participant			Scenario	Questions	Answers	Right / Wrong
Number	Gender	Age				
			Bricks + ToM2(l/t)	1. What's on this brick now? (circle)		
				2. Where shall we put it now? (1)		
				3. Where will the mouse look for his brick first?		
				4. Where does the mouse think his brick first?		
				5. Where was the mouse's brick at the beginning?		
			Playdough + ToM1 (t/l)	1. What shape is the playdough now? (round)		
				2. Where shall we put it now? (1)		
				3. Where does the mouse think his playdough is?		
				4. Where will the mouse look for his playdough first?		
				5. Where was the mouse's playdough at the beginning?		
			Colour + CF neg	1. What colour is on this p of p now? (red)		
				2. Where shall we put it now? (1)		
				3. If I hadn't drawn anything on this p of p which box would it be in now?		
			Shape + CF Aff	1. What's on this p of p now? (circle)		
				2. Where shall we put it now?		
				3. If I had drawn a star on this p of p which box would it be in now? (2)		

Combination 8/b

## **APPENDIX D**

### **DESCRIPTION OF THE TASKS USED IN EXPERIMENT 2**

Testing the link between counterfactual situation and false belief task:  
affirmative and negative counterfactual, and standard false belief task using three different boxes in order to be able to analyse the mistakes children might make.  
Do they stick to reality or not?

For each scenario, the children are explained how the objects are classified in each box. A reminder card is placed in front of each box to prevent memory from getting overloaded. The children are then asked to physically classify a few objects.

During the task the experiment is helped by a mouse puppet.

#### **SCENARIO 1: pieces of paper and shapes**

3 boxes containing pieces of paper which have either a circle, a square or no shape on.

The experimenter is helped by the mouse. The mouse adds a blank piece of paper to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's piece of paper and draws a circle on it.

Q1: 'What is on this piece of paper now?'

Q2: 'Where shall we put this piece of paper now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't drawn a shape on this piece of paper, which box would it be in?' 3

CF Affirma: Q4: 'If I had drawn a square on this piece of paper, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with the piece of paper.

ToM : Q5: 'Where will the mouse look for his piece of paper first?'

Q5bis: 'Where does the mouse think his the piece of paper is?'

→ Memory question

Q6: 'Where was the mouse's piece of paper at the beginning?'

## **SCENARIO 2: Cards and colours**

There are three boxes containing cards with either a red patch on, a blue patch on or no colour.

The mouse adds a blank card into box 3, and then he goes to sleep in the experimenter's bag.

The experiment removes the mouse's card from box 3 and draws a red patch on it.

Q1: 'What colour is on this piece of paper now?'

Q2: 'Where shall we put this piece of paper now?' 1

→ Counterfactual situation

CF negative: Q3: 'If I hadn't drawn this piece of paper which box would it be in?' 3

CF affirmat: Q4: 'If I had drawn a blue patch on the mouse's card, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with his card.

ToM: Q5: 'Where will the mouse look for his piece of paper first?'

Q5bis: 'Where does the mouse think that his piece of paper is?

→ Memory question

Q6: 'Where was the piece of paper at the beginning?'

### **SCENARIO 3: Play dough**

3 boxes containing play dough which either has a round shape, a sausage shape or a 'squashed' shape.

The experimenter is helped by the mouse. The mouse adds a 'squashed' piece of dough to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's dough and rolls it into a ball.

Q1: 'What shape is this piece of dough now'

Q2: 'Where shall I put this piece of dough?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't rolled the play dough, which box would it be in now?' 3

CF Affirma: Q4: 'If I had rolled the play dough into a sausage, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with his dough.

ToM : Q5: 'Where will the mouse look for his play dough first?'

Q5bis: 'Where does the mouse think that his play dough is?'

→ Memory question

Q6: 'Where was the mouse's play dough at the beginning?'

### **SCENARIO 4: Bricks and stickers**

3 boxes containing bricks which either have a round sticker, a star sticker or no sticker on.

The experimenter is helped by the mouse. The mouse adds his brick with no sticker on to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's brick from the box and puts a round sticker on it.

Q1: 'What is on this brick now?'

Q2: 'Where shall we put this brick now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't put a sticker on the mouse's brick, which box would it be in?' 3

CF Affirma: Q4: 'If I had put a star sticker on the mouse's brick, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with his brick.

ToM : Q5: 'Where will the mouse look for his brick first?'

Q6bis: 'Where does the mouse think that his brick is?'

→ Memory question

Q7: 'Where was the mouse's brick at the beginning?'

## **SCENARIO 5: Faces and Glasses**

3 boxes containing pictures of men wearing either some round glasses, square glasses or no glasses.

The experimenter is helped by a mouse. The mouse adds picture with no glasses on to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's picture from the box and draws some round on it.

Q1: 'What sort of glass does this man wear now?'

Q2: 'Where shall we put this man now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't drawn glasses for this man, which box would he be in?' 3

CF Affirma: Q4: 'If I had drawn a pair of square glasses for this man, which box would he be in?' 2

→ False belief condition

The mouse comes back and he wants to play with his picture.

ToM : Q6: 'Where will the mouse look for piece of paper first?'

Q6bis: 'Where does the mouse think that his piece of paper is?'

→ Memory question

Q7: 'Which box was the mouse's piece of paper at the beginning?'

### **SCENARIO 6: Teddies and Bow ties**

3 boxes containing pictures of teddies wearing either a blue bow tie, a red bow tie or a bow tie with no colour.

The experimenter is helped by the mouse. The mouse adds his own picture of a teddy with a bow tie with no colour to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's picture from the box and colour the bow tie in red.

Q1: 'What colour is the bow tie now?'

Q2: 'Where shall we put it now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't colour the bow tie, which box would it be in now?' 3

CF Affirma: Q4: 'If I had coloured the bow tie in blue, which box would it be in now?' 2

→ False belief condition

The mouse comes back and he wants to play with his brick.

ToM : Q6: 'Where will the mouse look for his Teddy first?'

Q6bis: 'Where does the mouse think that his Teddy is?'

→ Memory question

Q7: 'Where was the mouse's Teddy at the beginning?'

### **SCENARIO 7: Faces and Hats**

3 boxes containing drawings of faces (the character was called Bob) which either wear a pointy hat, a round hat or no hat.

The experimenter is helped by the mouse. The mouse adds his own doll with no hat in to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's drawing of Bob from the box and puts a pointy hat on it.

Q1: 'What sort of hat does Bob wear now?'

Q2: 'Where shall we put Bob now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't a hat on Bob, which box would it be in?' 3



CF Affirma: Q4: 'If I had put a round hat on Bob, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with Bob.

ToM : Q6: 'Where will the mouse look for Bob first?'

Q6bis: 'Where does the mouse think that Bob is?'

→ Memory question

Q7: 'Where was the mouse's Bob at the beginning?'

### **SCENARIO 8: Teletubby and balls**

3 boxes containing pictures of Teletubby which either carry a red ball, a blue ball or a white ball.

The experimenter is helped by the mouse. The mouse adds his own picture of a Teletubby with a white ball in to box 3 and then goes sleep in the experimenter's bag.

The experimenter removes the mouse's Teletubby from the box and colour his ball red.

Q1: 'What colour is the Teletubby's ball now?'

Q2: 'Where shall we put the Teletubby now?' 1

→ Counterfactual reasoning

CF Negative: Q3: 'If I hadn't coloured the Teletubby's ball, which box would it be in?' 3

CF Affirma: Q4: 'If I had coloured the Teletubby's ball in blue, which box would it be in?' 2

→ False belief condition

The mouse comes back and he wants to play with his Teletubby.

ToM : Q6: 'Where will the mouse look for his Teletubby first?'

Q6bis: 'Where does the mouse think that his Teletubby is?'

→ Memory question

Q7: 'Where was the mouse's Teletubby at the beginning?'

## APPENDIX E

### EXPERIMENT 2: ANALYSIS WITH THE REDUCED SAMPLE

*Theory of Mind (excluding data when children did not answer the control question correctly)*

For the following analysis the children who answered at least one control question incorrectly were excluded from the analysis. This sample will be called the reduced sample, it includes 25 children.

Questions	Mean	Standard Deviation
Think	2.52	1.41
Look	2.60	1.38

Table 1 Means and standard deviations for the think or look questions (n=25).

A t-test showed that the difference between the 'think' and the 'look' questions was not significant,  $t(24) = -.258, p=.799$ .

The percentages of right answers for think1, look2, look1 and think2 were calculated and are presented in the following chart.

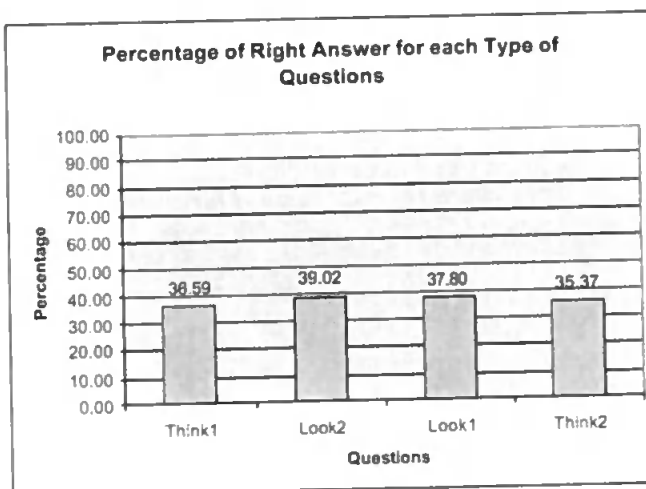


Figure 1. Percentage of right answers for each false belief question

An ANOVA showed that there was no significant difference between the false belief questions,  $F(3; 72) = .065, p=.978$ . Therefore there is no difference between 'think' and 'look', and there is no order effect.

*Errors on the false belief tasks*

As explained in the results section in chapter 3 children can make 'realist' and 'non-realist' errors. The total number of each type of errors was computed and is presented in the following table.

	Realist errors	Non realist errors	Total
Total number	64	28	92

Table 1: Number of errors by type

A one sample t-test showed that the number of children making realist errors ( $t(24) = 4.575, p<.001$ ) was highly significantly different from chance level.

*Comparison between counterfactual thinking and false belief task*

The performances on counterfactual tasks and false belief tasks were compared. The total number of correct answers on counterfactual tasks was computed and the total number of correct answers on false belief questions was calculated and divided by two to obtain a score out of four for each set of data. The means and the standard deviations were calculated for each type of questions. The results are presented in the following table.

Type of Question	Mean	Standard Deviation
Counterfactual questions	3.20	1.08
False Belief Questions	2.44	1.16

Table 2: Mean and standard deviation for the total of correct answer for counterfactual questions and false belief questions (n=25).

A t-test showed that the difference of performance between counterfactual questions and false belief questions was significant,  $t(24)=2.27$ ,  $p<.001$ .

*Correlations between counterfactual thinking and false belief tasks for the reduced sample.*

The link between counterfactual thinking and false belief task was investigated by running correlations. The results are presented in the following table.

	T.o.M. Total	C.F. Total	T.o.M. Think	T.o.M. Look	C.F. Aff	C.F. Neg
T.o.M Total	1					
C.F. Total	-.026	1				
T.o.M Think	.837(**)	.011	1			
T.o.M. Look	.828(**)	-.056	.386	1		
C.F. Aff	-.029	.733(**)	.087	-.138	1	
C.F. Neg	-.016	.864(**)	-.049	.024	.290	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

Table 3 Correlations between counterfactual thinking and false belief task questions.

The results show that there is no significant correlation between any of the theory of mind questions and the counterfactual questions.

The link between the number of realist errors and the false belief questions were investigated by running correlations. The results are presented in the following table.

	Think 1	Look 2	Look 1	Think 2	Think Total	Look Total	Realist Errors
Think 1	1						
Look 2	.425(*)	1					
Look 1	.249	.495(*)	1				
Think 2	.552(**)	.178	.254	1			
Think Total	.876(**)	.340	.286	.886(**)	1		
Look Total	.391	.867(**)	.862(**)	.250	.362	1	
Realist Errors	.274	-.034	-.006	.142	.235	-.023	1

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

Table 4: Correlations between the number of realist errors and the false belief questions for the reduced sample (n=25).

The results showed no significant correlations between the number of realist errors on counterfactual questions with any of the false belief task questions.

The number of realist errors on the theory of mind questions and the other tasks have been compared. The results are presented in the following table.

	No Realist Errors CF	ToM realist errors	TOM	CF Total
No Realist Errors CF	1	-.219	.079	-.791(**)
ToM realist errors	-.219	1	-.619(**)	.033
TOM	.079	-.619(**)	1	-.026
CF Total	-.791(**)	.033	-.026	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

Table 5: Correlations between number of errors and tasks. (n=25)

The results showed that the number of realist errors in counterfactual and theory of mind tasks did not correlate and that there was not correlation between the number of realist errors on the false belief tasks and the performance on counterfactual tasks.

## **APPENDIX F**

### **COUNTERFACTUAL SCENARIOS USE IN EXPERIMENT 3**

#### **Scenario 2: Yes / No**

One day the tree was standing up like this. But guess what! There was a storm and the wind came along and blew the tree over like this.

Now control question: Is the tree standing up now? (No)

Before control question: Was the tree standing up before? (Yes)

Test question: What if the wind had not come – would the tree be standing up?  
(Yes)

#### **Scenario 3: Selective pointing**

One day Ruddy the Rat has some chocolate, he holds the chocolate in his mouth like this. But guess what! He drops the chocolate and it falls on the floor.

Now control question: Where is the chocolate now? Can you show me?

Before control question: Where was the chocolate before? Can you show me?

Test question: What if Ruddy had not dropped the chocolate – where would the chocolate be? Can you show me?

#### **Scenario 4: Selective pointing**

Robert the bunny lives in this box. But guess what! One day the door is not shut properly. So the bunny rabbit opens the door and he goes to sit here, outside the box.

Now control question: Where is the bunny rabbit now? Can you show me?

Before control question: Where was the bunny rabbit before? Can you show me?

Test question: What if the door of the box had been shut properly – where would the bunny rabbit be? Can you show me?



## **APPENDIX G**

### **FALSE BELIEF TO OTHERS TASK USED IN EXPERIMENT 3**

#### **Pooh bear task**

A closed wooden box and an open plastic bowl of similar dimensions were placed on the table. The child was introduced to a 'Winnie-The-Pooh' puppet who was enjoying a snack from his honey-pot, which was then placed in the bowl prior to his departure. 'Piglet' then entered, took the honey-pot from the bowl, ate from it, and placed it in the closed box.

The child was then reintroduced to Pooh, and told that he was hungry.

FB question: 'Where do you think Pooh will look for the honey?' Wooden Box

Reality question 1: 'Where is the honey really?' Silver Box

Memory question 2: 'Where did Pooh put the honey?' Wooden Box

#### **Maxi**

'Mother returns from her shopping trip. She bought chocolate for her cake. Maxi may help her put away the things. He asks her: 'Where should I put the chocolate?' 'In the blue cupboard' says the mother.

Maxi put the chocolate into the blue cupboard. [A toy chocolate is put into the blue cupboard box.] Maxi remembers exactly where he put the chocolate so that he could come back and get some later. He loves chocolate. Then he leaves for the playground. [The boy doll is removed] Mother starts to prepare the cake and takes the chocolate out of the cupboard. She grates a bit into the dough and then she does not put it back into the blue but the green cupboard. [Toy chocolate is thereby transferred from the blue to the green cupboard] Now she realises that she forgot to buy eggs. So she goes to her neighbour for some eggs. Then Maxi

comes back from the playground, hungry, and he wants to get some chocolate.

[Boy doll appears]. He still remembers where he had put the chocolate.'

Belief question: 'Where will Maxi look for the chocolate?'

Reality question: 'Where is the chocolate really?'

Memory question: 'Where did Maxi put the chocolate?'

## **APPENDIX H**

### **FALSE BELIEF TO SELF TASKS USED IN EXPERIMENT 3**

#### **Cereal Box Task**

The material included: 2 boxes: cereals + opaque box. Before the experiment was run the content of the cereal box was moved to the opaque box. The children were told the following instructions and questions:

The children were shown both closed boxes and asked the false belief question: 'Where are the breakfast cereals?' (Right answer: in the cereal box)

The children were then asked to open both boxes, one after the other and check their content. They discovered that the cereals were in the unmarked box and that the cereal box was empty. Both boxes were closed again, and the children were then asked the test question: 'Where did you think the breakfast cereals were before you opened the boxes?'

Finally the children were asked the reality control question: 'Where are the breakfast cereals really?'

#### **Crayon Task**

Before the experiment crayons were removed from a crayon box and replaced by a rubber.

The children were shown the crayon box and asked the false belief question: 'What is inside the box?' (Right answer: crayons) The children were then asked to open the box and found a rubber instead. They were then asked the test question: 'What did you think was inside the box before you opened it?' Finally they were asked the reality control question: 'What is inside the box really?'

# APPENDIX I

## CONTRARY-TO-FACT SYLLOGISMS USED IN EXPERIMENT 3

### Instructions:

'Let's try this one together. Now, I'm going to tell you something that may sound funny, but I want you to think about it. I want you to imagine a picture in your head. It could be; All fish live in trees. So, the fish you are thinking about, in the picture in your head, do they live in trees? Yes? Good. So now I'm going to tell you some stories and every time I tell you a story I want you to think about it. I want you to make another picture inside your head of it.'

Yes problems:	No problems:
1.All cows quack Susie is a cow Does Susie quack? (Yes)	5.All cats bark George is a cat Does George meow? (No)
2.All the grass is blue. The garden is covered in grass. Is the garden blue? (Yes)	6.All water feels dry Ann had her hand in some water Does Ann's hand feel wet? (No)
3.All ice feels hot. Jill has her hand in some ice. Does Jill hand feel hot? (Yes)	7.All snow is black Len is a snowman made of snow Is Len white? (No)
4.All zebras have spots on their back Alex is a zebra Is Alex spotty? (Yes)	8.All hedgehogs feel soft Harry is a hedgehog Does Harry feel prickly? (No)

## **APPENDIX J**

### **COUNTERFACTUAL SCENARIOS USED IN EXPERIMENT 4**

#### **Scenario 2:**

This is Lucy. One day, in her house the floor was nice and clean like that. But guess what! Lucy comes home and she doesn't take her shoes off. She comes inside and she makes the floor all dirty with her shoes.

Before control question: Was the floor clean before?

Now control question: Is the floor clean now?

Test question: What if Lucy had taken her shoes off – would the floor be dirty? (no)

#### **Scenario 3:**

A turtle has some chocolate; he holds the chocolate in her mouth like this. But guess what! He drops the chocolate and it falls on the floor.

*Now control question:* Where is the chocolate now? Can you show me?

*Before control question:* Where was the chocolate before? Can you show me?

*Test question:* What if the Turtle had not dropped the chocolate – where would the chocolate be? Can you show me?

#### **Scenario 4:**

Robert the bunny lives in this box on the magic farm. But guess what! One day the door is not shut properly. So the bunny rabbit opens the door and he goes to sit here, outside the box.

*Now control question:* Where is the bunny rabbit now? Can you show me?

*Before control question:* Where was the bunny rabbit before? Can you show me?

*Test question:* What if the door of the box had been shut properly – where would the bunny rabbit be? Can you show me?

## APPENDIX K

### SECOND ANALYSIS FOR EXPERIMENT 4

The data from experiment 4 can be analysed in a different way. It can be decided to remove all the participants who failed at least one control question in the counterfactual scenario and all the participants who presented a yes or a no bias. By doing this the sample was reduced to 36 participants.

The mean and the standard deviation were calculated for each type of task in each context and the results are presented in the following table.

	Standard context (n= 18)	Fantasy context (n=18)
Counterfactual tasks	2.57 (1.53)	2.00 (1.13)
Contrary-to-fact syllogisms	.89 (.80)	1.44 (1.29)

Table 6: Mean and standard deviation of the scores for each type of tasks in the two different contexts.

A mixed design ANOVA 2\*2 (context \* task) was computed and showed a significant main effect of task,  $F(1, 35) = 15.52, p < .01$ .

The test of within-subject contrast showed that this difference was statistically significant,  $F(1, 35) = 23.18, p < .01$ . The results suggested that the participants found the counterfactual task easier than the contrary-to-fact syllogisms independently of the context.

The test of between-subjects showed that this effect was not significant,  $F(1, 35) = .003, p = .959$ . There was no effect of context on its own.

The statistical analysis showed a near significance interaction between task and context,  $F(1, 35) = 3.942, p = .055$ . This result suggested that the context might not have the same effect on each type of task.

## REFERENCES

- American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders (third edition)*. Washington, DC: APA.
- Astington, J. W. (2003). Sometimes necessary, never sufficient: false belief understanding and social competence. In B. Repacholi and V. Slaughter (Eds.), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. New York: Psychology Press.
- Baron-Cohen, S. (1987). Autism and symbolic play. *British Journal of Developmental Psychology*, 5, 139-148.
- Baron-Cohen, S. (1994). Infantile autism. In A. A. Lazarus, A. M. Colman (1994). *Abnormal Psychology*. London: Routledge.
- Baron-Cohen, S. (1995). *Mindblindness: An essay in autism*. Cambridge, MA: MIT Press.
- Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg and D.J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford: Oxford University Press.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the Autistic-Child Have A Theory of Mind. *Cognition*, 21, 37-46.
- Bartsch, K. & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, 60, 946-964.
- Birch, S.A.J., and Bloom, P. (2004). Understanding children's and adult's limitations in mental states reasoning. *Trends in Cognitive Sciences*, 8, 6, 255-260.
- Bloom, P. & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77, B25-B31.



- Brent, E., Rios, P., Happé, F., Charman, T. (2003). Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism*, 8 (3), 283-299.
- Byrne, R. M. J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, 6, 10, 426-431.
- Carpendale, J., and Lewis, C. (2006). *How children develop social understanding*. Oxford: Blackwell Publishing.
- Courtin, C., and Melot, A.M. (1998). Development of theory of mind in deaf children. In *Psychological perspectives on deafness*. Eds M. Marschark and M.D. Clark, 79-102. Mahawah, NJ: Lawrence Erlbaum.
- Deleau, M. (1996). L'attribution d'états mentaux chez les enfants sourds et entendants: Une approche du rôle de l'expérience langagière sur la théorie de l'esprit. *Bulletin de Psychologie*, 5, 48 -56.
- Dias, M. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British of Journal of Developmental Psychology*, 6, 207-221.
- Dias, M. G. & Harris, P. L. (1990). The Influence of the Imagination on Reasoning by Young-Children. *British Journal of Developmental Psychology*, 8, 305-318.
- Evans, J. St. B. T. (1989). *Bias in Human Reasoning: Causes and Consequences*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Evans, St. B. T. (1998). Matching bias in conditional reasoning: do we understand it after 25 years? *Thinking and Reasoning*, 4 (1), 45-82.
- Evans. St. B. T., and Handley, S.J. (1999). The role of negation in conditional inference. *Quarterly Journal of Experimental Psychology– Human Experimental Psychology*, 53A(3), 739-769.
- Evans, J. St. B. T., Newstead, S. E., Byrne, R. M. J. (1997). *Human Reasoning: The Psychology of Deduction*. Hove: Lawrence Erlbaum Associates.

- Evans, J. St. B. T., and Over, D. E. (1996). *Rationality and Reasoning*. Hove, UK: Psychology Press.
- Flynn, E., O'Malley, C., and Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science*, 7:1, 103-115.
- Fodor, J.A. (1983). *The modularity of the mind: An essay on faculty psychology*. Cambridge: The MIT Press.
- Keenan, T. (2003). Individual differences in theory of mind. In B. Repacholi and V. Slaughter (Eds.), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. New York: Psychology Press.
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994a). The children's test of nonword repetition: a test of phonological working memory. *Memory* 2[2], 103-127.
- Gathercole, S.E, and Adams, A-M. (1994b). Children's phonological working memory: Contribution of long-term knowledge and rehearsal. *Journal of Memory and Language*. 33 (5), 672-688.
- German, T. (1999). Children's causal reasoning: counterfactual thinking occurs for 'negative' outcomes only. *Developmental Science*, 2, 442-447.
- Gopnik, A. & Astington, J. W. (1988). Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction. *Child Development*, 59, 26-37.
- Goswami, U. (1998). *Cognition in Children*. Hove: Psychology Press Ltd.
- Goswami, U. (2002). Inductive and deductive reasoning. In U. Goswami (Eds.), *Childhood Cognitive Development*. Oxford: Blackwell Publishing.
- Grant, C. M., Riggs, K. J., and Boucher, J. (2004). Counterfactual and mental state

- reasoning in children with autism. *Journal of Autism and Developmental Disorders*, 34, 2, 177-188.
- Guajardo, N. R., and Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, 19, 53-80.
- Happé, F. (1994a). *Autism: an introduction to psychological theory*. London: UCL Press Limited
- Happé, F.G.E. (1994b). An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129-154.
- Harris, P. L. (2000). *The Work of Imagination*. Frome: Blackwell Publishing.
- Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61, 233-259.
- Hawkins, J., Pea, R., D., Glick, J., & Scribner, S. (1984). 'Merds that laughs don't like mushrooms': Evidence for deductive reasoning by preschoolers. *Developmental Psychology*, 20 (4), 584-594.
- Kahnemann, D. and Tversky, A. (1982). The simulation heuristics. In Kahneman D., Slovic P. and Tversky A. (Eds.). *Judgement under uncertaintyL Heuristics and Biases*. (p.201-208). New York: Cambridge University Press.
- Keenan, T. (2003). *Individual Differences in Theory of Mind: The Preschool Years and Beyond*. In B. Repacholi and V. Slaughter (Eds.), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. New York: Psychology Press.
- Kline, P. (2000). *Handbook of Psychological Testing*. London: Routeledge.
- Kucjaz, S.A., and Daly, M.J. (1979). The development of hypothetical reference in

- the speech of your children. *Journal of Child Language*, 6, 563-579.
- Koslowski, B., and Masnick, A. (2002). The development of causal reasoning. In *Childhood Cognitive Development*. Eds U. Goswami. (2002). Oxford: Blackwell Publishing.
- Leevers, H. J. & Harris, P. L. (1999). Persisting effects of instruction on young children's syllogistic reasoning with incongruent and abstract premises. *Thinking-and-Reasoning*, 5, 145-173.
- Leevers, H. J. & Harris, P. L. (2000). Counterfactual syllogistic reasoning in normal 4-year-olds, children with learning disabilities, and children with autism. *Journal of Experimental Child Psychology*, 76, 64-87.
- Leslie, A. M. (1994). Pretending and Believing - Issues in the Theory of Tomm. *Cognition*, 50, 211-238.
- Leslie, A. M, and Polizzi, P. (2000). Inhibitory processing in false-belief task: Two conjectures. *Developmental Science*, 1:2, 247-253.
- Lewis, C. and Osborne, A. (1990). Three-year-olds' problems with false belief: conceptual deficit or linguistic artifact? *Child Development*, 61, 1514-1519.
- McKinnon, M., and Moscovitch, M. (2007). Domain-general contributions to social reasoning re-explored. *Cognition*, 102, 179-218.
- Markman, K.D., Gavanski, I., Sherman, S.J., and McMullen, M.N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, 71, 450-463.
- Markovits, H. (1995). Conditional reasoning with false premises: fantasy and information retrieval. *British Journal of Developmental Psychology*, 13, 1-11.
- Markovits, H. & Vachon, R. (1989). Reasoning with Contrary-To-Fact Propositions. *Journal of Experimental Child Psychology*, 47, 398-412.

- Markovits, H., Venet, M., JanveauBrennan, G., Malfait, N., Pion, N., & Vadeboncoeur, I. (1996). Reasoning in young children: Fantasy and information retrieval. *Child Development, 67*, 2857-2872.
- McMullen, M. N., Markman, K.D., and Gavanski, I. (1995). Living in neither the best nor the worst of all possible worlds: antecedents and consequences of upward and downward counterfactual thinking. In N.J. Roese and J.M. Olson (Eds). *What might have been: the social psychology of counterfactual thinking* (p. 133-167). Mahwah: Erlbaum.
- Mitchell, P. (1997). *Introduction to Theory of Mind: Children, Autism and Apes*. London: Arnold.
- Peterson, C.C., and Siegal, M. (1995). Deafness, conversation and the theory of mind. *Journal of Child Psychology and Psychiatry, 36*, 459-474.
- Peterson, C.C., and Siegal, M. (1995). Changing focus on the representational mind: Concepts of false photos, false drawings and false beliefs in deaf, autistic and normal children. *British Journal of Developmental Psychology, 16*, 301-320.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge MA: MIT Press.
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). 3-Year-Olds Difficulty with False Belief - the Case for A Conceptual Deficit. *British Journal of Developmental Psychology, 5*, 125-137.
- Perner, J., Sprung, M., Steinkogler, B. (2004). Counterfactual conditionals and false belief: a developmental dissociation. *Cognitive Development, 19*, 179-201.
- Perner, J. and Wimmer, H. (1985). 'John thinks that Mary thinks that...' Attribution of second-order beliefs by 5-10 year-old children. *Journal of Experimental Child Psychology, 39*, 437-471.
- Peskin, J. and Ardino, V. (2003). Representing the mental world in children's social

- behaviour: Playing hide-and-seek and keeping a secret. *Social Development*, 12 (4), 496-512.
- Peterson, C.C and Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry*, 36, 459-474.
- Peterson, D.M., and Bowler, D.M. (2000). Counterfactual reasoning and false belief understanding in children with autism. *Autism*, 4(4), 391-405.
- Peterson, C.C and Siegal, M. (1999). Representing inner worlds: theory of mind in autistic and normal hearing children. *Developmental Science*, 10, 126-29.
- Pollard, P. and Evans, J. St. B. (1980). The influence of the origin of logic on conditional reasoning performance. *The Quarterly Journal of Experimental Psychology*, 32 (4), 605-624.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural Brain Sciences*, 1, 515-526.
- Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 13, 73-90.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, (1), 133-148.
- Roese, N.J. and Olson, J.M. (2003). Counterfactual thinking. In L. Nadel, D. Chalmers, P. Culicover, B. French, and R. Goldstone. (Eds). *Encyclopedia of cognitive science*. (p.858 – 861). New York: Macmillan.
- Scott, F. J., Baron-Cohen, S., & Leslie, A. (1999). 'If pigs could fly': A test of counterfactual reasoning and pretence in children with autism. *British Journal of Developmental Psychology*, 17, 349-362.
- Scholl, B.J., and Leslie, A.M. (2001). Modularity, development and 'theory of mind'.

*Mind and Language*, 14, (1), 131-153.

Scholl, B.J., and Leslie, A.M. (2001). Minds, modules, and meta-analysis. *Child Development*, 72, 696-701.

Slaughter, V., & Repacholi, B. (2003). *Introduction : Individual Differences in Theory of Mind. What are we investigating?* In B. Repacholi and V. Slaughter (Eds.), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. New York: Psychology Press.

Steele, S., Joseph, R. M., and Tager-Flusberg, H. (2003). Brief report: Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders*, 33, 4, 461-467.

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

Wellman, H.M. (2002). Understanding the psychological world: developing a theory of mind. In U. Goswami (Eds), *Childhood Cognitive Development*. Oxford: Blackwell Publishing.

Wellman, H.M., and Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, 30, 239-277.

Wellman, H. M. and Bartsch, K. (1994). Before belief: Children's early psychological theory. In C. Lewis & P. Mitchell (Eds). *Children's early understanding of the mind: Origins and Development*. Hove: Lawrence Erlbaum Associates Ltd.

Wellman, H. M., Cross, D., & Watson, J. (2000). A meta-analysis of theory of mind development: The truth about the false-belief. *Child Development*, 72, 655-684.

Wellman, H., M., & Lagattuta, K., H. (2000). *Developing understandings of mind*. In Baron-Cohen, S., Tager-Flushberg, H., & Cohen, D., J., *Understanding other minds: perspectives from developmental cognitive neuroscience*. 2<sup>nd</sup> edition. New York: Oxford University Press.

Wellman, H., M., & Liu, D. (2004). Scaling of Theory-of-mind Tasks. *Child Development*, 75, 2, 523-541.

- Wellman, H., M., and Woolley, J., D. (1990). From simple desires to ordinary beliefs: the early development of everyday psychology. *Cognition*, 35, 245-275.
- 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.
- Woolfe, T., Want, S., and Siegal, M. (2002). Signposts to development: theory of mind in deaf children. *Child Development*, 73, 768