



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

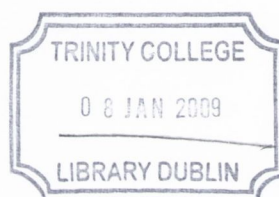
I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

**Attention and Reasoning: A task-switching
investigation of the mental model theory**

Alexandre Pereda Baños

A thesis submitted for the degree of Doctor of Philosophy of the
University of Dublin, Trinity College Dublin. Research carried
out in the School of Psychology and Institute of Neuroscience

2008



THESIS
8084

Declaration

I declare that this work has not been submitted previously as an exercise for degree at this or any other university and that it is entirely my own work. I give my permission for the Trinity College Library to lend or copy this thesis upon request.

Signed

Date

Alex 

19-09-2008

Acknowledgements

First of all I have to thank my supervisors, Hugh Garavan and Ruth Byrne for allowing me the privilege of spending three years of my life completely devoted to my own research interests. Thanks for your teachings, for your encouragement, and for your patience too.

Thanks to all my colleagues, Gloria, Liam, Eric, Cristina, Elena, Michelle, Julie and Sarah, for your kindness, for so much help received and for the efforts you have all made in understanding my often untidy English. Thanks also go to the School of Psychology and TCIN crews, June, Eddie, Marcella, Pat, Maeve, and Barbara, who we sadly missed last year.

I cannot be thankful enough to Mary and Caren. You made my transition to Dublin very (perhaps too) easy. Without your friendship and your support these three years would not have been at all such an enjoyable time.

Thanks to Borja, Davinia, Giovanna, Ignasi and Abdullah for the fun times. And last, but not least, thanks to mum, dad and to my brother for making it all possible.

Summary

Despite the fact that most theories of reasoning agree on the main role played by limitations in WM capacity in reasoning performance, little research has been directed to elucidate how the executive processes that deal with these limitations contribute to complex mental operations such as calculation or deduction. The present thesis aims to demonstrate how the methods employed in the investigation of one such executive process, namely, attention switching, may help to shed light on some debated questions about the way people tackle reasoning problems. Particularly we are interested in those questions referring to the influence of attentional limitations in the ease with which reasoning problems can be solved.

In chapter 2 we present three behavioral studies aimed to obtain a direct test the mental model theory's principle of parsimony, and of the nature of the mental representations and processes involved in carrying out conditional inferences from indicative and counterfactual conditionals. The studies employs of RT measures and a task-switching paradigm. The results support the referred hypotheses and demonstrate the potential utility that the methods employed might have in answering questions about the nature of the representations and processes involved in human reasoning.

The main proposal of the behavioral studies presented in chapter 3 is that, what the mental model theory refers to as the principle of truth, may reflect a basic cognitive phenomenon. Consequently, this tendency should be reflected on the latency of responses in a simple categorization task, not only in "higher-level" processes such as deductive reasoning or hypothesis-testing. Results confirm this hypothesis. People are faster to confirm than to reject that a given triple of numbers follows a previously specified rule. Furthermore, switching between response modes (confirmation and rejection) entails an asymmetrical switching cost, which is interpreted to reflect that confirmation is a less demanding operation than rejection. Experiments 2 and 3 reveal that these response mode switching costs are due to a general switching process which is also evident in switching between triple categorization rules. Finally in experiment 4, we establish some boundary conditions for the observed switching costs, specifically, we address the role of the interference between rules.

In chapter 4 we aim to provide convergent evidence to the results reported in chapter 2 and the interpretations we derived from them by means of an fMRI experiment. More specifically, we were first interested in comparing the effects on brain

activation of switching between rules governing the behaviour of number triples, and switching between response modes (confirmation or rejection of the triple as an instance of the tested rule). Results support the view that attention-switching is related to widespread areas in frontal and more posterior regions, which were mostly common for the two types of switching observed. Also, asymmetrical switching costs seemed to be reflected in the activation level some of the areas involved in switching.

In summary, this thesis: provides strong convergent evidence that the mental model theory's principles of parsimony and truth reflect basic cognitive phenomena; demonstrates the close relationship between different types of attention switching by showing that object and task switching costs by are operated by a common switching mechanism; unveils a direct relation between asymmetric costs and brain activation and an object similarity effect for object switching; and finally, highlights the fact that hypotheses about the representations and processes involved in reasoning should consider specific models of WM.

Table of contents

	Page
Declaration	i
Acknowledgements	ii
Summary	iii
Table of Contents	v
List of Tables	ix
List of Figures	x
Chapter 1	
Introduction	1
Attention-switching	3
Reasoning	6
<i>Deductive Reasoning</i>	6
<i>The Mental Model Theory of Deduction</i>	7
<i>Conditional Reasoning</i>	8
<i>Counterfactuals</i>	10
<i>Hypothesis Testing</i>	11
Attention and Reasoning	14
<i>The Psychometric Perspective</i>	14
<i>Dual-Task Methodology</i>	16
<i>Other Methodologies</i>	18
Organisation of the thesis	20
Chapter 2	
An attention-switching study of mental representations of indicative and counterfactual conditionals	23
Introduction	23
Switching between conditional inferences	24
Experiment 1	28
<i>Method</i>	29

<i>Participants</i>	29
<i>Materials</i>	29
<i>Design</i>	30
<i>Procedure</i>	30
<i>Results</i>	31
<i>Discussion</i>	32
Experiment 2	34
<i>Method</i>	35
<i>Participants</i>	35
<i>Materials</i>	35
<i>Design</i>	36
<i>Procedure</i>	36
<i>Results and Discussion</i>	37
Experiment 3	39
<i>Method</i>	40
<i>Participants</i>	40
<i>Materials and Procedure</i>	40
<i>Design</i>	40
<i>Results</i>	41
<i>Discussion</i>	45
General Discussion	45
 Chapter 3	
Attentional Limitations and True Possibilities	49
Introduction	49
Hypotheses and overview of the experiments	50
Experiment 4	52
<i>Method</i>	52
<i>Participants</i>	52
<i>Materials</i>	52
<i>Design</i>	53
<i>Procedure</i>	53
<i>Results</i>	54
<i>Discussion</i>	55

Experiment 5	56
<i>Method</i>	56
<i>Participants</i>	56
<i>Materials</i>	56
<i>Design</i>	57
<i>Procedure</i>	58
<i>Results</i>	58
<i>Discussion</i>	59
Experiment 6	59
<i>Method</i>	60
<i>Participants</i>	60
<i>Materials</i>	60
<i>Design</i>	60
<i>Procedure</i>	61
<i>Results</i>	61
<i>Discussion</i>	64
Experiment 7	65
<i>Method</i>	66
<i>Participants</i>	66
<i>Materials</i>	66
<i>Design</i>	67
<i>Procedure</i>	67
<i>Results</i>	67
<i>Discussion</i>	70
General Discussion	71
<i>Implications for reasoning</i>	71
<i>Implications for attention-switching</i>	75

Chapter 4

An fMRI investigation of rule and response mode switching costs	78
Introduction	79
Functional magnetic resonance imaging (fMRI)	80
Neural basis of attention-switching	80

Brain activity related to confirmation versus rejections	81
Experiment 8	82
<i>Method</i>	83
<i>Participants</i>	83
<i>Task design</i>	83
<i>fMRI parameters</i>	85
<i>fMRI analyses</i>	85
<i>Results</i>	88
<i>Behavioral Results</i>	88
<i>Event-related functional activation</i>	90
<i>Block-related functional activation</i>	94
Discussion	96
<i>The role of the frontal and parietal lobes in attention-switching</i>	96
<i>Brain correlates of asymmetrical switching costs</i>	99
<i>Cingular activity and rejection</i>	100
Conclusions	101
 Chapter 5	
General Discussion	103
Summary of the thesis findings	104
Implications for reasoning	106
Implications for attention	110
Attention and reasoning	114
Concluding comments	117
 References	119
 Appendix	
Instructions for the experiments presented in chapter 2	137

List of Tables

Table	Page
Table 4.1	Clusters active for bout response mode and rule switching
Table 4.2	Clusters of significant activation for both switching conditions
Table 4.3	Task activation for “high confirmation” and “high rejection” blocks

List of Figures

Figure	Page
Figure 2.1	Interaction between Inference and Inference Switching in exp. 1
Figure 2.2	Interaction between Inference and Inference Switching in exp. 2
Figure 2.3	Interaction between Inference and Inference Switching for indicative and counterfactual conditionals in exp. 2
Figure 2.4	Interaction between inference and inference Switching in exp. 3
Figure 2.5	Interaction between Inference and Inference Switching for indicative and counterfactual conditionals in exp. 3
Figure 2.6	Interaction between Mood and Mood Switching in exp. 3
Figure 2.7	Interaction between Mood and Mood Switching for MP and MT trials
Figure 3.1	Reaction times to confirm and reject triples on repeat and shift trials in exp. 4
Figure 3.2	Reaction times to confirm and reject triples and response mode repeat and response mode shift trials in exp. 5
Figure 3.3	Reaction times to confirm and reject triples for trials under the ascending and descending rules
Figure 3.4	Reaction times to rule repetition and rule shift trials for Response mode shift and repeat trials
Figure 3.5	Reaction times to the 3 levels of the Response Mode variable
Figure 3.6	Reaction times to the 3 levels of the Response mod variable for the two rules in exp. 7
Figure 3.7	Switching costs for the three levels of Response Mode variable in exp. 7
Figure 4.1	Switching costs for confirmations and rejections in the response mode switching condition

- Figure 4.2 Switching costs for both rules in the rule switching condition
- Figure 4.3 Clusters activated by rule and response mode switching.
-

Chapter 1 Introduction

One of the main explanatory constructs of human cognitive functioning is the limitation in the capacity to process information. Consider an everyday example: over the last few years the ban on the use of mobile phones while driving a motor vehicle has been enforced in many European countries. This legislation acknowledges the well known fact that we can only deal with a limited amount of information at a given moment. Even though most of us would argue that, in time, driving becomes an automatic skill and that we can deal with the demands of simultaneously maintaining a telephone conversation, we would surely reconsider if runover by a driver who was on his mobile at the time. This difficulty in dealing simultaneously with large amounts of information has often been referred to by cognitive psychologists as a limitation in working memory (WM) capacity (Baddeley, 1986; 2002). WM is widely understood as a state of higher activation in which currently relevant mental representations are held. There is also evidence of a further division within WM, the focus of attention, which would be a yet higher state of activation, often equated to consciousness, whose capacity is of just one element (e.g., Cowan, 1995; Garavan, 1998; McElree and Doshier, 1989; Oberauer, 2002; but see Cowan 2001). Importantly, in order to deal with this capacity limitations, a set of processes is needed in order to switch, inhibit, update and manipulate information, the so-called “executive processes”.

The present work aims to demonstrate how the methods employed in the investigation of one such executive process, namely, attention switching, may help to shed light on some debated questions about the way people tackle reasoning problems. In particular, we are interested in those questions referring to the influence of attentional limitations in the ease with which reasoning problems can be solved. Despite the fact that most theories of reasoning agree on the main role played by WM capacity limitations in explaining reasoning performance, little research has been directed to elucidate how the executive processes that deal with these limitations contribute to complex mental operations such as calculation or deduction. We hope that this work will constitute a step in this direction.

One likely cause for this lack of research is what some have referred to as the “divide and conquer strategy” in cognitive science. Briefly, many psychologists work under the assumption that the human cognitive system is divided into a series of independent modules (e.g., Fodor, 1983), whose functioning can be understood in isolation from other modules’ influence. This idea is reflected in the very organization of cognitive science, with attention, memory, reasoning, etc., constituting independent fields of research. This strategy has been useful in improving our comprehension of human cognition, however, an effort is required as well to integrate the knowledge each discipline has acquired about the cognitive system. This thesis constitutes an attempt to build a bridge between the fields of attention and reasoning, again, by showing how the methodologies of the former can be of much use in answering some of the questions posed by the latter. However, the aim of this thesis is not only methodological, we also

Attention and Reasoning

hope to highlight the important role played by executive processes (in this case, attention switching) in the performance of reasoning problems.

In this thesis we present three series of investigations, a set of three experiments that examine attention switching when making conditional inferences; a set of four experiments that examine attention switching in a hypothesis categorization task; and an fMRI study that examines the neural correlates of some of the effects discovered in the behavioral studies. This chapter starts with an introduction to the experimental methodology employed in all three sets of studies, namely, the *attention-switching* paradigm. We will then turn our attention to the field of reasoning, with special emphasis on the theoretical framework for this thesis, namely, the *mental model theory*. The different experimental approaches to the issue of the relation between attention and reasoning will then be briefly reviewed. Finally, an overview of our own methods and hypotheses will be offered.

Attention-switching

Capacity limitations in cognitive processing have been known to psychologists since the early days of scientific psychology. We need go no further than William James' seminal work, *Principles of Psychology* (1890), to find the hypothesis that we can only focus our attention on just one "object" at a time. Soon after the inception of the "cognitive revolution", results obtained with the Sternberg memory task (e.g., Sternberg, 1969) and the visual search paradigm (e.g., Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977) suggested a serial comparison of items in WM that necessitated switches between these items. However, only in the last decade have we witnessed a surge of interest in the

attention-switching process, and many methodologies have been developed in order to investigate its nature. Two of these methodologies are of key relevance in this thesis, the “object-switching” and the “task-switching” paradigms.

When people maintain two counts while switching randomly between them, there is a sizeable increase in reaction time when switching between counts, that is count-switching trials are slower than count-repeating trials. These “object-switching costs” may reflect the operation of a switching mechanism that controls the focus of attention. Recall that the focus of attention is a subset of WM whose capacity is generally believed to be only as large as one single element (Cowan, 1995; Garavan, 1998; McElree and Doshier, 1989; Oberauer, 2002; but see Cowan, 2001).

When people are shown a series of number-letter pairs (e.g., A7), and are asked to perform two different tasks depending on the colour of the stimulus, there is an increase in reaction time when switching between tasks, the task-switching cost (e. g., Allport, Styles and Hsieh, 1994; Gonzalez, Milan, Pereda and Hochel, 2005; Rogers and Monsell, 1995; Tornay and Milan, 2001; see Monsell, 2003, for a review). For example, if the stimulus is red the task is to say whether the letter is a vowel or a consonant, whereas if the stimulus is blue the task is to say whether the number is odd or even. There is a cost for trials in which the task switches (e.g., blue T3 → red U4), that is, task switching trials are slower than trials in which the task is the same as the previous trial (e.g., red N9 → red E5). The task-switching cost may represent a fundamental constraint with respect to the possibility of representing within WM more than one task set at a given moment (e.g., Allport and Wylie, 1999; Rubinstein, Meyer and Evans, 2001; Mayr and Kliegl, 2000).

Note that in the object switching paradigm participants perform the same operation (e.g., counting) throughout the experiment while switching the object of this operation, whereas in the task-switching paradigm the operation itself switches (e.g. from an even-odd to a vowel-consonant categorization). A noticeable difference between object and task switching costs is the existence of a pattern known as “task-switching asymmetry” (e. g., Allport & Wylie, 2000; Yeung and Monsell, 2003), which is typically observed when switching between tasks of different dominance or “task-strength”. The concept of dominance makes reference to a relation between two tasks, such that one of them is more practiced and/or automatic. For example, in the well known Stroop paradigm (Stroop, 1935) the stronger task is to read the word, as compared to naming the colour. This asymmetrical pattern of reaction time switching costs reflects the somewhat counterintuitive fact that *it is easier to switch into the more difficult (weaker) task from the easier (stronger) task than to switch in the opposite direction*. The fact that this is not observed in object-switching costs may be due to the objects consuming a similar amount of resources. Asymmetrical switching cost will be investigated in many of the experiments reported in this thesis.

We propose that the attention-switching methodology might be particularly useful for providing convergent evidence to test the hypotheses of the mental model theory of reasoning (e.g., Johnson-Laird, 1983; Johnson-Laird and Byrne, 1991; 2002) about the representations and processes involved in the correct performance of reasoning problems, to which we now turn.

Reasoning

Reasoning has been defined as the central activity in intelligent thinking, an activity that becomes necessary every time we need to solve a problem whose solution is not stored in memory (Evans, 1993). We can find many definitions of reasoning: a form of goal-directed thinking which always starts from a series of observations or premises, and which is different from associative or creative thinking (Johnson-Laird & Byrne, 1991); a form of thinking consisting of explicit thought processes that are effectively equivalent to the application of rules of some formal system (Gilhooly, 1998). Following Johnson-Laird & Byrne (1991), we can identify three main forms of reasoning: calculation, deduction and induction. Our work will focus mainly on deductive and inductive reasoning: *deduction* is a reasoning process in which conclusions are drawn that necessarily follow from the information that is already known; *induction* consists in the derivation of general rules (hypotheses) from particular cases.

Deductive reasoning

Four main types of deductive reasoning theories have been examined extensively in recent years (see e.g. Johnson-Laird & Byrne, 1991; Evans, 1989): the *mental models* theory (e.g. Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; 2002); the *formal rules* theory (e.g. Braine & O'Brien, 1991; 1998; Rips, 1983; 1994), the *domain specific rules* theories (e.g. Cheng & Holyoak, 1985; Cosmides, 1989; Fiddick, Cosmides & Tooby, 2000); and the "*heuristics-and-biases*" approach (Pollard, 1982; Evans, 1989, 2003). Both the mental models and formal rules approaches relate the difficulties that people find when solving reasoning problems to the limitation in WM capacity. This tenet

differentiates these two approaches from the third one. The role ascribed to WM limitations by domain-specific reasoning mechanisms is unclear, given that they are triggered by a particular type of *content*, they may place little demand on attentional resources. Heuristics-and-biases approaches such as Evans' dual-process theory (e.g., Evans, 2003) emphasize the role played by preattentive processes that select task relevant information, therefore, this approach does not consider the influence of attentional limitations. Unclear is the role ascribed by WM as well by the *expected information gain model* (Oaksford and Chatter, 02), this model states that cognitive behaviour is adapted to the structure of the environment, it does not depend on rule or models, thus, participants are not testing rules but seeking information that will help them update their beliefs and to obtain a gain in information that decreases their uncertainty.

The hypotheses put forward in this thesis were motivated by the framework of the mental model theory. We will briefly introduce its main tenets, as well as those specific aspects of the theory on which this work will be based.

The mental model theory of deduction

According to the mental model theory (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991, 2002; Johnson-Laird, Byrne & Schaeken, 1992; 1994), the understanding of an assertion is achieved by constructing a *mental model*, which is a representation of the state of affairs described by the premises. This representation has the same structure as the situation it represents (Johnson-Laird, 1983). The theory makes no specific prediction about the modality of these representations: "*The tokens of mental models may occur in a visual image, or they may not be accessible to consciousness. What matters is not the*

phenomenal experience, but the structure of the models." (Johnson-Laird & Byrne, 1991, p. 39). Reasoning is a semantic process based on the construction and manipulation of representations rather than a process based on the application of syntactic rules (Braine & O'Brien, 1981; 1998; Rips, 1983; 1994). In other words, meaning is more important than syntax. Importantly, these representations are constructed and manipulated in a working memory of limited capacity: "*What is needed is a theory that reconciles the semantics of truth tables with the constraints of mental processing*". Reasoners tend to reduce the cognitive load by representing explicitly only some amount of the to-be-represented information, as stated by the *principle of parsimony*. Moreover, mental models represent only what is true, as stated by the *principle of truth* (see e.g. Johnson-Laird & Byrne, 2002). That does not mean that a mental model cannot represent negations, it can, whenever the negation is a real possibility. A corollary of this principle is that "*...possibilities are psychologically basic, not truth values.*" (Johnson-Laird & Byrne, 2002, p. 653). The principle of truth will play an important role in the experiments reported in chapters 3 and 4 of this thesis. There, we will provide evidence that this tendency to think about true possibilities is a fundamental aspect of our cognitive functioning. First, chapter 2 will examine how people tackle conditional inferences, therefore, we provide a brief introduction to conditional reasoning next.

Conditional reasoning

People spontaneously use conditional inferences in their everyday reasoning, indeed, most of our knowledge is conditional (Johnson-Laird and Byrne, 2002). For example, you probably know that if you get caught on your mobile while driving, then you will be

finer. As Evans and Over (2004) point out, the capacity of exerting conditional reasoning allows us hypothetical thought and therefore distinguishes us from other animals.

Conditional reasoning research (e.g., Braine and O'Brien, 1998; Johnson-Laird and Byrne, 2002) has established that a conditional premise of the form "If A then B" is commonly interpreted by adults as being consistent with three possibilities: "there is an A and there is a B"; "there is not an A and there is not a B", and "there is not an A and there is a B". Thus the fully explicit set of mental models of this conditional premise could be represented like this:

a	b
not a	b
not a	not b

People do not represent what is false (principle of truth), that is, "there is an A and there is not a B". The mental model theory proposes that, due to limitations in WM capacity, people represent initially as few possibilities as they can (principle of parsimony):

a	b
...	

The ellipsis is referred to by the model theory as the *implicit model*, and it reflects the proposal that people are aware that further possibilities are consistent with the assertion. From this "material implication" interpretation two different inferences are warranted: "A, therefore B", known as *modus ponens*; and "not B, therefore not A"; known as *modus tollens*. The model theory can explain certain widespread effects that occur when people perform conditional reasoning tasks. For example, most people make

modus ponens (MP) inferences readily, but in experiments usually only half the participants make the modus tollens (MT) inference, the other half say that nothing follows. The relative ease of modus ponens over modus tollens may be explained by the fact that reasoners need to think about just one explicit possibility to reach a valid ponens deduction whereas in order to reach a valid tollens deduction it is necessary to *flesh out* the implicit model (Johnson-Laird & Byrne, 1991), a process that places further demands on processing resources. These proposals on the interpretation of conditionals have received extensive empirical support (see Johnson-Laird, 2000, for a review), usually by means of accuracy measures. We propose that RT measures and an attention-switching paradigm might be particularly well suited to provide convergent evidence to test this account. For example, if MP and MT involve different representations in terms of the number of possibilities people think about, and the processes required, a cost of switching between these inferences should be observed. However, the mental model theory also posits that there are situations in which MP and MT inferences are not as different as described, as when the conditional is presented in counterfactual mood, issue to which we now turn.

Counterfactuals

Counterfactual conditionals make reference to a situation that could have happened, but that did not. Counterfactuality is conveyed in English often by means of the subjunctive mood, as opposed to the indicative mood (i.e. “if there had been an A, there would have been a B”; as opposed to “if there was an A, there was a B”). Counterfactual reasoning plays a central role in people’s everyday thinking about past possibilities and past or

present impossibilities (Byrne, 1997, 2002, 2005), being implicated in a great variety of cognitive activities ranging from fantasy and creativity to deductive and probabilistic reasoning. Counterfactual reasoning is also closely related to causal thought aiding the prediction and prevention of future outcomes, and it can even amplify certain emotional states (see Byrne, 2002 for a review). However, our interest in counterfactuals stems from the hypothesis that they induce an initial representation of two possibilities, rather than a single possibility. According to the mental model theory, when presented with a counterfactual conditional people keep in mind two possibilities, the one described by the counterfactual, and the factual situation in which the conjecture has not happened.

not a not c (factual)

a c (counterfactual)

...

The main aim of the experiments reported in chapter 2 will be to provide convergent evidence to test the mental model theory’s hypotheses about the representation of indicative and counterfactual conditionals, with reaction time (RT) measures and by means of an attention-switching paradigm. Specifically, we suggest that the pattern of costs observed when switching between MP and MT inferences should be affected by the mood of the conditional. RT measures have seldom been employed in deductive reasoning tasks, which usually depend on the manipulation of the number of models or inferential steps required to solve a reasoning task and the observation of the effects of this manipulation on the accuracy of reasoners conclusions. A complementary objective of the experiments in chapter 2 is to demonstrate that the hypotheses of the mental model theory can also be tested by RT data.

Hypothesis testing

The experiments presented in chapters 3 and 4 aim to test the idea that the tendency to think about true possibilities, one of the causes of MT inferences being so demanding (when in the indicative mood), is in fact a fundamental feature of the cognitive system. This test will be achieved by means of a novel computerised version of a well known hypothesis-testing paradigm, Wason's 2-4-6 task (Wason, 1960).

In the 2-4-6 or rule discovery task participants must test hypotheses to discover a rule. They are given an instance of the rule: 2, 4, 6. They test their hypotheses by generating triples. The experimenter tells them for each triple whether it is an instance of the rule or not. The task reflects some of the procedures of scientific discovery, where the participant is like a scientist attempting to work out the general rule to explain data. The rule is deliberately a very general one, 'ascending numbers' but most people generate the hypothesis that the rule is a specific one such as 'ascending even numbers' and they tend to generate triples such as '8, 10, 12'.

Only 6 of the 29 participants in Wason's original experiment gave the correct rule, "ascending numbers", after their first announcement. Participants who failed often tested few triples and used tests that were more likely to yield a "yes" response. Thus, most of the participants in the experiment were apparently unable or unwilling to use a falsifying strategy (that is, testing triples that would disconfirm their hypothesized rule), an observation that was referred to as "*confirmation bias*" by Wason (1960).

Klayman and Ha (1987) proposed that participants followed a *positive test strategy* (Klayman and Ha, 1987, 1989). Consider the inversion of the relation between the experimenter's rule and the participants hypothesis, for example, the experimenter's rule

Attention and Reasoning

is “ascending even numbers” and the participant’s hypothesis is “ascending numbers”. A positive test such as “1, 2, 3” falsifies the hypothesis. The difference between a “confirmation bias” and a “positive test strategy” account is subtle, what the former describes as a motivational tendency to prove ourselves right, the latter describes as a natural bias towards considering positive information that fails when applied in certain conditions.

There is a way to dramatically improve participants’ performance on the 2-4-6 task (Tweney, Doherty, Worner, Pliske, Mynatt, Gross, and Arrkelin, 1982). Instead of asking participants to figure out a rule by proposing triples and receiving positive or negative feedback, they were asked to discover two rules, arbitrarily called DAX and MED. Thus, rather than receiving positive or negative feedback, each triple was classified as fitting to one of the rules. With this version of the task, more than half of the participants made a correct first announcement of the rules. One can find hints throughout the hypothesis-testing literature of the idea that what is important in the success of this manipulation stems from its effect on the mental representation of the rules (e.g., Gorman, Stafford and Gorman, 1987; Tweney et al., 1982). Gorman et al. (1987) hypothesized that this manipulation fostered a unified mental representation of the two rules. On the other hand, Tweney et al. (1982) suggested that the logical structure of the problem was altered in a way that encouraged participants to switch from one hypothesis to another, thus avoiding a rigid testing of a single hypothesis. Evans (1989) proposes a *positivity bias*, a selective attention process that causes people to attend to positive rather than negative information. The dual-goal manipulation thus avoids the labelling of information as positive or negative.

We suggest that the mental model theory's principle of truth is key in explaining performance on the 2-4-6 task. Testing this suggestion is the aim of chapter 3, and again, this will be achieved by means of RT measures and a series of attention-switching experiments.

We can summarize this brief review on reasoning by noting that the concept of attention is central to reasoning theories in general and the mental model theory in particular. This theory proposes that attentional limitations and a tendency to focus on real possibilities are the key explanatory factors of deductive reasoning performance. As we will show in the next section, the interest on this close relation between attention and reasoning has fostered the development of various lines of research.

Attention and reasoning

Attentional limitations play a central explanatory role in the main theories of reasoning performance (e.g. Braine and O'Brien, 1998; Johnson-Laird & Byrne, 2002; Rips, 1994). Both mental models and formal rules theorists agree that the main cause of the widespread performance errors observed in reasoning tasks is the inability to maintain large amounts of information active simultaneously in WM. However, the interest in the specific relations between attentional and reasoning processes has only recently begun to develop. In what follows, we will provide a brief summary of this literature.

The Psychometric perspective

If WM limitations are the key explanatory factor of reasoning performance, then a positive correlation between WM capacity (WMC) and reasoning performance should be

expected. This relation between WMC and reasoning is strongly emphasized in a study by Kyllonen & Christal (1990) whose title is illustrative enough: “*Reasoning ability is – little more than- working memory capacity?*”. These authors assessed general reasoning ability, using various tests from the psychometric literature, and WMC by means of tests constructed by themselves according to Baddeley’s (1986) definition of working memory. It was found that the correlations between these two constructs were always higher than .80. Many studies have reported high correlations between WMC and reasoning ability (e.g., Conway, Kane and Engle, 2003; Engle and Kane, 2004; Engle, Kane and Tuholski, 1999; Oberauer, Süß, Wilhelm and Wittman, 2000; Oberauer, Schulze, Wilhelm and Süß, 2005; but see Ackerman, Beier and Boyle, 2005). WMC is often understood as an executive attention ability supporting the active maintenance of goal-relevant information, an ability that becomes essential in interference-rich situations in which correct responding cannot be achieved by automatic spreading of activation between representations or by habitual responding (see Ackerman et al., 2005; Conway et al., 2003 for reviews).

This psychometric perspective often described *g* or general reasoning ability as a high-level process eluding precise definition and, perhaps as a consequence of this, the development of good measures of reasoning ability has been almost an art form (Kyllonen and Christal, 1990; see also Stanovich, 1999). The Raven’s Progressive Matrices test, which involves identifying patterns, applying abstract rules and detecting subtle relationships is the prototypical measure of *Gf* for the psychometric tradition. Thus, reasoning in the psychometric literature is probably not the same construct we are referring to when describing theories of reasoning competence. However, there are some

studies that relate individual differences in WMC with performance in reasoning tasks which are more akin to the definition of reasoning in the reasoning literature. For example, Capon, Handley and Dennis (2003) conducted a study in which they assessed the relation between WMC as measured by a series of complex and simple verbal and spatial WM measures and syllogistic and five-term spatial inferences. They found that when WMC is related to performance in syllogistic and spatial reasoning tasks, rather than tasks measuring general reasoning ability of *Gf*, the correlations are still high. In this literature is often considered that the most suitable model of WM is Baddeley's (see e.g. Baddeley, 1986) with two slave systems (the phonological loop, which deals with verbal information; and the visuospatial sketchpad, which deals with visuospatial information) and a central executive. Reasoning performance is mostly related to the action of the central executive, which is widely understood as a mechanism of controlled attention as proposed by Engle and colleagues (e.g. Engle, Kane & Tuholski, 1999; Engle, Kane & Conway, 2003). The close relationship between the central executive and reasoning performance is also demonstrated by the set of studies reviewed in the next section.

Dual-Task methodology

The dual-task methodology was instrumental in the development of Baddeley's model of WM. Broadly, it involves the concurrent performance of tasks designed to disrupt specific subsystems of WM. Thus, a straightforward strategy to ascertain which WM subsystems are most related to reasoning would be to have participants solving reasoning problems while performing concurrent tasks disrupting either the verbal subsystem, the visuospatial subsystem or the central executive. Studies using this strategy have often

Attention and Reasoning

found that reasoning is mainly disrupted by tasks which are supposed to disrupt the central executive (e.g., Baddeley and Hitch, 1974; Hitch and Baddeley, 1974; Toms, Morris and Ward, 1993). Toms et al. (1993) interpreted their data as implying that the representations of reasoning problems are in the form of abstract tokens, which are at a higher level of representation than the visuospatial and phonological codes, with the memory load interfering with the strategic deployment of “memory constellations” and occupying some of an abstract and general workspace.

Further investigation has confirmed that secondary tasks loading the central executive are accompanied by an impairment of reasoning performance, for example, Klauer and colleagues (Klauer, Stegmaier and Meiser, 1997) compared the effects of different secondary tasks affecting the three WM components on propositional reasoning and on spatial reasoning. In one of their experiments (Klauer et al., 1996, cited in Klauer et al., 1997), they showed that a visual tracking task disrupted spatial reasoning but not propositional reasoning, and this effect was independent of the difficulty of the propositional reasoning task (see also, Gilhooly, Logie, Wetherick and Wynn, 1993; Gilhooly, Logie and Wynn, 1999).

It is worth mentioning an interesting set of studies devised by Barrouillet and Lecas (1999; Barrouillet, Grosset and Lecas, 2000), on the grounds of the mental model theory. These authors had their participants divided in three age groups (children, young adolescents and adults), and their task was to list all the possibilities compatible with a conditional statement. A measure of WMC was also obtained. The data showed that a conditional of the form “If A then B” is interpreted in different ways depending on age: children make the conjunctive interpretation (requiring 1 model), adolescents the

biconditional one (requiring 2 models) and adults make the conditional interpretation (requiring 3 models). Despite this, WMC was more predictive of reasoning performance than was age (Barrouillet and Lecas, 1999).

The main conclusion to draw from these studies is that the relation between reasoning performance and WM is mainly mediated by the central executive subsystem which is widely understood as a general pool of resources for controlled processing (e.g., Engle et al., 1999). Reasoning performance relates to WM inasmuch as it requires the deployment of attentional resources in order to enhance processing efficiency. We review now a set of studies that provide new ideas so as to how to study the relation between attention and reasoning.

Other methodologies

A direct demonstration of the close relation between attention and reasoning is provided by the “premise order” or “focusing” effect (Giroto, Mazzocco & Tasso, 1997; Legrenzi, Giroto & Johnson-Laird, 1993). This effect consists in the overcoming of the superiority of MP over MT inferences by changing the order of presentation of the premises. For example, when a conditional of the form “If there is a circle at the top of the card there is a triangle at the bottom” is presented before the premise “There is a square at the bottom of the card”, the usual poor performance on MT inferences is observed. However, if the order of presentation is switched (Giroto et al., 1997; Legrenzi et al., 1993) the superiority of MP disappears. A possible explanation (Legrenzi et al., 1993) is that the inverted order may lead individuals to switch the order of the elements in their models of the conditional. Therefore, an MT inference would not force reasoners to work

backwards from the most recent item to enter working memory to a previous item: “...they can work in the same direction as the information entered working memory from the denial of the consequent to the denial of the antecedent” (Legrenzi et al, 1993). Related results were obtained by Barrouillet, Grosset and Lecas (2000, exps 1 and 2; amongst other related references), who found that MT inferences took longer to verify than MP inferences. In other words, they found a strong directionality effect (see also Evans, 1993b), that is, inferences from antecedent to consequent (MP) took less time to verify than inferences in the opposite direction (MT).

The use of reaction time measures in reasoning research is not widespread, perhaps because of the view that theories of reasoning competence predict failures rather than slow responses (Toms et al., 1993). However, as Moreno-Ríos and García-Madruga (2002) point out, reasoning tasks as studied by theorists of reasoning competence were often made deliberately complex in order to detect differences in accuracy. There is an increasing number of studies based on latencies (e.g., Barrouillet et al., 2000; García-Madruga, Moreno, Carriedo, Gutiérrez & Johnson-Laird, 2001; Glenberg, Meyer & Lindem, 1987; Moreno-Ríos & García-Madruga, 2002; Quelhas & Byrne, 2000; Schaeken, Johnson-Laird & d’Ydewalle, 1996) that show that reaction time measures support the mental models theory’s predictions. For example, a series of studies have used the *priming* effect and latency measures in order to investigate reasoning performance. First, Moreno-Ríos & García-Madruga (2002), devised a simplified reasoning task in which the premises gave information about the relative position of three letters. The priming effect was supposed to rise, for example, from the spread of activation from premises of the form “B C –” or “–C–” to the model “B C A”. The results

showed a priming effect both when reasoners had to evaluate the premises (Moreno-Ríos & García-Madruga, 2002, exp. 1) and when they had to generate the conclusion (Moreno-Ríos & García-Madruga, 2002, exp. 2) and even when they were told to ignore the prime (Moreno-Ríos & García-Madruga, 2002, exp. 3). Another interesting usage of the priming methodology in reasoning was that of Galinsky & Moskowitz (2000), whom had some of their participants reading paragraphs about different counterfactual scenarios before facing the Duncker candle problem (Galinsky & Moskowitz, 2000, exp. 1) or the Wason's selection task (Galinsky & Moskowitz, 2000, exp. 2). The results showed that participants primed with a counterfactual were more likely to solve these problems, this result was interpreted as an effect of priming of a mental simulation mind-set in which relevant alternatives are more likely to be considered. In similar fashion, Santamaría, Espino and Byrne (2005) showed that counterfactual conditionals primed both the possibility described by the counterfactual, and the factual situation in which the conjecture does not happen.

With this section, we hope to have provided enough evidence of the close relation between reasoning and executive processing and of the necessity of further investigating how executive processes contribute to the successful performance of reasoning problems. In carrying this investigation, methodologies and measures not initially designed to deal with reasoning problems have already provided fruitful results, this thesis constitutes an attempt to further advance in that direction.

Organisation of the thesis

This thesis presents an investigation of the mental representations and processes involved in the performance of reasoning problems and categorization tasks. Specifically, we intend to provide convergent evidence to test some of the mental model theory's hypotheses about the main factors that constrain our ability to think, by means of a measure (RT) and a paradigm (attention-switching) that, despite being infrequently employed in reasoning studies, might be particularly well suited to shed light on these issues.

In chapter 2 we use the task-switching paradigm in order to provide a more direct test of the nature of the mental representations constructed for factual and counterfactual conditionals. Participants switched between *modus ponens* (MP) and *modus tollens* (MT) conditional inferences, presented either in subjunctive or indicative mood. We test the idea that people represent counterfactual conditionals by keeping two possibilities in mind, the one described by the conditional and the presupposed facts in which the conjecture has not happened, thus making the counterfactual MT inference easier to draw. This idea is tested by examining the pattern of costs observed when switching between MP and MT inferences, and between the indicative and counterfactual mood.

In chapter 3 it is proposed that the tendency to think about true possibilities, one of the causes of indicative MT inferences being so demanding, is in fact a fundamental feature of the cognitive system. We intend to observe this tendency in a task where participants have to categorize instances as *confirming* or *rejecting* a given rule. We predicted that people would be faster to confirm than to reject that a given triple of numbers follows a previously specified rule. We also tested whether switching between

response modes (confirmation and rejection) entails an asymmetrical switching cost (e.g., Allport et al., 1994; Meuter and Allport, 1999). In one of the experiments participants had to switch both between response modes (confirm, reject) and categorization rules (which we interpreted as equivalent to switch between tasks and objects respectively) and we tested whether these switching costs reflected the operation of a common switching mechanism.

In chapter 4 we employed fMRI in order to investigate the neural basis of the switching costs observed in chapter 3, that is, costs of switching between response modes and between categorization rules. It is suggested that, inasmuch as these two types of switching seem to be operated by the same mechanism, a similar activation map should be observed for each of them. Furthermore, it is expected that any asymmetrical pattern of response mode switching will be related to functional activation.

In the final chapter of this thesis we consider the implications of the experimental results both for the fields of reasoning and attention switching.

Chapter 2 – An attention switching study of mental representations of indicative and counterfactual conditionals

In this chapter we present three empirical studies aimed to obtain a direct test of the nature of the mental representations and processes involved in carrying out conditional inferences from indicative and counterfactual conditionals. Specifically, we intend to provide convergent evidence of the mental model theory's hypotheses on counterfactual thinking by means of RT measures and a task-switching paradigm.

Introduction

Counterfactual conditionals make reference to a situation that could have happened, but that finally did not. As outlined in the previous chapter, our interest in counterfactuals stems from the kind of mental representation they seem to induce.

Fillenbaum (1974) asked his participants to read sentences such as: "*If he had caught the plane he would have arrived on time*" and then gave them unexpectedly a memory task. The results showed that participants tended to falsely recognize the negated antecedents and consequents. According to the mental model theory, when presented with a counterfactual conditional people keep in mind two possibilities, the one described by the counterfactual, and the factual situation in which the conjecture has not happened.

not a	not c	(factual)
a	c	(counterfactual)
...		

People think of two possibilities to understand the counterfactual, whereas they think of only one initially to understand the indicative conditional. The mental model theory predicts that the ease with which participants can make MP and MT inferences should not be different from a counterfactual. They should make each one readily, because they do not need to flesh out their models. This hypothesis has received extensive empirical support. For example, Byrne and Tasso (1999, exp. 1), had their participants solve a series of reasoning problems based on the four basic inferences from conditionals, including the two of interest in this thesis, that is: modus ponens (he caught the plane, therefore he arrived on time); and modus tollens (he did not arrive on time, therefore he did not catch the plane). Half of the participants received an indicative conditional and half a subjunctive conditional, and their task was to generate a conclusion. The results confirmed the predictions of the model theory, namely, the participants made more modus tollens inferences from the subjunctive conditional than from the indicative one. Moreover, the number of MT inferences in the subjunctive condition did not differ significantly from the number of MP inferences (see also Thompson and Byrne, 2002).

Switching between conditional inferences

As introduced in the *Attention-switching* section of the introductory chapter, switching costs are usually observed when switching attention between different operations, as shown by the “task-switching” paradigm (e. g., Allport, 1994; Gonzalez et al., 2005; Rogers and Monsell, 1995; Tornay and Milan, 2001; see Monsell, 2003, for a review). There is some agreement in that the switch cost reflects a tendency to keep active just one task-set at a given moment in working memory (e.g. Allport and Wylie, 1999; Mayr and Kliegl, 2000; Rubinstein, Meyer and Evans, 2001; Tornay and Milán, 2001). We

believe that this paradigm might provide a very straightforward test of the hypotheses of the mental model theory of reasoning (e.g., Johnson-Laird, 1983; Johnson-Laird and Byrne, 1991, 2002) about the representations and processes involved in the correct performance of conditional reasoning problems. For example, the mental model theory hypothesizes these representations and processes are different for indicative MP and MT inferences, but not for counterfactual MP and MT inferences, therefore, a different pattern of switching costs for both situations ought to be expected, as we will now explain.

A conditional in the indicative mood requires a single possibility to be represented in order to solve MP inferences, whereas a fleshing out of at least a second possibility is needed to solve MT inferences.

INDICATIVE

MP		MT	
a	b	a	b
...		not a	not b
		...	

On the other hand, when the conditional is in the subjunctive mood, the same two possibilities are required to solve modus ponens and modus tollens inferences.

COUNTERFACTUAL

MP		MT	
a	b	a	b
not a	not b	not a	not b
...		...	

Thus, the mental model theory presents us with four thoroughly detailed tasks which differ not only in the number of possibilities represented but also in the processes involved in their successful performance (as only in indicative MT inferences is the fleshing out required).

As already mentioned, an undetermined issue in the task-switching field is what makes two tasks different. Although the main aim of this study is to employ the task-switching paradigm to obtain evidence about the nature of the mental representation of conditionals, it is worth noting that the hypotheses of the mental model theory can be of help for task-switching research. Indeed, the mental model theory makes quite specific predictions about the mental representations and processes needed in order to solve conditional inferences, while an issue that remains elusive in the task-switching field is precisely what makes two tasks different. This question has been mainly approached in two different ways. The first approach, let us refer to it as the *task-similarity* approach, consists in defining a *task space* in which two tasks are different when task components are different (i.e. encoding, judgment, response selection, etc.). From this approach results have been reported that show reduced switching costs when similarity in task components is enhanced (see e.g. Arrington, Altmann & Carr, 2000; Kleinsorge, 1999). On this account, smaller switching costs would be predicted for shifts between MP and MT inferences from counterfactual conditionals, as compared to MP and MT inferences from indicative conditionals, as the operations and representations involved in the former are identical.

The second approach derives from the observation of a pattern usually present when switching between tasks of different dominance or *task-strength*. As introduced in the previous chapter, the concept of dominance makes reference to a relation between two tasks, such that one of them is more practiced and more automatic, for example, in

the famous Stroop paradigm (Stroop, 1935) the stronger task is to read the word, as compared to naming its colour. When switching between tasks that differ in dominance a “task-switching asymmetry” is observed (e. g., Allport & Wylie, 2000; Yeung and Monsell, 2003). This pattern reflects the somewhat counterintuitive fact that people are faster when switching into the weaker (or more difficult) task than into the stronger (easier) task. In the Stroop example, this pattern would be reflected in a larger cost when reading the letter after naming the colour, compared to the cost observed when naming the colour after reading the letter. In many cases the cost of switching into the weaker task is not significant, that is, switching costs are only observed for the stronger task. The prediction from this approach is that an asymmetrical inference switching-cost should be observed, with the cost being significant only for the stronger (more dominant) task, that is, MP inferences. It is worth noting that the task-similarity and task-strength approaches are not necessarily exclusive and, to our knowledge, have never been pitted against each other.

In summary, the task-switching paradigm can provide further evidence about the nature of the mental representation of conditional inferences, and the mental model theory can provide task-switching research with a detailed set of tasks, which could help to identify the factors that contribute to make two tasks different. Experiment 1 will test whether an inference switching cost is observed when repeatedly testing the same indicative conditional for MP and MT inferences. This first experiment will provide a baseline effect against which to test the mental model theory’s hypotheses about the representation of counterfactual conditionals in subsequent experiments. Both the task space and task strength approaches to task similarity would predict a cost for switching between MP and MT inferences, but only the task strength approach would predict an asymmetrical cost, that is, significant costs only for MP inferences (which, we claim, is

the stronger task). In other words, the asymmetry prediction is that switching from MT to MP should entail a higher cost than switching in the opposite direction. In experiment 2 two groups of participants are compared, one is presented with indicative and the other with counterfactual conditionals. Given that no fleshing out is needed in counterfactual MT inferences, the prediction in terms of task-strength would be a symmetrical cost, that is, switching in either direction should take the same amount of time. The predictions of the task-similarity would be either lower or absent costs for the counterfactual condition. Experiment 3, provides a thorough test of the effect of mood by means of a within-participants manipulation of this variable. In this case, the task-strength approach would also predict an asymmetrical cost of switching mood, with significant costs only for counterfactual inferences, given that no fleshing out is needed for these.

Experiment 1

The main aim of this study was to employ the task switching paradigm to test predictions of the mental model theory regarding the mental representations involved in making modus ponens (MP) and modus tollens (MT) inferences. The first hypothesis was that switching between MP and MT entails a task switching cost. Should this effect be observed with an indicative conditional, it would provide us with a baseline against which to test the possible effects of the mood of the conditional.

All of the inferences in this experiment were made from the same conditional sentence: "If there is a vowel letter, then there is an even number". An example of the MP inference is: If there is a vowel letter then there is an even number; A; 4". An example of the MT inference is: "If there is a vowel letter then there is an even number; 7; B".

Participants were reminded of this conditional at the beginning of each block of trials. Then, in each trial they were shown a premise and a conclusion (i.e. a specific letter and number), and their task was to evaluate if the inference was valid.

Method

Participants

Thirteen undergraduate psychology students (3 male, 10 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 20 and 29 years (mean 22.3 years). Their vision was normal or corrected to normal.

Materials

The materials consisted of all of the alphabet letters and the digits 1 to 9. Participants were instructed to evaluate the validity of each premise-conclusion pair, in light of the conditional sentence “if there is a vowel letter there is an even number”, which was presented at the beginning of each block of trials. The premise was always shown on the left side of the screen, and the conclusion on the right side, participants were prompted to interpret the inference always in this “direction”, by presenting first the premise, and after a short delay, the conclusion (see procedure). The participant had to decide if the inference was valid. For example, if in a given trial the letter “a” and the number “2” were shown, this conclusion was a correct MP inference, and the response was “valid”. On the other hand, if they were shown the stimulus pair “7” and “e”, this was an incorrect MT inference, and the response was “invalid” (see instructions in Appendix 1). The response was made by pressing one of the two responses key (keys “m” and “z”), one key for “valid” and the other for “invalid”, labelled as such. The mapping of

Chapter 2 An attention switching study of indicative and counterfactual conditionals

keys to responses was counterbalanced. Participants performed a total of 360 trials, with 60 trials of each of the six types possible: 2 inferences (MP or MT) x 2 types of conclusion (valid, invalid) x 2 switching conditions (repeat, shift). There were 10 blocks of 36 trials each. The first block was designed in advance as practice and discarded prior to any data analyses. These blocks were separated by means of short rest periods. The experimental session lasted approximately 30 minutes.

Design

A repeated measures design with two independent variables was used. The variables were: *inference*, with two levels (modus ponens, modus tollens); and *inference switching*, with two levels (inference repetition, that is, either ponens followed by ponens, or tollens followed by tollens; and inference shift, either ponens followed by tollens, or tollens followed by ponens).

Procedure

The stimuli were presented on a computer screen controlled by a PC (Pentium III) that also recorded participants' responses. E-prime software (Schneider, Eschman and Zuccolotto, 2002) generated and controlled the stimulus presentation. Every trial began with a fixation point (a cross) in the centre of the screen lasting 200 ms. Immediately after the fixation and on the left side of the screen appeared the premise (a letter in the case of a modus ponens inference, a number in the case of a modus tollens inference). 400 ms. after the premise, the conclusion appeared on the right side of the screen and the pair remained until participants responded. A beep sounded for one second after participants made an incorrect response. One second after the response the next trial began.

Results

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. There was a main effect of the variable Inference, $F(1, 12) = 55.97, p < .000001$, due to responses to modus ponens inferences being about 300 ms. faster than responses to tollens inferences (996 ms. for MP; 1240 ms. for MT). There was no main effect of Inference Switching, and this result was due to the interaction with Inference, $F(1, 12) = 5.23, p < .041$. The interaction stemmed from an inference switching cost that was significant only for modus ponens inferences, $F(1, 12) = 64.17, p < .000004$ (see figure 2.1). Modus ponens inferences were 85 ms. faster on "repeat" trials, that is, when they followed another modus ponens inference, compared to modus ponens inferences on "shift" trials, that is, when they followed a modus tollens inference (908 ms. and 972 ms.). Modus tollens inferences took the same amount of time regardless of whether they were "repeat" or "shift" trials (1231 ms. and 1248 ms.; repeat trials were only 17 ms. faster). The results thus show a cost asymmetry pattern which reflects the observation that when switching between tasks of differing dominance, switching costs are observed only for the stronger task, in this case, modus ponens inferences. Indeed, the switching cost for MP inferences is about four times larger than the switching cost for MT inferences, with only the former being significant.

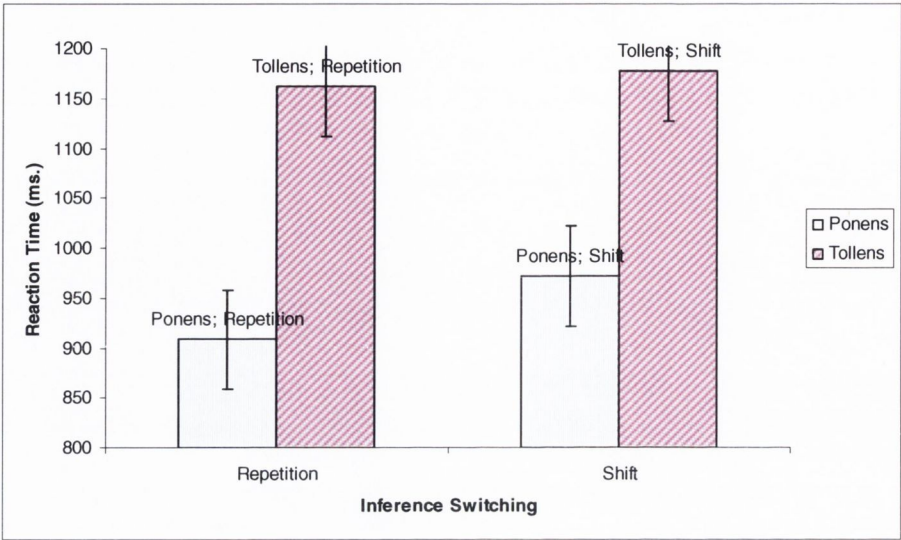


Figure 2.1. Interaction between Inference and Inference Switching in exp. 1 (switching costs are only significant for MP inferences).

Overall accuracy was 94% , 97% for MP inferences and 90% for MT inferences. The difference was significant $F(1, 12) = 59.43, p < .0023$. No other effect or interaction was significant in the accuracy data.

Discussion

Switching between MP and MT inferences displays the same asymmetrical switching cost pattern that is typically observed when switching between tasks of different dominance. That is, switching costs are observed only for MP inferences, indicating that it is harder to switch into stronger task. Given this pattern of switching costs, the difference between MP and MT inferences may reflect task dominance, that is, MP is a more practiced and less resource-demanding task than MT. It is worth noting that, insofar as indicative and counterfactual MP and MT inferences are clearly a different task, the task-similarity approach would also predict an inference switching cost.

However, it is not evident how this approach would accommodate the fact that the cost is not significant for MT inferences.

These results constitute further evidence for the well established ease of the MP inference by means of RT measures that are seldom employed in deductive reasoning tasks. The results are consistent with the mental model theory's proposal that MP requires a single model whereas MT requires multiple models to be fleshed out. However, it should be stressed that these results do not allow us to distinguish between model and rule accounts of conditional reasoning. Rule accounts of deductive reasoning (e.g. Braine & O'Brien, 1998; Rips, 1994) propose that people solve inferences by applying syntactic rules, and that the advantage of MP over MT inferences derives from the greater number of inferential steps needed in order to solve an MT inference. The asymmetrical pattern of inference switching costs is also consistent with a greater number of inferential steps needed to correctly solve an MT rather than an MP inference (e.g., Johnson-Laird et al., 1992, 1994). Nonetheless, this experiment has established a baseline effect against which to test the hypotheses of the model theory regarding the representation of counterfactual conditionals, and for which the rule theory offers no explanation. Testing these hypotheses is the objective of the next two experiments.

It could be argued that testing the same conditional throughout the whole experiment eliminates the need to reason at all. Conversely, it could be that the use of implicit negations (i.e., "7" instead of "not even") made the task too complicated (see Evans, 1989). Therefore, in the next experiment the setup is modified in order to rule out the possible influence of these factors. A different conditional is presented on each trial, and explicit rather than implicit negations are used.

Experiment 2

The aim of this experiment is to compare the mental representation of indicative and counterfactual conditionals using the task switching methodology. The first experiment has shown that switching between MP and MT inferences from an indicative conditional showed the same features as switching between tasks of differing dominance. However, insofar as MT inferences from a counterfactual conditional are known to be easier (e.g., Thompson and Byrne, 2002), the switching cost pattern should be affected.

In their study on asymmetrical switching costs, Meuter and Allport (1999) proposed that the more similar the two tasks are in terms of dominance, the more symmetrical the switching costs should be (that is, switching costs should be similar and significant for both tasks). From this perspective, a more symmetrical switching cost should be observed for counterfactual than indicative inferences, as MP and MT are more similar in the former than in the latter mood. In other words, the RT to inference switch and inference repeat trials should not differ for counterfactual conditionals.

Two predictions in terms of task-similarity can be put forward. First, it could be argued that the task-similarity approach should not predict inference switching costs for counterfactual conditionals, as the mental representations of MP and MT inferences are the same, and no fleshing out is required in either case. However, results from exp. 4 in Byrne and Tasso's study (1999) suggested that participants focused on just one of the two possibilities represented for counterfactual conditionals. No clear tendency to focus on either possibility was observed though, half of the participants seemed to be focusing on the alternative where the conjecture happened, whereas the other half focused on the alternative in which the conjecture did not happen. In this case, a cost for switching between counterfactual MP and MT inferences would be predicted, as the possibility

one needs to focus on to solve either inference is different. It is worth noting this latter prediction is not incompatible with the task-strength prediction, that is, asymmetrical costs only for indicative conditionals.

Method

Participants

Twenty-four undergraduate psychology students (10 male, 14 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 22 and 45 years (mean 23.6 years). Their vision was normal or corrected to normal. Participants were randomly assigned to two groups (n=12).

Materials

The materials consisted of indicative and counterfactual abstract conditional sentences, which made reference to one of four possible materials: shapes, numbers, letters and coloured patches. Once again there were two types of inferences (MP, MT) x 2 types of conclusion possible (valid, invalid). The materials were presented on the centre of the screen with the conditional, premise and conclusion on separate lines and in downwards direction, e.g.:

If there is not an A there is a B

not-B

therefore

A

The response to the above example is “valid”, as it represents a valid MT inference (see instructions in Appendix 1). The responses were equal to those used in experiment 1. Participants performed a total of 120 trials, with 20 trials of each of the

Chapter 2 An attention switching study of indicative and counterfactual conditionals

six types possible: 2 inferences (MP or MT) x 2 types of conclusion (valid, invalid) x 2 switching conditions (repeat, shift). There were 6 blocks of 20 trials each. The first block was designed in advance as practice and discarded prior to any data analyses. These blocks were separated by means of short rest periods. The experimental session lasted approximately 30 minutes.

Design

We used a mixed design with three independent variables. The first two were manipulated within-participants: *Inference*, with two levels (MP, MT); and *Inference Switching*, with two levels (inference repetition, that is, either ponens followed by ponens, or tollens followed by tollens; and inference shift, either ponens followed by tollens, or tollens followed by ponens). The third variable was *Mood* (indicative, counterfactual) and was manipulated between-participants. The indicative group thus performed the same task as did participants in experiment 1, although with the mentioned differences regarding materials.

Procedure

Every trial began with a conditional at the top of the screen, immediately after which the premise and the conclusion appeared one beneath the other. The three sentences were presented sequentially, with the presentation of all of them lasting 2 seconds. All three sentences remained on the screen until a response was made, after which feedback on that trial's execution was shown on screen for half a second, and then the next trial began.

Results and Discussion

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. There was a main effect of the variable Inference, $F(1, 22) = 11.63$, $p < .002$, due to responses to modus ponens inferences being 126 ms. faster than responses to tollens inferences (1603 ms. for MP; 1729 ms. for MT). There was no main effect of Inference Switching, and this result was due to the interaction with Inference, $F(1, 22) = 8.37$, $p < .008$. As in experiment 1, the interaction stemmed from an inference switching cost that was significant only for MP inferences $F(1, 22) = 3.95$, $p < .049$. MP inferences were 120 ms. faster on "repeat" trials, that is, when they followed another MP inference, compared to MP inferences on "shift" trials, that is, when they followed a MT inference (1532 ms. for MP; 1652 ms. for MT inferences). This time, for MT inferences a benefit instead of a cost was observed, although it did not reach significance, $F < 1$. MT inferences were 110 ms. slower when they followed another MT inference, compared to when they followed an MP inference (1795 ms. and 1683 ms). The results thus showed a cost asymmetry pattern which, as in experiment 1, reflected the observation that switching costs are observed only for the stronger task, in this case, MP inferences (see figure 2.2).

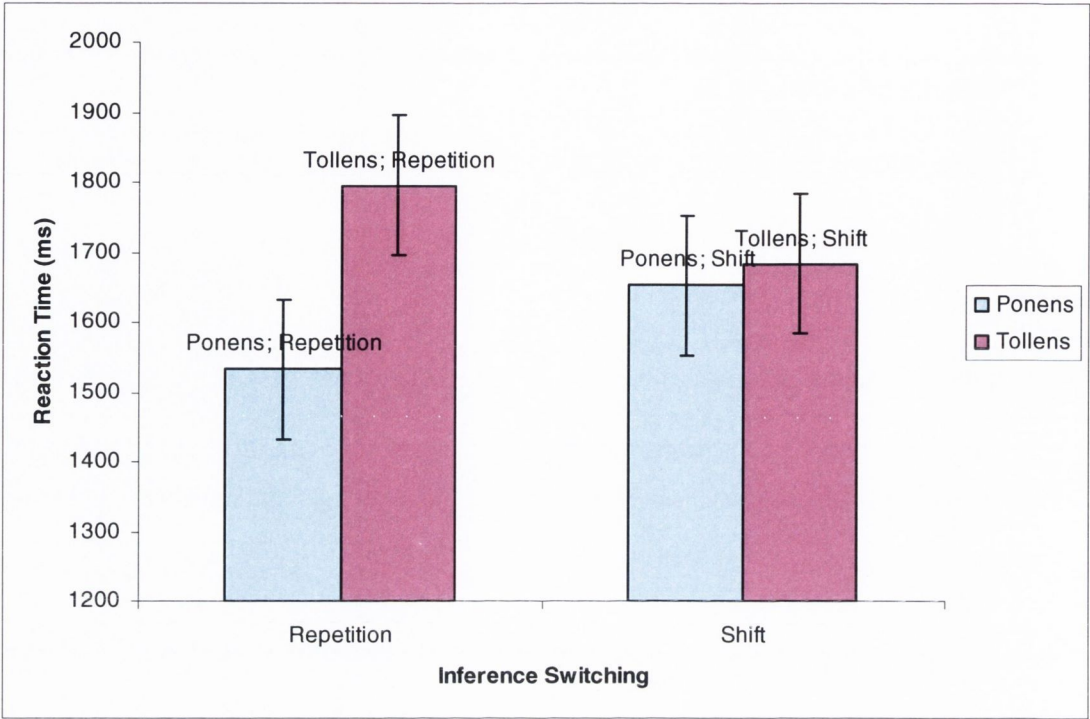


Figure 2.2. Interaction between Inference and Inference Switching in exp. 2 (switching costs are only significant for MP inferences, regardless of mood).

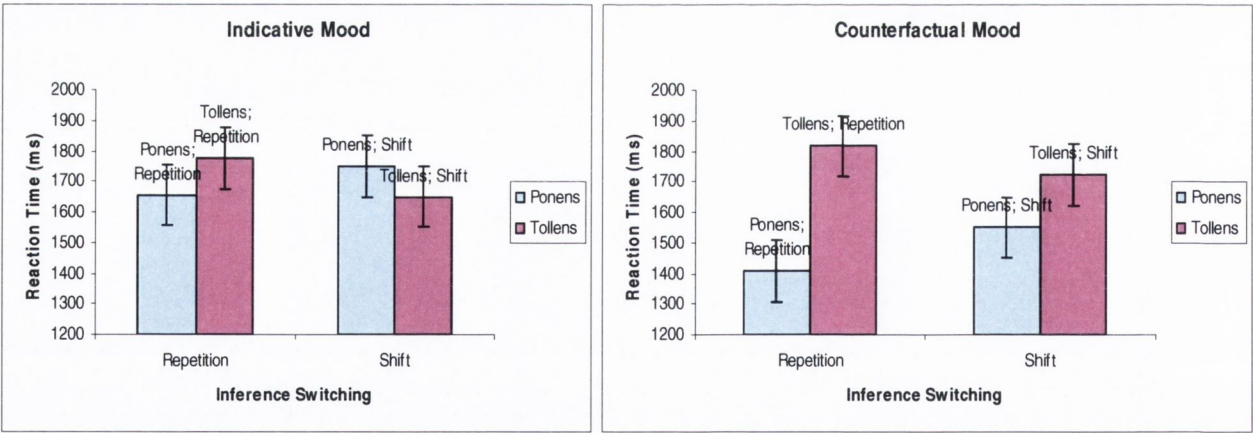


Figure 2.3. Interaction between Inference and Inference Switching for indicative and counterfactual conditionals in exp. 2.

Experiment 2 replicated the asymmetrical pattern for switches between MP and MT inferences, in other words, switching costs were only observed for MP inferences.

The absence of interactions involving the Mood variable implies that inference switching costs are also present for counterfactual conditionals, as figure 2.3 shows.

However, it is worth noting that only conditional sentences were presented throughout the whole experiment, which might have eliminated the need for participants to carefully read the sentences, therefore making it difficult to convey counterfactuality. The next experiment was designed to avoid this possible artifact.

Experiment 3

The aim of this experiment was to provide a more sensitive test of the possible differences in MP and MT switching-cost patterns between indicative and counterfactual conditionals by means of a within-participants manipulation of the Mood variable, and by making sure that participants carefully read the sentences. As introduced in experiment 2, depending on whether participants focus on one possibility or both for counterfactual conditionals, a task-similarity approach would predict inference switching costs or an absence of them for this condition.

However, the manipulation of the Mood variable also allowed us to introduce a second kind of switching, that is, switching between counterfactual and indicative conditionals. This manipulation allows to test whether there are costs of switching between mood without switching between inferences. It could thus be argued that the task-strength approach would predict an asymmetrical cost of switching mood. Only indicative conditionals require the fleshing out process, which could result in inferences from indicative conditionals being a weaker task than inferences from counterfactual conditionals. In this case, mood-switching costs should only be significant for counterfactual inferences, that is, a shift from an indicative to a counterfactual conditional should entail a higher cost than to shift in the opposite direction, consistent

Chapter 2 An attention switching study of indicative and counterfactual conditionals with counterfactual inferences being the stronger task. Furthermore, given that this fleshing out process is only required for indicative MT inferences, an interaction between Inference, Mood and Mood Switching should be observed as well. In other words, mood switching costs should be observed for the one inference, MT indicative, that is clearly weaker than the rest.

Method

Participants

Fourteen undergraduate psychology students (6 male, 8 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 21 and 32 years (mean 23.6 years). Their vision was normal or corrected to normal.

Materials and Procedure

The materials and procedure used in this experiment were the same as in experiment 2 (see instructions in Appendix 1). In this experiment we also included a small number of “fillers” in which disjunctions rather conditionals were evaluated, to ensure that participants read the sentences carefully (i.e. “there was an A or there was a B; there was an A; therefore there was not a B). Disjunction trials and every trial immediately following a disjunction were discarded from the statistical analyses.

Design

We used a repeated measures design with four independent variables: *Inference*, with two levels (MP, MT); *Inference Switching*, with two levels (inference repetition, that is, either ponens followed by ponens, or tollens followed by tollens; and inference shift,

either ponens followed by tollens, or tollens followed by ponens); *Mood* (indicative, counterfactual); and finally, *Mood Switching* (mood repetition, that is, either counterfactual followed by counterfactual or indicative followed by indicative; and mood shift, that is, counterfactual followed by indicative or indicative followed by counterfactual).

Results

The introduction of filler trials seemed to have served its purpose, as RT's for this experiment were much larger than those from the previous experiment. Indeed, when reading times were subtracted, RT's were similar to those of experiment 2, which further suggests that participants in that experiment were not reading the conditional sentences carefully, thus making it impossible to convey counterfactuality. Therefore, RTs in this experiment did not include reading times, rather, we focused on the times employed in considering the premise and conclusion.

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. There was a marginal main effect of the variable Inference, $F(1, 13) = 3.78, p < .073$, due to responses to MP inferences being about 301 ms. faster than responses to MT inferences (1864 ms. for MP; 2165 ms. for MT). There was no main effect of Inference Switching, and this result was due to the interaction with Inference, $F(1, 13) = 6.08, p < .020$. As in experiments 1 and 2, the interaction stemmed from an inference switching cost that was significant only for MP inferences $F(1, 13) = 4.70, p < .049$. MP inferences were 126 ms. faster on "repeat" trials, that is, when they followed another MP inference, compared to MP inferences on "shift" trials, that is, when they followed a MT inference (1798 ms. for repeats; 1924 ms. for shifts). A benefit instead of a cost was observed for

MT inferences, although as in experiment 2, even though it did not reach significance, $F < 1$. Nonetheless, the magnitude of this benefit is similar to that of the switching cost for MP inferences. MT inferences were about 136 ms. slower when they followed another MT inference, compared to when they followed a MP inference (2231 ms. and 2096 ms). The results thus replicate for the third time a cost asymmetry pattern which reflects the observation that when switching between tasks of differing difficulty, switching costs are observed only for the easier task, in this case, MP inferences (see figure 2.4).

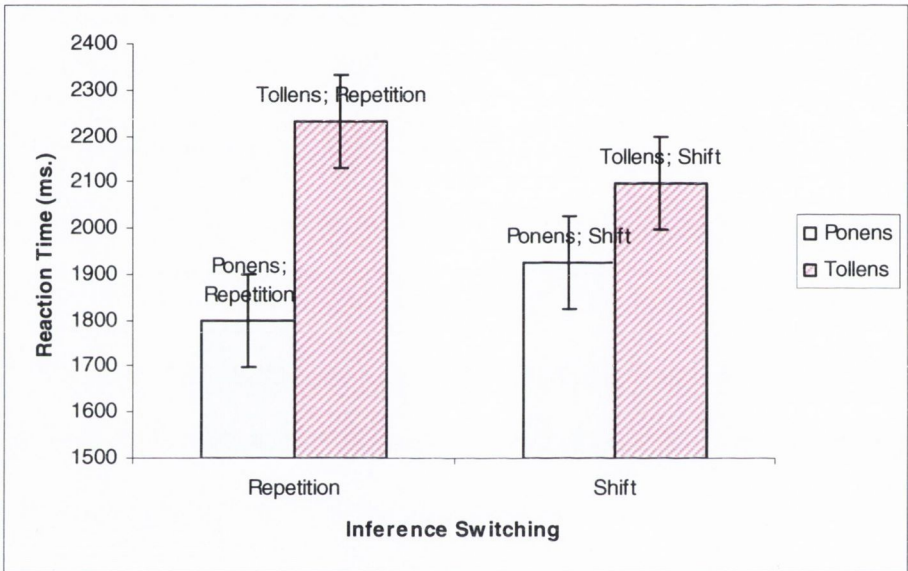


Figure 2.4. Interaction between Inference and Inference Switching in exp. 3 (switching costs are only significant for MP inferences, regardless of mood).

As in experiment 2, the predicted interaction between Mood, Inference and Inference switching was not significant (see figure 2.5). However, there was a marginal main effect of Mood Switching, $F(1, 13) = 3.86, p < .073$, due to a mood switching cost of 294 ms. (mood repetition: 1849 ms.; mood shift: 2155 ms.). The interaction between Mood and Mood Switching was significant, $F(1, 13) = 6.20, p < .027$, there was a significant 575 ms. mood switching cost for the counterfactual condition (1769 ms. for

repetitions; 2324 ms. for shifts), $F(1, 13) = 11.78, p < .004$, but not for the factual condition (1929 ms. for repetitions; 1993 ms. for shifts), $F < 1$ (see figure 2.6). This result reflects an asymmetrical switching cost pattern, that is, *mood-switching* costs are only observed for counterfactual conditionals, thus suggesting that MP and MT inferences are a stronger task when made from a counterfactual conditional rather than an indicative conditional.

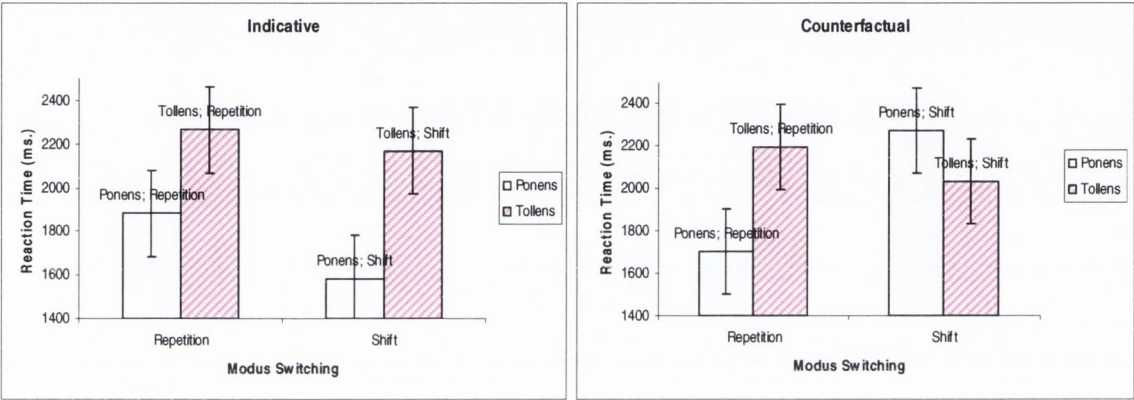


Figure 2.5. Interaction between Inference and Inference Switching for indicative and counterfactual conditionals in exp. 3.

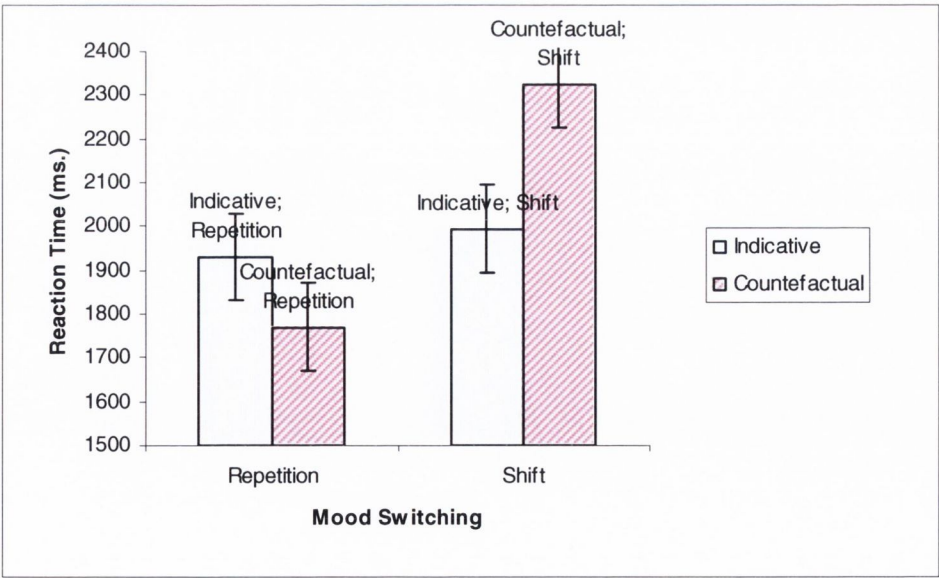


Figure 2.6. Interaction between Mood and Mood Switching, reflecting an asymmetrical pattern for the mood switching cost, exp. 3.

A marginally significant interaction between the variables Mood and Inference was observed as well, $F(1, 13) = 3.52, p < .083$, the effect of inference was significant only for indicative conditionals, $F(1, 13) = 4.70, p < .041$ (for counterfactual conditionals, $F < 1$). This replicates the finding that MT is harder than MP for indicative conditionals but not for counterfactual conditionals (Byrne and Tasso, 1999; Thompson and Byrne, 2002).

Finally, a two-way interaction between Inference, Mood and Mood Switching was observed, $F(1, 13) = 14.26, p < .002$. This interaction stemmed from the fact that a costs are not significant for MT trials in the counterfactual condition, $F < 1$ (see figure 2.7). In other words, the pattern further supports the proposal that indicative MT inferences are the weaker of the four tasks in this study, consistent with these being the only ones requiring the fleshing out process. Furthermore, mood switching costs are significant for the other three conditions, including counterfactual MP and MT.

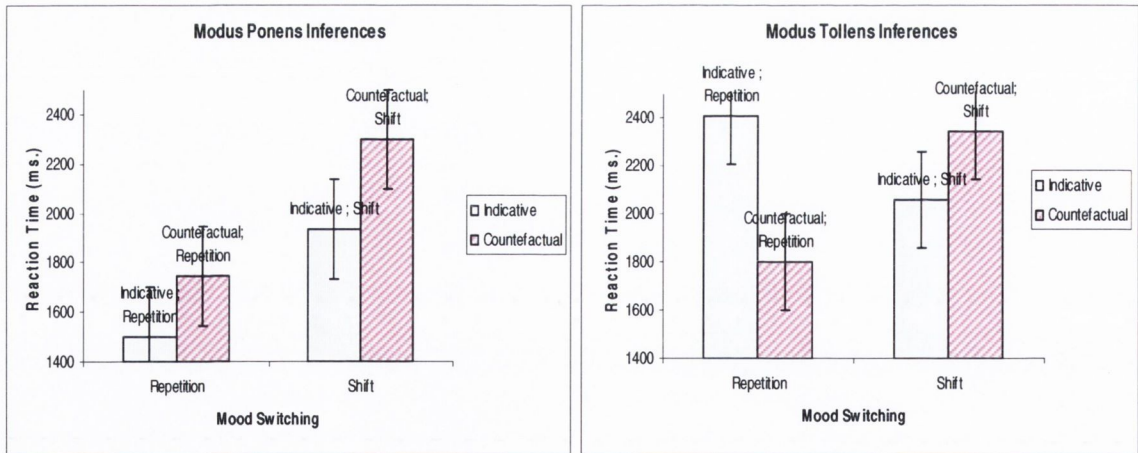


Figure 2.7. Interaction between Mood and Mood Switching for MP and MT trials.

Discussion

Experiment 3 replicated the asymmetrical pattern for switches between MP and MT inferences, significant inferences switching costs are significant for MT to MP shifts, but not for shifts in the opposite direction. The interaction between Mood and Inference is evidence in favour of the mental model theory's hypotheses about the representation of MP and MT inferences from factual and counterfactual conditionals. MT is harder than MP for indicative conditionals but not for counterfactual conditionals (Byrne and Tasso, 1999; Thompson and Byrne, 2002). An asymmetrical pattern of mood switching costs was unveiled as well, in this case switching costs are significant only for counterfactual trials, suggesting that carrying out an inference from a counterfactual conditional is a stronger or easier task than its counterpart, and consistent with the view that only in the indicative mood is the fleshing out process required. Finally, this interaction between Mood and Mood Switching is modulated by Inference, this result will be further analyzed below.

General discussion

The first main finding of this study is that an asymmetrical pattern of inference switching costs is observed in all three experiments. In every case, switching costs are observed only for MP inferences, that is, there is a cost for a MT to MP shift, but not for the opposite. The interpretation of these results in light of the mental models (e.g., Johnson-Laird & Byrne, 2002) and task-switching (Allport et al., 1994; Allport and Wylie, 1999, 2000) literatures is straightforward: compared to indicative MT inferences, indicative MP inferences are the stronger, more practiced and less demanding task, as they do not require the fleshing out of all of the possibilities consistent with the conditional.

The second main finding is the asymmetrical pattern of mood switching costs observed in experiment 3, reflecting the fact that switching costs are significant only for counterfactual trials, which suggests that carrying out an inference from a counterfactual conditional is a stronger or easier task than its counterpart, again consistent with the view that only in the indicative mood is the fleshing out process required. The third main finding is that this interaction between Mood and Mood Switching (i.e. the asymmetrical pattern of mood switching costs) is modulated by Inference. Mood switching costs are asymmetrical for MT inferences but not for MP inferences. This asymmetry reflects the non significance of the cost for indicative MT inferences, suggesting that these are a weaker task than counterfactual MT inferences. Again, this is consistent with the view that only for the indicative MT inference is the fleshing out process required. Finally, we also replicated the observation that the advantage of MP over MT inferences is not significant when the conditional is in subjunctive mood (see Byrne and Tasso, 1999).

The predictions put forward in this chapter regarding different patterns of inference switching costs depending on mood are not supported by the results from experiment 3. Rather, we observe different patterns of mood switching costs depending on which inference is being carried. Nonetheless, these mood switching patterns provide useful insights into the processes and representations involved in conditional reasoning. The asymmetrical pattern of mood switching for MT inferences is consistent with counterfactual MT inferences being a stronger task than indicative MT inferences, which again, we interpret to reflect the need of the fleshing out process for the latter inferences. This result suggests in turn that no fleshing out is needed for counterfactual MT inferences. The mental model theory account on counterfactual conditionals was criticised for proposing that people represent two possibilities (Evans and Over, 2004),

however, the observation that a fleshing out process is not necessary for counterfactual MT inferences implies that the relevant possibilities are already represented. Furthermore, it is interesting to note how the cost of switching between indicative and counterfactual MP inferences is symmetrical. we further elaborate on these issues in the final chapter.

With regard to attention-switching, it is interesting to note that switching costs were similar in all three experiments regardless of the very different time intervals between switches, which amounted to various seconds in experiment 3 as compared to about 3 seconds in experiments 1 and 2 (as only in experiment 3 were participants reading the sentences carefully). This result replicates the well established robustness of the switching cost (e.g. Rogers and Monsell, 1995; but see Tornay and Milán, 2001). On the issue of what makes two tasks different, our results are better interpreted in terms of task-strength as opposed to task-similarity. However, it is possible that the manipulations of this experiment are not optimal to observe task similarity effects. In principle, the mental model theory is not concerned in principle with issues such as attention-switching and asymmetrical costs, it is therefore interesting to see how such phenomena seem to behave in accordance with the predictions of this theory. This issue will also be further discussed in the final chapter.

In summary, the results of the three experiments provide convergent evidence for the mental model theory's hypotheses about the representations involved in carrying out conditional inferences from indicative and counterfactual conditionals. This evidence is convergent in that it is obtained by means of a paradigm (task-switching) and a measure (RT) that have not often been employed in deductive reasoning tasks. A novel contribution of this study is that when considering what makes inferences from counterfactual and indicative conditionals different, the emphasis is on the need to

Chapter 2 An attention switching study of indicative and counterfactual conditionals

engage in the fleshing out process. This process is only needed when carrying out MT inferences from indicative conditionals and, indeed, this is the only condition that differs from the rest. Even though there is wide agreement that the limiting factor in reasoning performance is the limitation in WM capacity (e.g., Braine and O'Brien, 1991; Johnson-Laird and Byrne, 1991; Rips, 1994), little effort has been directed at exploring the role played by low-level attentional processes (e.g. attention-switching), aimed to deal with these capacity limitations. This study constitutes a firm step in that direction.

This chapter has thus highlighted the importance of the principle of parsimony by observing the effects of the fleshing out process in an attention-switching paradigm. In the next chapter we turn our attention to another main tenet of the mental model theory, the principle of truth.

Chapter 3 - Attentional limitations and true possibilities

Introduction

As mentioned in the introductory chapter, capacity limitations in cognitive processing are a well established phenomenon in cognitive psychology. The mental model theory reflects this phenomenon in its principle of parsimony. Other theories of reasoning would agree with this hypothesis (e.g. Braine and O'Brien, 1991), however, only the mental model theory proposes that people tend to think about what is true, as stated in the principle of truth. Indeed, a variety of studies in several areas of cognitive science have provided evidence that people show a widespread tendency to think about what they consider to be true. As introduced in chapter 1, in the literature on hypothesis-testing behaviour and scientific thinking, the nature of an effect first called "confirmation bias" has been a matter of discussion for more than four decades (e.g., Kareev and Halberstadt, 1993; Tukey, 1986; Tweney, et al., 1980; Wason, 1960). This paradigm has provided extensive evidence that when people engage in hypothesis-testing tasks, they show a marked preference towards confirming rather than falsifying the hypothesis they consider to be true, and they exhibit a natural tendency towards considering positive information. But, why do people exhibit this tendency? Again, one view is based on the mental model theory (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991, 2002), which proposes that, due to limitations in processing capacity they tend to only represent information they consider to be true (the principle of truth) and to represent explicitly as little information as possible (principle of parsimony).

Research on attentional capacity and processes (e.g., attention switching paradigms) has provided extensive evidence that people can attend to limited amounts

of information and that only one object or operation can be held in the focus of attention at a given moment (e.g., Cowan, 1995; Garavan, 1998; Mayr and Kliegl, 2000; Oberauer, 2002; Tornay and Milán, 2001). However, the mental model theory suggests that the tendency to focus on true possibilities is also a key explanatory factor of reasoning performance. Thus, what the mental model theory describes as the principles of parsimony and true possibilities (Johnson-Laird and Byrne, 2002), may both reflect fundamental features of our cognitive system, which we intend to demonstrate by applying an attention switching methodology to a very simple rule testing paradigm.

The main proposal of the series of experiments reported in this chapter is thus that the tendency to think about true possibilities may reflect a basic cognitive phenomenon. Consequently, it should be reflected on the latency of responses in a simple categorization task, not only in “higher-level” processes such as deductive reasoning or hypothesis-testing.

Hypotheses and overview of the experiments

All four experiments reported rely on a simplified version of Wason’s 2-4-6 task, in which participants have to determine if triples of numbers either confirm or disconfirm a previously specified rule. In this novel task we constructed, participants do not have to generate triples, nor do they have to discover the rule that governs the categorisation of the triples. Instead, the task requires categorization of confirmatory and disconfirmatory instances of a known rule. That is, participants are informed of a rule that describes the behavior of triples of numbers (i.e. ascending numbers) and shown a different triple on each trial (i.e. 2-4-6). Their task is to confirm or reject the triple as an instance of the rule (i.e. the triple 2-4-6 would confirm the rule ascending numbers). This task will

allow us to test whether participants are faster to confirm as compared to reject, and to test a series of predictions regarding switching costs that we turn to describe.

Our first prediction is thus that, due to the tendency to think about true possibilities, participants will be faster to identify triples that confirm a rule. The second prediction is that reaction time switching costs should be observed when switching between confirmation and rejection. Furthermore, we propose that confirming triples is a stronger and less resource demanding operation, and therefore that an asymmetrical switching-cost pattern should be observed. Specifically, there should be a higher cost when switching into the stronger operation which our first prediction anticipates will be confirmations.

In subsequent experiments we aim to test whether the asymmetrical response mode switching costs observed in experiment 4 are indeed the reflection of a task-switching process that is operating upon two tasks that differ in dominance. Based on the observation that a common neural basis subserves the process of switching between tasks and objects (Wager et al., 2004), it is proposed that if participants have to switch both between rules (objects) and response modes (tasks) and both are operated by a common mechanism, then a correlation and an underadditive interaction should be observed between object (rule) and task (response mode) switching costs. To our knowledge, object and task switching (see section *Attention-switching* of the introductory chapter) have not been observed in the same task before, therefore, this manipulation is also of interest for attention-switching research. The aim of the experiment 5 was therefore to establish if a rule switching cost is observed and to address two competing views on the effect of the dual goal manipulation of the 2-4-6 task (see *Hypothesis-testing* section in the introductory chapter). Experiment 6 sought to test whether response mode and rule switching costs were operated by a common

mechanism. Finally in experiment 7, we established some boundary conditions for the observed switching costs, specifically, we addressed the role of the interference between rules. These results were relevant with respect to some hypotheses on the effect of the dual-goal manipulation.

Experiment 4

Participants had to determine if each of a series of number triples (e.g., 2, 4, 6) either confirmed or rejected a previously specified rule. We hypothesized that a tendency to focus on true states of affairs should result in confirming being faster than rejecting. We also argued that if confirming and rejecting are different operations with the former being more dominant than the latter, an asymmetrical pattern of switching costs should be observed.

Method

Participants

Twenty undergraduate psychology students (9 male, 11 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 21 and 30 years (mean 23.4 years). Their vision was normal or corrected to normal.

Materials

The materials consisted of triples of numbers separated by hyphens (e.g. 2-4-6), which were presented on the centre of the screen, in a black font (Courier New, 18 mm). The triples were composed of one and two digit numbers, in either ascending or descending order, with the constraint that for every triple, the numerical difference between the first

and second numbers was equal to the difference between the second and the third. In other words, numbers within the triple varied in constant amounts, which ranged from a minimum of 1 to a maximum of ten (that is, from 0-1-2 to 79-89-99). Any triple within these constraints could be presented on a given trial.

Participants were told that they would see a series of number triples. The participant's task was to decide, for every triple, if it conformed to the rule "ascending numbers". This was communicated by pressing one of the two responses key (keys "m" and "z"). The mapping of keys to responses was counterbalanced across participants. There were 360 trials, with 90 trials of each of the 4 types possible (i.e., 2 responses (confirm, reject) x 2 types of triple (ascending, descending). Participants performed 10 blocks of 36 trials each. The first block was designed in advance as practice and discarded prior to any data analyses. These blocks were separated by means of short rest periods.

Design

The design was a two-way repeated measures with the factors Response Mode (with two levels: confirm and reject), and Response Mode Switching (with two levels: response mode repetition, e.g., confirm to confirm; and response mode shift, e. g., reject to confirm).

Procedure

The stimuli were presented on a computer screen controlled by a PC (Pentium III) that also recorded participants' responses. The E-prime software (Schneider, Eschman and Zuccolotto, 2002) generated and controlled the stimulus presentation. Every trial began with a fixation point (a cross) in the centre of the screen lasting 500 ms. and the number

triple appeared immediately after the fixation and in the same position, and remained until the participant made a response. After the response, participants received feedback (correct or incorrect on the centre of the screen) for 1500 ms. and the next trial began.

Results

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. Mean accuracy was 97%. The accuracy data mirrored the main effects found in the RT data (both for this and the subsequent experiments) and will not be reported.

In the reaction time data, there was a main effect of the variable Response Mode, $F(1, 19) = 11.02, p < .005$. due to participants being 71 ms. faster at confirmation (1171 ms. for confirming trials; 1243 ms. for rejecting trials). There was no main effect for Response Mode Switching, $F < 1$, but there was an interaction between this variable and Response Mode, $F(1, 19) = 16.30, p < .001$ (see figure 3.1). As the planned comparisons reveal, there was a switching cost for confirmations, $F(1, 19) = 9.75, p < .0008$, but not for rejections, for which a switching benefit was observed instead, $F(1, 19) = 13.35, p < .005$.

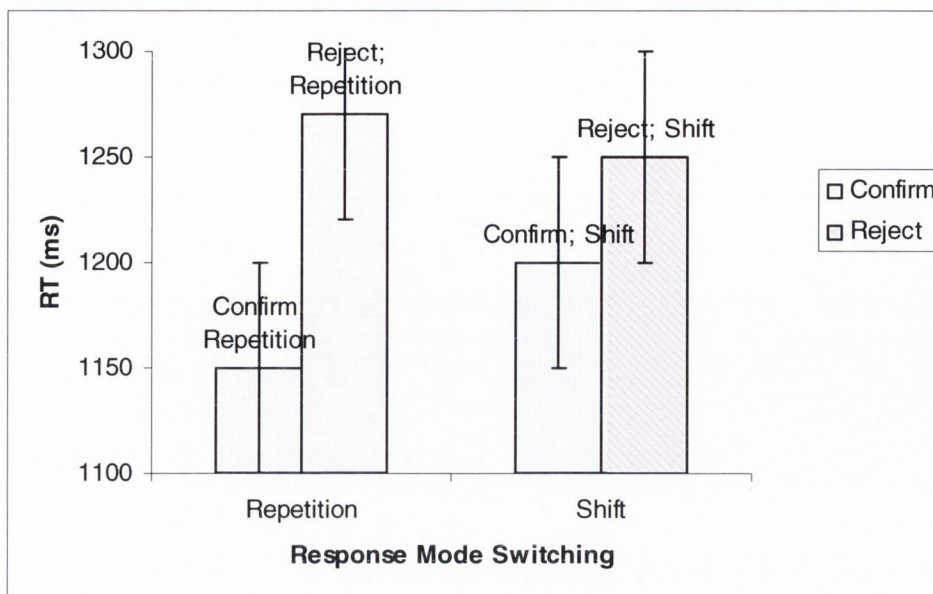


Figure 3.1. Reaction times to confirm and reject triples on repeat and shift trials in exp. 4

Discussion

As predicted, participants were significantly faster when confirming a triple than when rejecting one, a result consistent with the tendency to focus on true possibilities observed in studies of hypothesis testing in the 2-4-6 task. The experiment provides further evidence that participants tend to think about true possibilities, by means of RT measures, and obtained in a task that required participants to categorize instances according to whether they confirmed or rejected an hypothesis, and in which there was no motivational component to defend one's own hypothesis. Importantly, there is a switching cost when participants confirm after a rejection (a switch from reject to confirm), but a switching benefit when they switch in the opposite direction (confirm to reject). As explained in chapter 1, this pattern of switching costs is typically observed when switching between tasks that differ in difficulty or "task-strength" (e. g. Allport

and Wylie, 1999; 2000; Yeung & Monsell, 2003). The result is consistent with the idea that confirming is a less demanding or more automatic task than rejection.

Experiment 5

In the same way experiment 4 established a task-switching like cost for response modes, this second experiment aims to establish whether an object-switching like cost is observed when switching between rules. The two rules were “ascending numbers” and “numbers descending by 3s” (e. g. 9, 6, 3). Note that the participants task remains the same, that is, either to confirm or reject that a triple fits the relevant rule, therefore switches of response mode are unavoidable. However, the number of response mode switches required in this experiment is kept at low numbers, and the response mode switching variable will not be considered in the analyses.

Method

Participants

Twenty undergraduate psychology students (8 male, 12 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 23 and 32 years (mean 25.2). Their vision was normal or corrected to normal.

Materials

The only difference with the previous experiment is that triples are presented in a different font colour, depending on the rule to be tested, thus, half of the participants had to test the “ascending rule” when testing a red coloured triple, and the “descending by 3s” rule when testing a blue coloured triple. For example, the triple “2, 4, 6” in a red

font prompts the participant to test the “ascending rule”, so the response is confirm; if the same triple is presented in a blue font, the participant has to test the “descending by 3s” rule, and the response is reject. This mapping of colours to rules was reversed for the other half of the participants as a control measure. In every trial there was a 50% likelihood for a confirm response to be required. When the ascending rule was being tested, half of the triples presented were in ascending order and the other half were in descending order. Conversely, when testing the “descending by 3s” rule, half of the triples descended by threes, with the other half composed of triples that either ascended or descended by a constant other than three. The participant’s task was to decide, for every triple, if it conformed to the rule cued by its colour, pressing one of the two response keys (one for confirming and another for rejecting, the keys used were “m” and “z”). The mappings of keys to responses was counterbalanced across participants, and triples and rule were presented in random order.

Design

A repeated measures design was used, with three independent variables: rule, with two levels (ascending, descending by 3s); rule switching, two levels (repeat, that is, testing a rule after testing the same rule in the previous trial, e. g., “ascending” to “ascending”; and shift, e. g., “ascending” to “descending by 3s”); and response mode, with two levels (confirm, reject). Long sequences of triples requiring the same response were used in order to minimize the number of response mode switches, although a small number of short sequences was used as well to prevent the participants from noticing the manipulation.

Procedure

The experimental procedure was the same as in the previous experiment, except for the fact that participants were instructed to associate the colour of the triple to the corresponding rule.

Results

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. Mean accuracy was 96%.

In the reaction time data, the main effect of Response Mode is replicated, $F(1, 19) = 10.52$, $p < .0005$, participants are 133 ms. faster when the response mode is confirmation (1702 ms. for confirmations; 1835 ms. for rejections). There was a main effect of the variable Rule $F(1, 19) = 7.27$, $p < .016$, due to participants being 210 ms. faster at testing the ascending rule (1667 ms. for ascending rule trials; 1877 ms. for descending by 3s rule trials: 1877 ms.). A main effect of Rule Switching, $F(1, 19) = 18.78$, $p < .0005$, reflected an average rule switching cost of 156 ms. (1688 ms. for rule repeating trials; 1839 ms. for rule shifting trials). A marginal interaction between Rule and Response Mode, $F(1, 19) = 3.37$, $p < .081$, revealed that the effect of Response Mode was not significant for trials under the descending by 3s rule, $F < 1$, but was significant for the easier ascending rule, $F(1, 19) = 6.37$, $p < .011$. No interactions between Rule and Rule Switching were observed, that is, no asymmetrical pattern is observed for rule switching costs.

Discussion

The advantage for confirming responses observed in experiment 4 is replicated in the present experiment. In addition, the results show a sizeable rule switching cost of about 150 ms. indicating that participants are able to hold just one rule at a time in their focus of attention. This result replicates the object-switching costs found in other studies (e. g., Garavan, 1998; Oberauer, 2002).

Even though the “ascending rule” takes on average 200 ms. less to test than the “descending by 3s rule”, no interaction between Rule and Rule Switching is observed. In other words, there is no asymmetrical pattern in the rule switching cost, consistent with asymmetrical cost patterns being observed for task switches, but not for object switches.

Experiments 4 and 5 have thus established an asymmetrical task switching-like cost for response modes and an object switching-like cost for rules, respectively. Experiment 6 will test the hypothesis that response mode and rule switching costs are operated upon by the same mechanism, which would provide further evidence that switching between confirmation and rejection is equivalent to switching between tasks of differing dominance. In other words, if the response mode switching cost reflects something different from a switching operation, this effect should not correlate to and interact with the rule switching cost.

Experiment 6

In this experiment participants switched both between rules and the confirmation-rejection operations, bringing together task-switching and object-switching. The aim of this experiment was to evaluate to what extent these effects are related. Thus, based on recent neuroimaging evidence suggesting a common neural basis for different types of

switching (Wager, Jonides and Reading, 2004), we hypothesized that a correlation and an underadditive interaction should be observed between object (rule) and response mode (task) switching costs. To our knowledge, object and task switching have not been observed in the same task before.

Method

Participants

Twenty undergraduate psychology students (10 male, 10 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 19 and 27 years (mean = 22.3). Their vision was normal or corrected to normal.

Materials

This experiment used triples produced by participants during a previous hypothesis-testing experiment (Cowley & Byrne, 2000), thereby enabling us to use a more “ecological” stimulus set. Practically every triple in this set was classifiable as fitting the rules “ascending numbers” and “descending numbers” so these rules and only the triples that fit them were used in this experiment. All triples were presented in ascending or descending order. Otherwise, the materials are the same as in the previous experiment.

Design

A repeated measures design was used, with four independent variables: rule, with two levels (ascending, descending); rule switching, two levels (repeat, shift); response mode, with two levels (confirm, reject); and response mode switching (repeat, shift).

Procedure

The experimental procedure was the same as in the previous experiment.

Results

Only correct trials with response times that occurred between 2 standard deviations around each participant's mean were included in the reaction time analyses. Mean accuracy was of 93%.

First, this experiment replicates the result of experiments 4 and 5. In the reaction time data, the main effect of Response Mode found in exps. 4 and 5 was replicated again, $F(1, 19) = 46.95$, $p < .0000005$; participants were 110 ms. faster when the response mode was confirmation (1472 ms. for confirmations; 1582 ms. for rejections). The main effect of Rule found in exp. 5 was replicated again, $F(1, 19) = 36.71$, $p < .00001$, due to participants being 150 ms. faster at testing the ascending rule (average RT for ascending rule trials: 1454 ms.; average RT for descending trials: 1603 ms.). The main effect of Rule Switching found in exp. 5 was replicated, $F(1, 19) = 7.05$, $p < .016$, and reflected an average rule switching cost of 76 ms. (1483 ms. for rule repeating trials; 1559 ms. for rule shifting trials. Once again, there was no main effect of Response Mode Switching, $F < 1$, but as in experiment 4, there was an interaction between Response Mode and Response Mode Switching, $F(1, 19) = 23.79$, $p < .0001$. As the planned comparisons reveal, the Response Mode Switching cost for confirming trials turns into a slight benefit for rejecting trials, that is, we replicate the asymmetrical pattern for Response Mode switching costs observed in experiment 4 (compare figure 3.2 and figure 3.1 in exp. 4). As in exp. 5 the interaction between Rule and Rule

Switching was not significant, that is, no asymmetrical pattern was observed for rule switching costs.

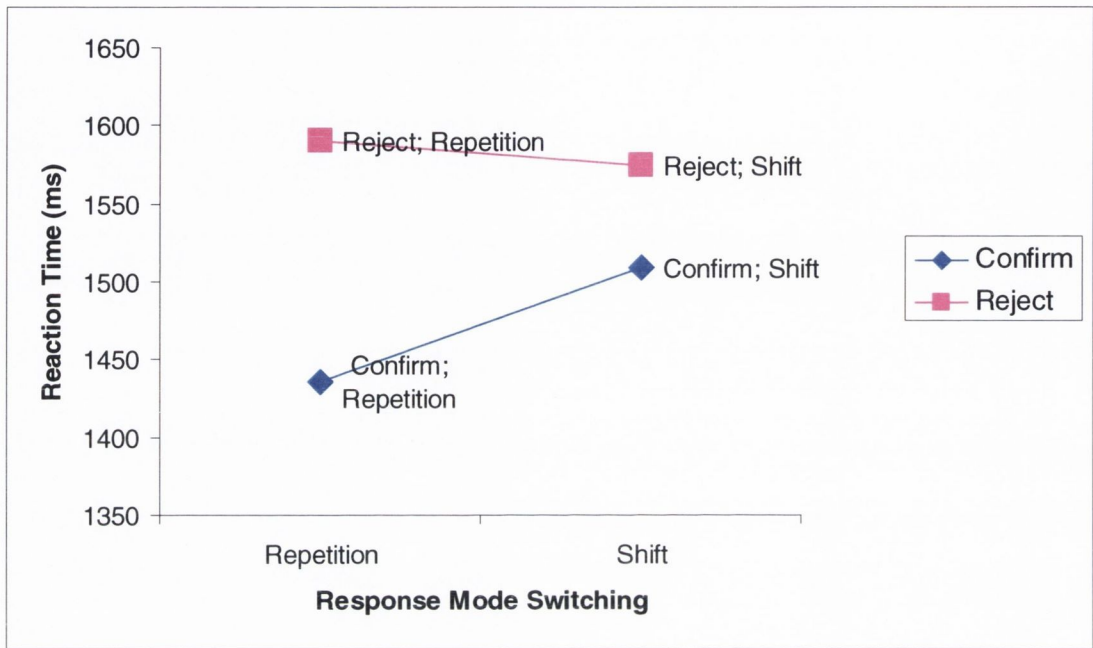


Figure 3.2. Reaction times to confirm and reject triples on response mode repeat and response mode shift trials in exp. 5.

The two-way interaction between the variables Rule and Response Mode (see figure 3.3), $F(1, 19) = 11.44$, $p < .003$, replicated the one observed in experiment 5. The effect of response mode was smaller when testing the descending rule (although it should be noted that the effect was still significant in this case, $F(1, 19) = 8.77$, $p < .008$). Thus, the advantage for confirmation over rejection was 158 ms. when testing the ascending rule, but only 62 ms. when testing the descending rule.

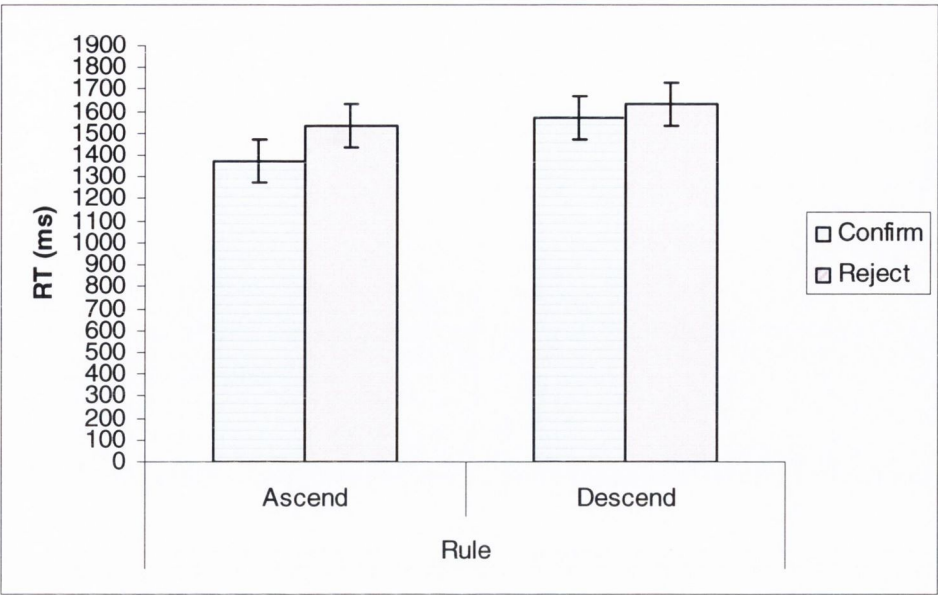


Figure 3.3. Reaction times to confirm and reject triples for trials under the ascending and descending rules.

The primary novel finding of this experiment is that, as predicted, a one-way interaction between Rule Switching and Response Mode Switching was observed, $F(1, 19) = 13.04$, $p < .0018$, due to the Rule Switching effect being significant only in trials where the response mode repeats $F(1, 19) = 12.50$, $p < .0022$, (see figure 3.4). This result means that response mode switching costs are only present when one does not switch rules, and conversely, that rule switching costs are only present in response repeating trials. In other words, for either variable, there is no cost of switching if the other variable switches as well. As this last interaction seems to be pointing to a mutual facilitation between both kinds of switching, we calculated the Pearson correlation for the Rule switching and Response Mode switching costs, which was marginally significant, $R = .55$, $p < .073$. Taken together, the interaction and correlation between both switching costs suggest that they are operated by a common mechanism, or that at least they share some components of the switching operation.

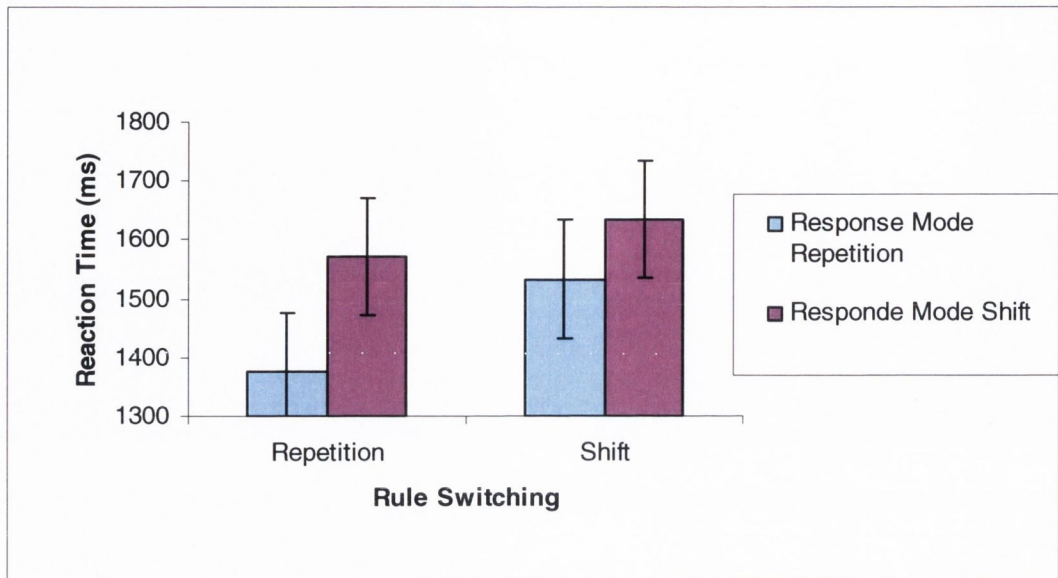


Figure 3.4. Reaction times to rule repetition and rule shift trials for Response Mode shift and repeat trials.

Discussion

The advantage for confirming trials was replicated again, over three experiments and with three different rules, all results point to an advantage for confirming over rejecting. As in experiment 5, there was an interaction between Rule and Response Mode, that is, the advantage for confirming trials is larger when testing the easier “ascending numbers” rule. This result is difficult to explain from the view that the advantage for confirmations is the result of a resource-independent process (e.g., Evans, 1989; 2003), if that were the case, the effect should be equally present for both rules, regardless of one being more demanding than the other.

The asymmetrical pattern for confirm-reject switching costs observed in experiment 4 is also replicated. There is a cost when switching from rejection to confirmation (reject to confirm), but a slight switching benefit when switching in the opposite direction (confirm to reject). Finally, the novel results in this experiment is that

the response mode switching and rule switching effects seem to be interrelated, as switching costs in one domain (rule or response mode) disappear when its counterpart switches as well. Such a result suggests a general switching operation capable of operating in parallel over different processes. This result, along with the asymmetrical cost of switching response mode, provides further support for the view that the response mode switching costs observed in expts. 4 and 6 reflect the need to switch between tasks of differing strength, with confirmation being the stronger or more practiced task. Before addressing the broad implication of these results for the fields of attention and reasoning, we report a final experiment that establishes some boundary conditions for the effects reported so far.

Experiment 7

The final experiment examines some boundary conditions for the switching costs reported so far, specifically we address the issue of the interference between rules. In experiment 6, where the rules were “ascending numbers” and “descending numbers”, all of the triples presented matched either the rule being tested or the alternate one. On the other hand, in experiment 5, while all of the triples were also ascending or descending, the more specific “descending by threes” rule caused triples that descend by a constant other than three to be incongruent with either rule. It is possible that the slower times when rejecting triples stem from the fact that triples that do not fit the tested rule are congruent with the alternate rule, thus creating an interference effect resulting in a larger latency of responses. In other words, in experiment 5 there were two different kinds of “rejecting” triples: those congruent with the alternate rule and those not congruent with any rule. A separate analysis was performed, but no significant differences were observed between the RT’s to either type of triple. However, it could be argued that the

fact that every triple in experiment 6 fell into the categories “ascending” or “descending” might have narrowed the possibility of observing any possible interference effects. In the experiments involving rule switching only two types of triple were presented (exps. 5 and 6). Our aim in experiment 7 is to observe the effect of introducing triples that did not match any rule (and are thus clearly different from the “rule-congruent” triples). Hence we introduce two very similar and much more specific rules, “ascending by ones” and “ascending by twos”. We also modify the triples in order to obtain a wider variety of “rule-incongruent” triples (see Materials section).

Method

Participants

Nineteen undergraduate psychology students (9 male, 10 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 19 and 37 years (mean = 26.9). Their vision was normal or corrected to normal.

Materials

Triples in this experiment belonged to one of six different categories: 1. numbers ascending by ones; 2. numbers ascending by twos; 3. numbers ascending by another constant, 4. numbers ascending in variable amounts; 5. numbers descending by another constant; and 6. numbers descending in variable amounts. In every trial there was a 50% likelihood for a confirm response to be required. When either the “ascending by 1s” or the “ascending by 2s” rule was being tested, half of the triples presented conformed to that rule. The other half were composed of the remaining five types of triples in equal amounts.

Design

The design of the experiment was based on the Response Mode variable, which had 3 levels, with rejecting triples divided into those fitting to the alternate rule (we will refer to them as “reject-congruent”); and those not fitting either rule (we will refer to them as “reject-incongruent”). Thus, the design was a 3 x 2 x 2 x 2 repeated measures: Response Mode (confirm, reject-congruent, reject-incongruent) x Response Mode Switching (repeat, shift) x Rule (ascending by 1s, ascending by 2s) x Rule Switching (repeat, shift).

Procedure

The procedure was the same as in the previous experiments.

Results

Only correct trials with response times that occurred between 2 standard deviations around each participant’s mean were included in the reaction time analyses. Mean accuracy was 94%.

There was a main effect of Response Mode, $F(2, 36) = 20.82$, $p < .000001$ (see figure 3.5). Participants were faster when rejecting incongruent triples (1457 ms. for confirm; 1485 ms. for reject-congruent; 1290 for reject incongruent).

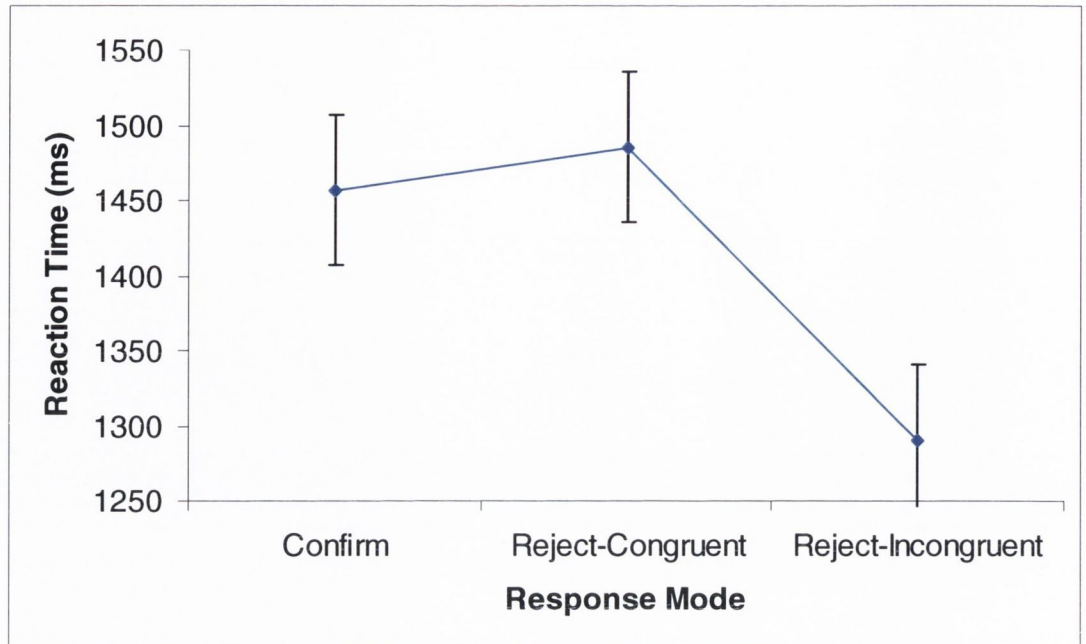


Figure 3.5. Reaction times to the 3 levels of the Response Mode variable.

There was a main effect of the variable Rule $F(1, 19) = 8.58, p < .008$, due to participants being 156 ms. faster at testing the ascending by 1s rule (average RT for ascending by 1s: 1389 ms.; average RT for ascending by 2s: 1455 ms.). An interaction between the variables Rule and Response Mode, $F(1, 19) = 4.84, p < .013$ was observed as well. As the planned comparisons revealed, the difference between confirm and reject-congruent trials is not significant for trials under the ascending by 2s rule, $F < 1$. If we do not consider incongruent trials, there is a confirmation bias for the easier rule, replicating previous results (see figure 3.6).

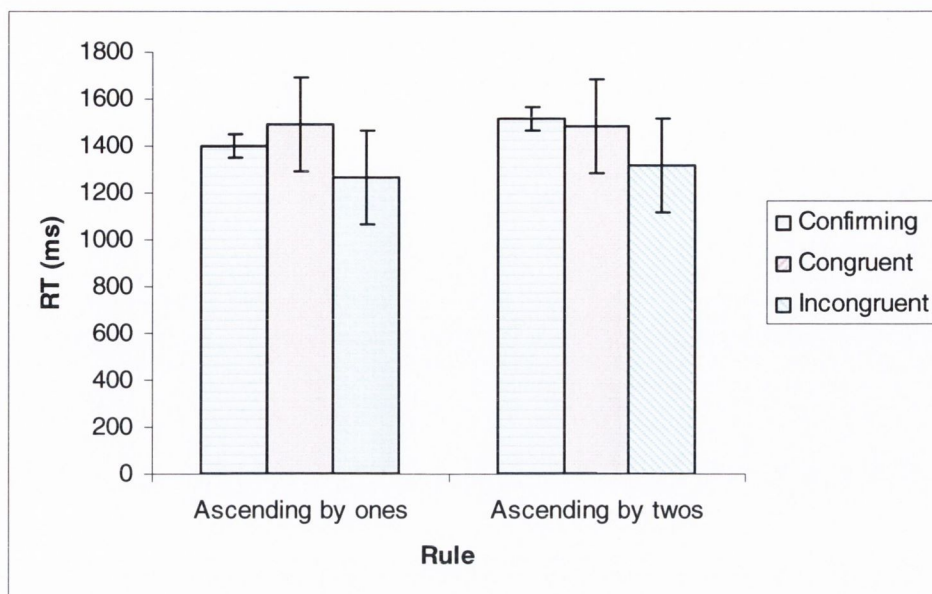


Figure 3.6. Reaction times to the 3 levels of the Response Mode variable for the two rules in exp. 7.

Although the analysis does not reveal an interaction between Triple Type and Rule Switching, a visual inspection of figure 3.7 and a posterior analysis that did not consider reject-incongruent trial clearly showed that the absence of rule switching costs holds only for “reject-incongruent” trials. Indeed, when considering only “confirming” and “congruent” trials, the effect of rule switching is significant, $F(1, 18) = 4.27$, $p < .0053$.

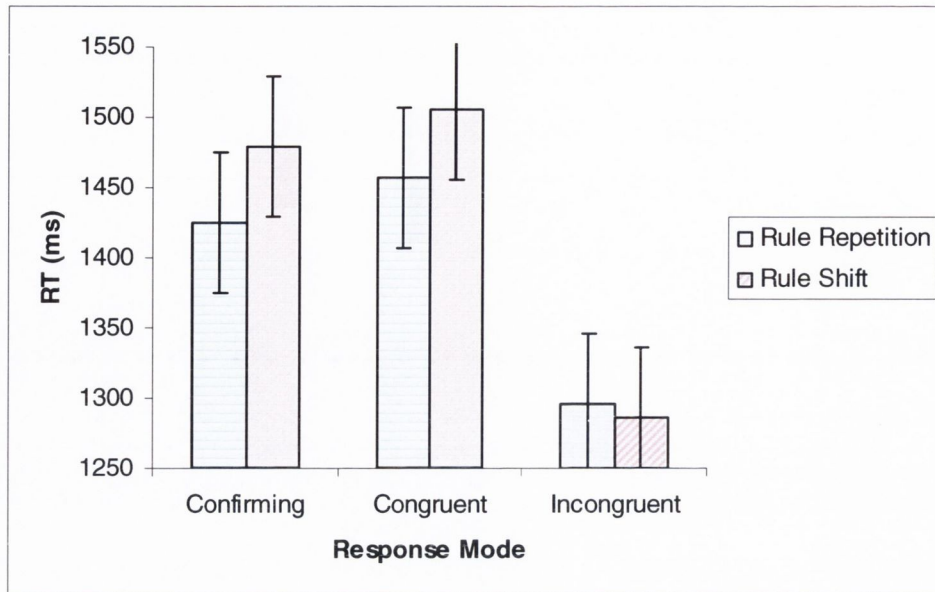


Figure 3.7. Switching costs for the three levels of Response Mode variable in exp. 7.

Discussion

The main difference between this experiment and the previous ones is the introduction of a wider variety of triples, thus leaving us with three levels of Response Mode (that is, confirm, reject-congruent and reject-incongruent). In experiment 5, the use of the “descending by 3s” rule caused triples that descended by a constant other than 3 to be incongruent with either rule, but no differences were observed. The introduction of a wider variety of triples in experiment 7 has had an interesting result. The main effect of Response Mode reflects in this case the shorter latency of responses to reject incongruent triples. Further analyses showed that when “incongruent” triples were not considered in the analyses, the main effect of Rule, and the interaction between Rule and Response Mode observed in previous experiments are replicated. Significant rule switching costs are observed only in “confirming” and “reject-congruent” trials, and there is a significant difference between confirm and rule-congruent trials for the easier rule. The fact that rule switching costs were only observed for confirm and reject-

congruent triples is interesting. It is possible that explicitly representing a rule makes the triples that fall under its scope subject to a range of attentional effects that are not observed when a triple is easily identifiable as not fitting any of the current rules. We will return to this issue, and its implications for hypothesis-testing research in the next section.

General discussion

Before discussing the implications of these results for both the reasoning and attention fields, let us summarize briefly the main findings:

Over three experiments, with different rules, and regardless of whether rule switching was required, the results point to an advantage for confirming over rejecting. Also, an asymmetrical cost of switching response mode was observed in experiments 4 and 6. These results support the view that confirmation and rejection are different tasks with confirmation being the stronger or more practiced task.

Rule switching costs were observed in experiments 5, 6 and 7, consistent with the view that only one rule can be held at a given moment in the focus of attention. In experiment 6, these rule-switching costs show a pattern of interaction and a correlation with the response mode switching costs which strongly suggests that a common switching mechanism is operating upon both rules and response modes. This result is a further piece of evidence that the response mode switching cost reflects the functioning of a switching process that is operating between tasks of differing dominance.

Implications for reasoning

As expected, a tendency towards confirmation is present in all four experiments, to reject that a triple of numbers fits a given rule seems to be a more demanding operation

than to confirm. Thus, these results support the idea that people tend to focus on positive information, and that this tendency can be observed in “low-level tasks” that do not have the complexity of hypotheses testing or conditional reasoning tasks. Indeed, our task involved categorization of number triples as fitting to a given rule, but even in experiment 4, where participants test a very simple rule (ascending numbers), differential effects of response mode were found.

There was a pattern across experiments 5, 6 and 7, in all of them a smaller advantage for confirmations was observed when testing the harder rule. The clearest example of this effect occurs when testing the more difficult “descending by 3s” rule in experiment 5, here the advantage towards confirming responses for this rule was half the size of the advantage observed when testing the easier “ascending numbers” rule. Thus, when involved in a more demanding operation, in this case testing a harder rule, the confirmation advantage is weakened, which would suggest that this effect is not resource-free or preattentive. This result is difficult to explain from the view that the advantage for confirmations is the result of a resource-independent process (e.g., Evans, 1989; 2003), if that were the case, the effect should be equally present for both rules, regardless of one being more demanding than the other. That is, if the advantage for confirmations is the result of a resource-independent process, then this effect should be not be affected by the difficulty of the tested rule. (pace Evans, 1989; see Gale and Ball, 2003).

An explanation of these results in terms of information gain (Oaksford and Chatter, 2002) could also been put forward. As the probability of a confirming response in our experiments was always of 50% further experiments would be needed in order to rule out this possibility.

There has been a growing consensus in the hypothesis-testing field about the importance (and the difficulty) of directing participants' attention to negative information in order to succeed in the 2-4-6 task (e.g. Gale and Ball, 2003). What Wason (1960) regarded initially as a kind of "motivational bias" towards confirming our own hypothesis, is regarded today as a more general "cognitive bias" towards positive information (Eysenck and Keane, 2005). Our results support this interpretation, and constitute a convergent piece of evidence in favour of the general proposal that the cognitive system shows a marked preference towards thinking about true possibilities (Johnson-Laird and Byrne, 2002).

There was an effect of Response Mode Switching in experiments 4 and 6. This effect displays an asymmetrical pattern, that is, response mode switching costs were present only when switching into the easier task, which in this case is "confirming". This pattern has been typically observed when switching between tasks of unequal dominance (e.g. Allport et al., 1994; Allport and Wylie, 1999; 2000; Yeung and Monsell, 2003). Thus, not only is confirming faster than rejecting, this result is also evidence that the response mode switching cost reflects the functioning of a switching process that is operating between tasks of differing dominance. As suggested by the model theory (Johnson-Laird & Byrne, 1991, 2002), limitations in processing capacity are as responsible for the tendency to focus on positive instances as they are for the tendency to focus on single elements. . Research on hypothesis-testing (e.g., Kareev and Halberstadt, 1993; Tweney et al., 1980; Wason, 1960, 1962) and reasoning (e. g. Johnson-Laird & Byrne, 1991, 2005), suggests that people tend to think more easily about what is true and not about what is false.

Further research would be needed in order to ascertain the ultimate cause of this advantage for confirming responses. It could be argued that, as people tend to think

about true possibilities (Johnson-Laird and Byrne, 2002), these possibilities display a lower threshold of activation, in other words, they would be primed (e.g. Meyer and Schvaneveldt, 1977). These results could also be explained by higher processing demands for rejections, consistent with the view that people represent what is false by negating what is true (see e.g., Santamaría and Espino, 2000, 2006). These are only tentative explanations, and perhaps they are not mutually exclusive, but the aim of this experiment was to establish that this tendency to think about true possibilities is a basic cognitive phenomenon which can be observed in more basic tasks than those typically employed in thinking paradigms. In this sense, it is worth noting that whereas capacity limitations are a well established phenomenon for cognitive psychology, the ultimate cause for these limitations remains a very disputed matter (see Meyer and Kieras, 1997a, 1997b for an approach in terms of strategies rather than limitations).

The rule switching effects found in this experiment might shed light on the cause of the success of the dual-goal or DAX-MED manipulation in improving participants hypothesis-testing performance in the 2-4-6 task (Tweney et al., 1980, Wason, 1960). In those studies, instead of asking participants to figure out a rule by proposing triples and receiving positive or negative feedback, they were asked to discover two rules, arbitrarily called DAX and MED. Thus, rather than receiving positive or negative feedback, each triple was classified as fitting to one of the rules. What is important for the success of this manipulation would appear to stem from its effect on the representation of the rules (e.g., Gorman et al., 1987; Tweney et al., 1982). Gorman et al. (1987) hypothesized that this manipulation fostered a unified mental representation of the two rules. On the other hand, Tweney et al. (1982) suggested that the logical structure of the problem was altered in a way that encouraged participants to switch from one hypothesis to another, thus avoiding rigid testing of a single hypothesis. Our

results suggest that a single representation of two rules cannot be easily constructed (pace Gorman et al., 1987). For example, in experiment 6, where both rules are known, simple and similar (ascending or descending integers), there was a clear rule switching cost, indicative of two distinct representations for the two rules. Moreover, in experiment 7, rule switching costs were only observed for triples that match one of the rules being tested, and those triples that were incongruent with both rules were easily rejected. In this sense, the DAX-MED manipulation could be seen as a way to induce participants to consider all of the triples that do not fit the original rule (DAX) as instances of an explicitly represented alternative rule (MED), thus increasing the search space by inducing participants to search for examples of MEDs. Thus, this manipulation would make all information positive, that is, no rejections are needed, which fits well with our observation that people find confirmation easier than rejection.

Evans (1989) gives an account of the effect of the dual goal instructions which also highlights the potential difficulty to deal with negation. Concretely, Evans proposes a *positivity bias*, a selective pre-attentional process that causes people to attend to positive rather than negative information, in the case of the rule-discovery task, participants tend to test positive cases that match the condition of their hypotheses. Thus, the better performance observed with dual goal instructions may be due to the linguistic labelling of the feedback to the triples that do not fit the rules as MED, thereby avoiding a negative label and not disturbing the participant's attention to this information (but see Gale and Ball, 2003). Even though the interaction between Rule and Response Mode did not reach significance, planned comparisons showed no advantage for confirming responses when testing the descending by 3s rule. Results show that the difference was due to confirmation being slower when testing the more difficult rule. This interaction is evidence counter to the view that the advantage for confirmations is the result of a

resource-independent process (e.g., Evans, 1989; 2003), if that were the case, the mentioned effect should be equally present for both rules, regardless of one being more demanding than the other.

Implications for attention-switching

Sizeable rule switching costs were obtained in experiments 5 and 6, supporting the view that people can only consider one rule at a time. Moreover, these rule switching costs show an interesting pattern in relation to the response mode switching costs. This interaction, found in experiment 6, is an underadditivity between rule switching costs and response mode switching costs. The cost for switching both rule and response mode is no higher than the cost of switching only one of them. Moreover, both switching costs were moderately correlated. These results favour a view in which a single attention switching mechanism is capable of operating simultaneously over different inputs (objects and operations), rather than different switching mechanisms for each domain. The correlation and interaction between switching costs and the asymmetrical cost of Response Mode switching, together with the advantage for confirming responses, all points to confirmation and rejection being two tasks differing in strength, rather than two outputs of the same operation.

Arrington, Altmann and Carr (2003; see also Kleinsorge, 1999) showed that task switching costs can be significantly reduced as the similarity between components of both tasks is increased. We could argue that the similarity between the rules for every experiment varies in this order: experiment 7 has the more similar rules (“ascending by ones” and “ascending by twos”; next experiment 6 (“ascending” and “descending”); and finally experiment 5 has the more dissimilar rules (“ascending” and “descending by 3s”). The differences in RT between these pairs of rules (averaging over rules switches

and rule repetitions) follow that arrangement: 60 ms difference between rules in experiment 7; 150 ms. in experiment 6; and 200 ms. in experiment 5. But, interestingly, this pattern is followed as well by the rule-switching costs (50, 80 and 150 ms. respectively), an observation that could be interpreted as an effect of similarity in object-switching, thus drawing similarities between the fields of task and object switching-costs.

Our results are relevant to attention switching research, insofar as they have unveiled the possibility of an attention switching operation capable of operating simultaneously over rules and responses, consistent with the evidence that different types of switching share a common neural substrate (e.g., Pereda et al., in preparation; Wager et al., 2004). Also, we have provided evidence of a task-switching-like similarity effect on object-switching costs.

Finally, these experiments constitute an example of how a very simple task can be used in order to ascertain the low-level attentional determinants for “higher” reasoning phenomena. In the next chapter, we intend to provide convergent evidence for our interpretation of the switching costs reported in this chapter by means of a neuroimaging technique.

Chapter 4 – An fMRI investigation of rule and response mode switching costs

In this chapter we present a study which employed functional magnetic resonance imaging (fMRI) to investigate the neural basis of the attention switching costs and the “truth” effect. The aim is to provide convergent evidence for the results reported in chapter 3, and the interpretations of the data we drew. More specifically, we intend to compare the effects on brain activation of switching between rules governing the behaviour of number triples, and switching between response modes (confirmation or rejection of the triple as an instance of the tested rule). In chapter 3, we suggested that these two switching types are seemingly operated upon by a single attention-switching mechanism, capable of operating simultaneously over rules and response modes. However, an important difference between rule and response mode switching was observed, only the latter entailed an asymmetrical switching cost pattern. It is suggested that, inasmuch as these two types of switching seem to be operated by the same mechanism, a similar activation map should be observed for each of them. Furthermore, it is expected that the asymmetrical pattern of response mode switching will be related to functional activation. Therefore, the study presented in this chapter has two main conditions, a response mode switching condition necessary and a rule switching condition.

Blocks with high and low numbers of confirming responses were also compared. We expected blocks with low numbers of confirming responses (i.e. high numbers of rejecting responses) to elicit greater activity in areas related to the monitoring of demands for cognitive control, such as the anterior cingulate area.

Introduction

The main aim of this study is to provide further convergent evidence of the main findings reported in chapter 3 by means of fMRI. In addition, the study will allow us to replicate the behavioral effects reported there. In chapter 2, it was shown that people are faster to confirm than to reject that a given triple of numbers follows a previously specified rule and that switching between response modes (confirmation and rejection) entailed an asymmetrical switching cost. It was suggested that these results were consistent with confirmation and rejection being different operations differing in dominance. Also in chapter 3, we reported that rule and response mode switching costs entered into an underadditive interaction, such that switching both elements was easier than switching either of them separately. In addition to this, both switching costs were correlated. These latter results led us to propose that a common switching mechanism operated between both rules and response modes (objects and tasks). Furthermore, response mode but not rule switching costs exhibited an asymmetrical pattern, in other words, costs were only observed for the easier response mode, while no difference in costs was observed for the two rules tested.

There are three main objectives in this study: first, to further test the hypotheses that both switching operations share the same mechanism by observing the brain activation related to each of them, which we expect to be similar; second, to provide some insight into the relation between asymmetrical switching costs and brain activation, an issue that has seldom been addressed in the literature on the neural basis of attention switching; and third, to compare the effects on brain activation of performing blocks with a high or low proportion of confirming responses.

Functional magnetic resonance imaging (fMRI)

fMRI is a method of imaging both the structure and function on the brain in vivo. When placed in a large magnetic field the magnetisation of water molecules in the blood lines up with the external field. Microwave radiofrequency can then be applied, which has the effect of disturbing the magnetic alignment of the water, when this radiofrequency is turned off the magnetisation of the water relaxes back to alignment with the external field, re-emitting the radiofrequency in the process. This re-emission of the radiofrequency is the basis for the magnetic resonance signal. The majority of fMRI studies infer brain activation from the BOLD (blood-oxygen level dependence) effect. This effect is based on the blood flow increase produced when neurons fire, due to the higher consumption of oxygen there is an decrease in the concentration of deoxyhaemoglobin and an increase in the MR signal. fMRI is a safe and widely used technique that offers excellent spatial resolution.

Neural basis of attention-switching

Neuroimaging studies of task switching have identified numerous brain areas that are more active in switching trials, mainly in medial and lateral regions of the prefrontal cortex, parietal lobe and cerebellum (see e.g. Monsell, 2003, for a review). Not surprisingly, activation maps for attention-switching often show a significant overlap with those attributed to WM processes, such as storage and rehearsal (e.g., D'Esposito, Aguirre, Zarahn, Ballard, Shin, and Lease., 1998; Garavan, Ross and Stein, 2000; Smith and Jonides, 1999). There is much agreement as to the observation that attention switching is achieved by means of a broadly distributed set of areas, mainly frontal and parietal (e.g., Wager et al., 2004). In a meta-analyses of several neuroimaging studies on attention-switching, Wager et al. (2004) found that several posterior (parietal and

occipital) and premotor regions were reliably activated across studies, however, anterior frontal regions were only active at lower statistical thresholds. Importantly, this meta-analysis included studies involving rather different types of switching (i.e., switching between rules, objects, attributes, etc.) and found, broadly, that the same regions of the medial prefrontal, premotor and parietal cortices were active. These authors (Wager et al., 2004) highlighted the necessity of studies involving multiple shift types in the same participants which is one of the main features of this study.

Brain activity related to confirmation versus rejection

We expect that blocks with higher levels of rejection will produce higher activation levels in the cingulate gyrus, consistent with the view that rejection is a different and more demanding operation than confirmation. The cingulate gyrus has been ascribed a main role in selective attention and performance monitoring (e.g., Cohen, Aston-Jones and Gilzerant., 2004; Duncan and Owen, 2000; McDonald, Cohen, Stenger and Carter, 2000; Venn, Cohen, Botvinick, Stenger and Carter, 2001). Cingulate activity has been observed as well when tasks require divided attention, novel responses, overcoming of a prepotent response (Dehaene, Posner and Tucker, 1994; Carter, Braver, Barch, Botvinick, Noll and Cohen, 1998), and it has been also related to reasoning (e.g. Fangmeier, Knauff, Ruff and Sloutsky, 2006; Ruff, Knauff, Fangmeier and Spreer, 2003). Many authors agree that the anterior cingulate monitors the demands for cognitive control by detecting conflict situations and implementing this control when needed (e.g., McDonald et al., 1998). Thus, given that people tend to think about true possibilities (Johnson-Laird and Byrne, 2005), and thinking about false possibilities or counterexamples places a higher demand on attentional resources, an increased cingular

activity is to be expected in blocks where participants have to reject triples, rather than confirm them.

Experiment 8

The task was a computerised derivative of Wason's 2-4-6 task, in which participants have to determine if triples of numbers either *confirm* or *reject* a previously specified rule (i.e., ascending integers). Participants do not have to generate triples, nor do they have to ascertain the rule that governs the behaviour of the triples, as they are informed about it in advance, hence the task requires categorization of confirmatory and disconfirmatory instances in light of an hypothesis (the rule) rather than testing a hypothesis. There are two main conditions in this study. One condition, Response Mode Switching, requires testing a single rule in each block of trials, the focus being on the activation associated with shifts in response mode. Shifts from confirmation to rejection are referred to as "type 1 shifts", while shifts from rejection to confirmation are referred to as "type 2 shifts". The second condition, Rule Switching, involves continuous switching between two different rules. The rules are "ascending numbers" and "numbers descending by 3s", the focus being on the activation associated with shifts of rule. Shifts from the ascending rule to the descending by 3s rule are referred to as "type 1 shifts", whereas shifts in the opposite direction are referred to as "type 2 shifts". We expect to observe similar maps for both conditions (i.e. both kinds of shift), involving medial prefrontal, premotor, parietal and occipital activation (e.g., Garavan et al, 2000; Wager et al., 2004). But, as noted above, a fundamental difference between both switching costs is that only for response mode switching is an asymmetrical pattern expected (i.e., higher switching costs for type 2 response mode shifts). We therefore

predict that different activation will be observed for type 1 and type 2 shifts for response mode switching only.

In the response mode switching condition, half of the blocks require a high proportion of confirming responses, while the other half require a high proportion of rejecting responses. We expect the “rejection blocks” to elicit a greater cingulate activity than “confirmation blocks”.

Method

Participants

Twelve undergraduate psychology students (5 male, 7 female) took part in this experiment and received course credits in exchange for their participation. Their ages ranged between 21 and 30 years (mean 23.4 years). Their vision was normal or corrected to normal.

Task design

The materials consisted of triples of numbers separated by hyphens (e.g. 2-4-6), which were presented on the centre of the screen, in a black font (Courier New, 18 mm) in the response mode switching condition, and in red or blue font in the rule switching condition (the colour identified the rule to be tested on the current trial). The triples were composed of one and two digit numbers, in either ascending or descending order, with the constraint that for every triple, the numerical difference between the first and second numbers was equal to the difference between the second and the third.

Participants were told that they would see a series of number triples. The participant's task was to decide, for every triple, if it conformed to the rules “ascending numbers” or “numbers descending by 3”. In the rule switching condition, the colour of

the triples cued the rule to be tested (red triples: ascending rule; blue triples: descending by 3s rule). In the response mode switching condition a single rule was tested in each block of trials, and the participants were informed about it at the beginning of the block. Participants pressed one of the two responses keys in a response pad with the middle and index fingers of the right hand, one for confirmation and one for rejection.

The stimuli were presented on a computer screen controlled by a PC (Pentium III) that also recorded participants' responses. The E-prime software (Schneider, et al., 2002) generated and controlled the stimulus presentation. Every trial began with the presentation of a number triple in the centre of the screen, which remained for 2000 ms. After the response, participants received feedback (correct or incorrect) during 1000 ms. and the next trial began. There were 30 triples per block, with a 4 second preparation period at the beginning and a 26 second rest period at the end, which makes a total of two minutes for each of the 12 blocks.

Prior to the scanning session participants were trained on both conditions. The scanning session was divided into 8 runs, with 4 runs for each of the two conditions. Every run consisted of three blocks of trials, lasting 6 minutes. The order of presentation of the rule switching and response mode switching conditions was counterbalanced. There were six randomly placed shifts per block in both conditions.

The response mode switching design was a 2 (Response Mode: confirm, reject) x 2 (Response Mode Switching: repeat, shift) x 2 (Rule: ascending, descending by 3s) repeated measures ANOVA. The last variable, Rule, was blocked, each rule was tested on half of the blocks, and the order of presentation of rules was counterbalanced. Event-related activation maps were obtained for type 1 and type 2 response mode shifts.

The rule switching design was a 2 (Rule: ascending, descending by 3s) x 2 (Rule Switching: repeat, shift) x 2 (Response Mode: confirm, reject) repeated measures

ANOVA. Even though in this version there also were shifts in response mode, these response mode shifts never occurred in the same trial as a rule shift, and the response mode switching variable was not considered in the analyses. Event-related activation maps were obtained for type 1 and type 2 rule shifts.

fMRI parameters

All scanning was conducted on a Philips Intera Achieva 3.0 Tesla MR system (Best, The Netherlands) equipped with a mirror that reflected the 640 x 480 pixel display, projected on a panel placed behind the participants' head outside the magnet. The mirror was mounted on the head coil in the participants' line of vision. Imaging started with 31.5 seconds of standard scout images to adjust the head positioning, followed by a reference scan to resolve sensitivity variations. Imaging used a parallel SENSitivity Encoding (SENSE) approach (Pruessmann et al., 1999) with reduction factor 2. 180 high-resolution T1-weighted anatomic MPRAGE axial images (FOV 230 mm, thickness 0.9 mm, voxel size 0.9 x 0.9 x 0.9) were then acquired (total duration 6 minutes), to allow subsequent activation localization and spatial normalization. Thirty-two non-contiguous (10% gap) 3.5 mm axial slices covering the entire brain were collected using a T2* weighted echo-planar imaging sequence (TE = 35 ms, TR = 2000 ms, FOV 224 mm, 64 x 64 mm matrix size in Fourier space). The functional scan had a total duration of 48 minutes.

fMRI analyses

The fMRI data were analysed using the AFNI software package (Cox, 1996). Time-series data were motion-corrected using 3D volume registration (least-squares alignment of three translational and three rotational parameters). Activation outside the brain was

also removed using edge detection. A block analysis was performed to estimate the activation for the “high confirmation” and “high rejection” blocks separately in the response mode switching condition. These ON-OFF block regressors were convolved with a standard haemodynamic response to accommodate the lag time of the blood oxygen level-dependent (BOLD) response. Multiple regression analyses were then used to determine the average level of block activation as a percentage change relative to the rest periods (baseline). The baseline activation was derived from averaging the rest periods in each block over all the runs of the task.

Deconvolution techniques for successful response mode and rule shifts calculated event-related activation. Separate haemodynamic response functions at 2 s temporal resolution were calculated. Although the stimulus stream was presented at 1 Hz, all events of interest were time-locked to the beginning of the two-second whole-brain volume acquisition. A multiple regression analysis was used to derive estimates for the time-point parameters of the haemodynamic response functions, by estimating the signal contributed by each individual event type to the overall time series. In the present analysis regressors for type 1 and 2 response mode and rule switches were entered, and the regression estimated the signal contributed by each of these events over and above that accounted for by the ongoing task. The haemodynamic response functions were then modelled voxelwise with a gamma-variate function using non-linear regression (Murphy and Garavan, 2005; Ward, Garavan, Ross, Bloom, Cox and Stein, 1998). An area-under-the-curve measure of the gamma-variate model was expressed as a percentage of the tonic baseline activity and served as the activation measure for the event-related responses. The activation map for shifts therefore represents the activation during successful switching events that is significantly greater than during the ongoing task.

Activation maps were skull stripped and warped into a standard stereotaxic space (Talairach and Tournoux, 1988) and spatially blurred with a 4.2-mm full-width at half-maximum isotropic Gaussian filter after performing a second edge detection on the skull stripped brain. Group activation maps were determined with one-sample *t* tests against the null hypothesis of zero activation changes (i.e. no change relative to tonic task-related activity). Significant voxels passed a voxelwise statistical threshold ($t = 4.14045$, $P < 0.005$) and were required to be part of a larger 185 μl cluster of contiguous significant voxels. Thresholding was determined through Monte Carlo simulations and resulted in a 5% probability of a cluster surviving due to chance. The activation maps were then combined deriving OR maps. An OR map includes the voxels of activation indicated as significant from either of the constituent maps. An OR map was obtained for blocks with high and low levels of confirmation in the response mode switching condition. Two activation maps were obtained for each switching condition by deriving an OR map for shifts in either direction, and those two maps were submitted to a further OR map and to a conjunction analyses. In the response mode condition experiment “type 1” shifts were those from confirmation to rejection, whereas “type 2” switches were those in the opposite direction. In the rule switching experiment “type 1” shifts were those from the “ascending” rule to the “descending by 3s” rule, while “type 2” shifts were those in the opposite direction. The mean activation for clusters in the combined OR maps was calculated for each subject and condition for the purposes of an ROI analysis, which consisted of a series of 2 x 2 ANOVAs, with the factors switching condition (rule switching, response mode switching) and shift type (type 1, type 2). The conjunction analyses involved deriving an AND map, that is, a map including the coinciding clusters of activation in both switching conditions.

Results

Behavioral results

In the response mode condition, a main effect of response mode switching, $F(1, 11) = 6.82, p < .001$, reflected a switching cost of 102 milliseconds (repetition: 1112 ms.; shift: 1214 ms.). The expected interaction between response mode switching and response (confirm or reject) was observed as well, $F(1, 11) = 23.81, p < .001$. Costs were significant both for confirmation, $F(1, 11) = 69.88, p < .001$, and rejection trials, $F(1, 11) = 11.55, p < .006$, however, the cost was about a hundred ms. larger for confirmations than rejections (146 vs. 57 ms., see figure 4.1). Thus, we replicate the asymmetrical switching pattern for response mode switching costs observed in experiments 4 and 6. A main effect of rule, $F(1, 11) = 6.82, p < .024$, reflected the fact that participants were 74. ms faster to test the ascending rule. Overall accuracy was 96%, no significant effects were observed for this dependent variable.

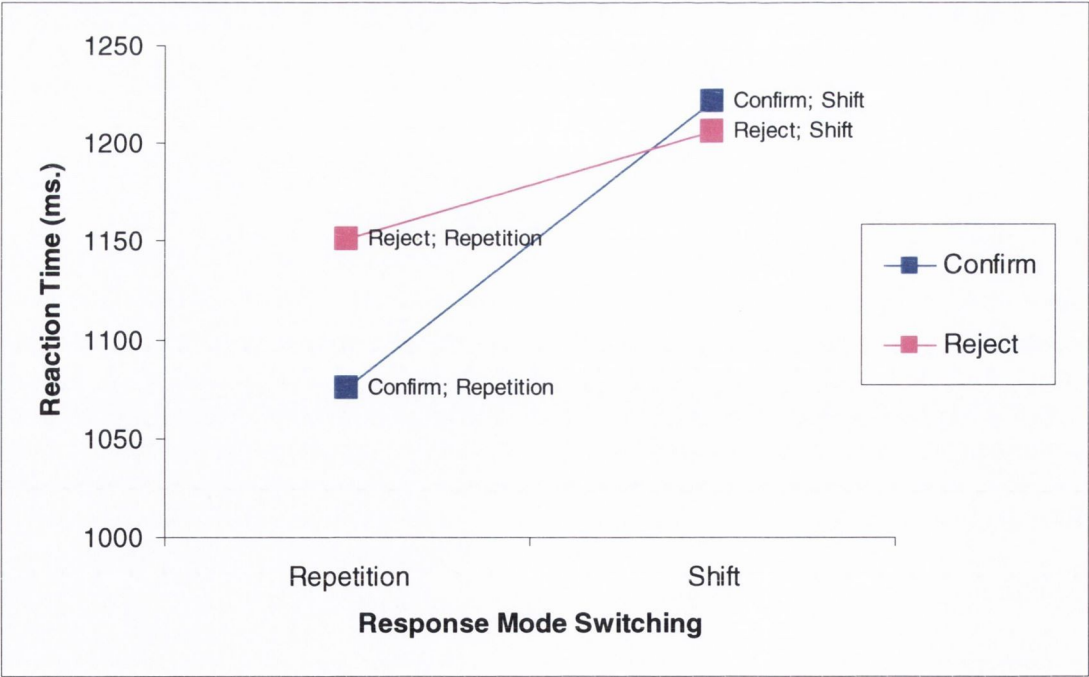


Figure 4.1. Switching costs for confirmations and rejections in the response mode switching condition.

In the rule switching condition, the repeated measures ANOVA showed a significant main effect of rule switching, $F(1, 11) = 83.29, p < .001$, due to a rule switching cost of 217 ms (see figure 4.2). There was a main effect of rule, $F(1, 11) = 23.81, p < .001$, reflecting that participants were 90 ms. faster when testing the ascending rule. In this case there was no interaction between rule and rule switching, in other words, no asymmetrical pattern was observed for rule switching costs as observed in experiments 5 and 7. Overall accuracy was 93%, again, no significant effects were observed for this dependent variable. As expected no asymmetrical pattern is observed for rule switching costs, contrary to what was observed for response mode switching costs.

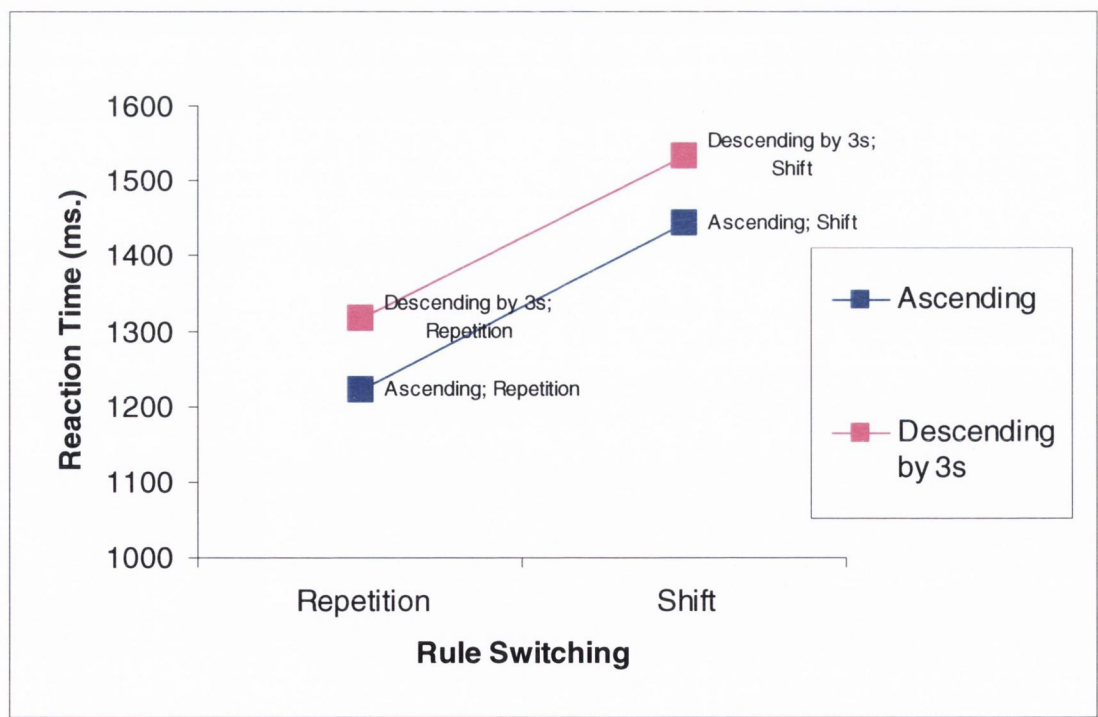


Figure 4.2. Switching costs for both rules in the rule switching condition.

Event-related functional activation

The results of the conjunction analysis, which consisted of an AND map including the voxels of activation indicated as significant in both the response switching and rule switching conditions, are reported in table 4.1 and figure 4.1 Large clusters in the inferior parietal lobule bilaterally overlap in the rule switching and response mode switching maps, with smaller clusters of common activation in the left superior frontal, left cingulate and left precentral gyrii, right superior parietal lobule and right inferior semi-lunar lobule of the cerebellum.

rule AND response mode	BA	Side	Centre-of-mass (x, y, z)	Cluster vol.
<i>Frontal Lobe</i>				
Precentral gyrus	6	L	-44, 0, 30	320
Superior frontal gyrus	6	L	-3, -4, 53	247
Cingulate gyrus	32	L	-9, -18, 35	162
<i>Parietal Lobe</i>				
Inferior parietal lobule	40	L	-33, 56, 37	3838
Inferior parietal lobule	40	R	43, 42, 40	3142
Superior parietal lobule	7	R	30, 61, 40	275
Postcentral gyrus	40	R	41, 25, 47	227
<i>Cerebellum</i>				
Inferior semi-lunar lobule	-	R	7, 69, -36	305

Table 4.1. Clusters active for both response mode switching
and rule switching.

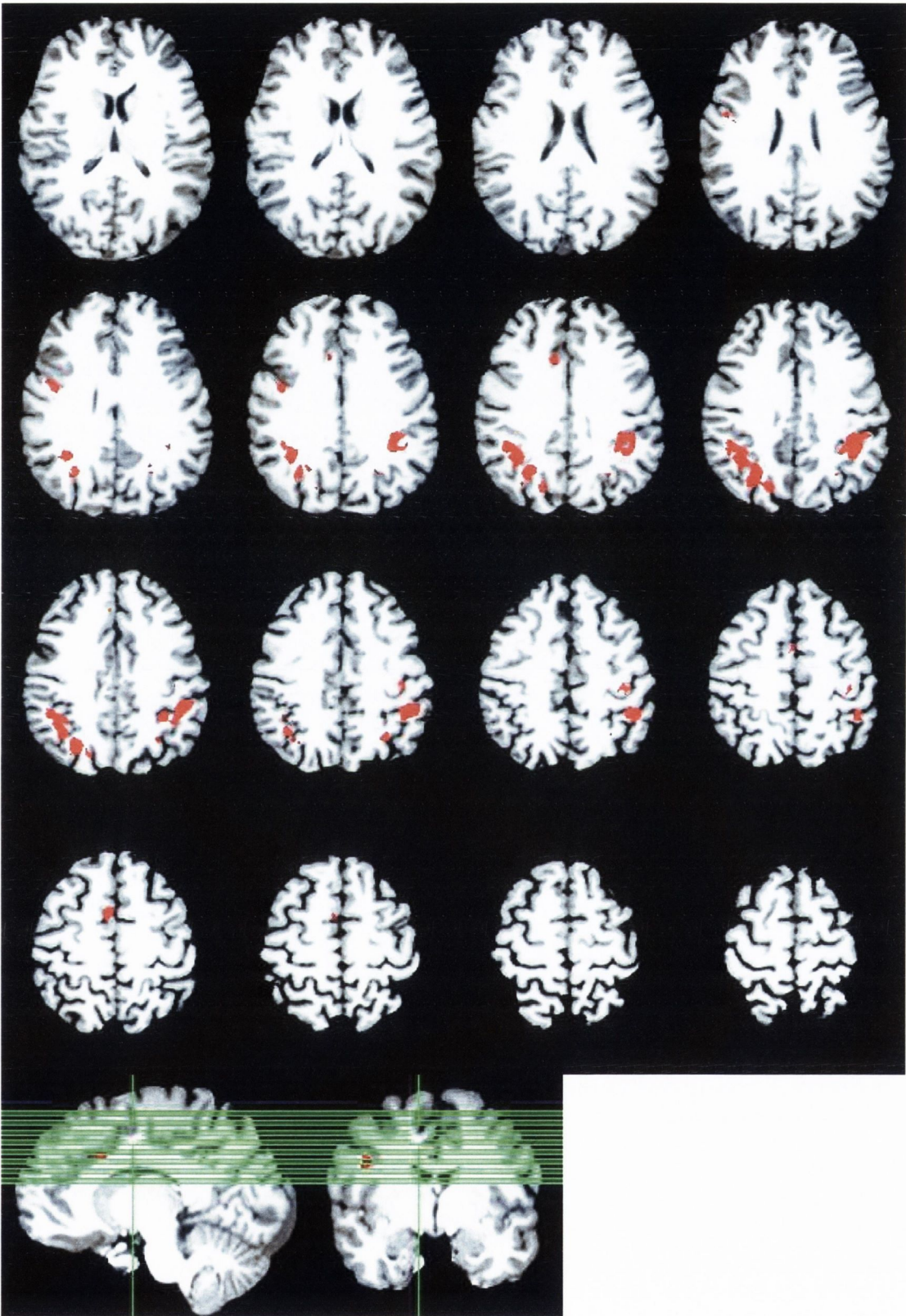


Figure 4.3. Clusters activated by rule and response mode switching.

This conjunction analysis was designed to find the common areas of activation. A further OR analysis was carried to determine if there are any areas unique to each condition and to unveil any effects in terms of shift type. Table 4.2 thus lists those regions that showed significant activation as a function of response mode switching or rule switching . It should be noted that only in the response mode switching condition did type 1 and type 2 switching costs differ, with the former in each case being significantly smaller than the latter, that is, there was lower activation when switching from confirmation to rejection.

rule OR response mode	BA	Side	Centre-of-mass (x, y, z)	Cluster vol.	Switching Condition	Shift Type	RMS T1 vs. T2	RS T1 vs. T2
Insula	47	L	-38, -21, -0	2278				
<i>Frontal Lobe</i>								
Medial frontal gyrus	10	L	-0, -3, 50	3886				
Inferior frontal gyrus	9	L	-44, -6, 30	3640	*	*	*	
Inferior frontal gyrus	9	R	42, -4, 27	2470			*	
Precentral gyrus	6	L	-27, 11, 54	2153			*	
Cingulate gyrus	32	L	-10, -22, 30	1475			*	
Middle frontal gyrus	9	R	44, -20, 33	1061			*	
<i>Parietal Lobe</i>								
Inferior parietal lobule	40	R	36, 40, 43	29912				
Inferior parietal lobule	40	L	-29, 62, 36	17236				
Postcentral gyrus	3	R	-54, 16, 41	1625			*	
<i>Temporal Lobe</i>								
Middle temporal gyrus	39	R	45, 1, -17	772	*		*	
<i>Occipital Lobe</i>								
Fusiform gyrus	18	R	23, 92, -15	1459				
Lingual gyrus	18	L	-16, 70, -2	552	*			
Inferior occipital gyrus	18	R	45, 75, -0	530	*			
Lingual gyrus	18	R	25, 87, -2	372				
<i>Cerebellum</i>								
Inferior semi-lunar lobule	-	R	1, 70, -39	3657				

Table 4.2. Clusters of significant activation for both switching conditions. The last four columns summarize the results of the ANOVA on the activation scores.

The combination of the rule and response mode switching maps reflected significant activations in the insula, left medial frontal gyrus, left cingulate gyrus, left precentral gyrus, right middle frontal gyrus and inferior frontal gyrus bilaterally. Large areas of activation were observed as well in the parietal lobe, namely, in the inferior parietal lobule bilaterally and left postcentral gyrus. The left middle temporal gyrus was also activated, and further clusters of activation were observed in the occipital lobe, insula and cerebellum.

The 2 x 2 ANOVAs on the activation change scores revealed significant results for some of the regions. There are main effects of both switching condition (that is, a significantly different activation for response mode and rule switching) and shift type in the left inferior frontal gyrus, and even though the interaction is not significant, the difference in activation between type 1 and type 2 shifts is significant for the response mode switching condition, $F(1, 11) = 6.07, p < .031$, but not for the rule switching condition, $F < 1$. In the right inferior frontal gyrus neither the main effects nor the interaction are significant, however, the difference between type 1 and type 2 shifts is still significant for the response mode switching condition, $F(1, 11) = 6.80, p < .024$. This pattern of results, namely, absence of significant main effects or interaction, but significant differences between shift types for the response mode switching condition is observed in three further areas: the left cingulate gyrus, where type 1 and type 2 shifts differ only for response mode switching, $F(1, 11) = 6.86, p < .023$; right middle frontal gyrus, $F(1, 11) = 6.14, p < .030$; and a marginal trend in the same direction is observed on the left precentral gyrus. No significant effects were observed for either the right or left inferior parietal lobules, but there was a significant effect of shift type in the response mode switching condition for the left postcentral gyrus, $F(1, 11) = 6.23, p < .029$ for response mode switching ($F < 1$ for rule switching), this is the same pattern

reported above for most of the frontal areas. Only in the occipital lobe did we observe significant effects of switching condition. Specifically, there were main effects of switching condition in the left lingual gyrus, $F(1, 11) = 4.98, p < .047$, and the right inferior occipital gyrus, $F(1, 11) = 13.55, p < .009$. In both regions activation was greater for the rule switching condition.

There was a significant correlation between the activation scores and their corresponding behavioral switching costs for the *left cingulate gyrus*. Type 2 response mode shifts correlated with the corresponding behavioral switching cost (although the result fell short of significance), $r = .502, p < .096$.

Block-related functional activation

Our task produced several significant activations in the frontal lobes: right medial frontal gyrus, cingulate gyrus bilaterally, right precentral gyrus, left middle frontal gyrus and left inferior frontal gyrus. Clusters of activation were also observed in the left precuneus, left superior and middle temporal gyri, right parahippocampal gyrus, cuneus bilaterally and several areas of the cerebellum. Differential effects of the level of confirmation were observed in the cingulate gyrus, with higher activation for “high rejection” blocks in the right and left cingulate gyrus, and higher deactivation for “high rejection” blocks in the posterior cingulate. The left superior and middle temporal gyri also showed higher deactivation for “high rejection” blocks (see table 4.3). However, for many of the areas activated by the response mode switching task, deactivations rather than activations were observed, an observation that we are not able to explain and that may undermine the validity of this block analysis. There were significant correlations between the activation scores and the behavioral RT scores for the *left posterior cingulate* and “high rejection” blocks, $r = -.70, p < .001$.

Attention and Reasoning

	Side	Centre-of-mass (x, y, z)	Cluster vol.	Mean activation change		
				High Conf.	Low Conf.	P
Frontal Lobe						
Clastrum	L	-37, 16, 0	632	-.202	-.156	.202
Middle Frontal Gyrus	L	-33, 5, 53	330	.138	.073	.169
Cingulate Gyrus	L	-6, 20, 37	308	.139	.152	.058*
Cingulate Gyrus	R	21, 32, 33	263	-.064	.001	.001*
Inferior Frontal Gyrus	L	-46, -28, 3	183	-.022	-.076	.195
Precentral Gyrus	R	33, 12, 52	149	.057	.177	.084
Posterior Cingulate	L	-11, 49, 21	137	-.059	-.127	.079
Medial Frontal Gyrus	R	5, -65, 13	133	-.161	-.116	.752
Superior Frontal Gyrus	L	-25, -56, 13	129	-.112	-.094	.741
Parietal Lobe						
Precuneus BA7	L	-3, 54, 51	12165	-.282	-.155	.022
Precuneus	L	-26, 59, 35	770	.202.	.194	.856
Temporal Lobe						
Superior Temporal Gyrus BA41	L	-46, 34, 12	1175	-.334	-.180	.008*
Middle Temporal Gyrus	L	-59, 45, -4	931	-.159	-.075	.035*
Parahypoccampal Gyrus	R	19, 41, 8	196	.093	.027	.324
Middle Temporal Gyrus BA37	L	-52, 63, 6	160	-.190	-.107	.108
Middle Temporal Gyrus	L	-47, 75, 16	153	-.206	-.131	.109
Occipital Lobe						
Cuneus	R	4, 79, 19	4557	-.277	-.234	.302
Cuneus BA30	L	-12, 62, 6	2104	-.207	-.162	.361
Cuneus	R	24, 69, 26	1490	.250	.198	.262
Middle Occipital Gyrus BA19	R	36, 78, -9	255	.171	.206	.567
Cerebellum						
Culmen	L	-22, 43, -16	1148	-.167	-.124	.157
Culmen	R	25, 43, -16	723	-.162	-.126	.271
Vermis	L	-2, 66, -36	292	.123	.148	.469
Vermis	L	0, 44, -36	177	.117	.117	.995
Uvula	R	32, 84, -24	150	.310	.268	.290
Vermis	L	-8, 49, -22	130	.021	.100	.044

Table 4.3. Task activation for “high confirmation” and “high rejection” blocks.

Discussion

The role of the frontal and parietal lobes in attention-switching

Historically, the frontal lobes have been considered the seat of the higher cognitive functions, as suggested by lesion studies revealing an impaired executive functioning in frontal lobe patients. However, studies involving neuroimaging of tasks that tap central executive functioning often report extensive cortical activations ranging from prefrontal to primary visual areas (e.g. Owen, Evans and Petrides, 1996; Prabhakaran, Smith, Desmond, Glover and Gabrieli, 1997). Indeed, the present results suggest an important role of the parietal lobe in attention switching. Even though there was some overlap in the areas activated in the frontal lobes both by rule switching and response mode switching, only in the inferior parietal lobe overlapping clusters of activation insensitive to either shift type or switching condition were observed. In other words, the parietal lobe were equally involved in both rule and response mode switching. This observation fits well with that noted by Wager et al. (2004) that parietal activations are more consistent than frontal ones across neuroimaging studies of attention switching and they were the most statistically robust activations in the present study.

Parietal activation has been found in many verbal and visuo-spatial working-memory tasks (e.g., Bunge, Klingberg, Jacobsen and Gabrieli, 2000; Jonides, Schumacher, Smith, Koeppe, Awh, Reuter-Lorenz, Marshuetz and Willis, 1998; Paulesu, Frith and Frackowiak, 1993). With respect to attention switching, Le et al. (Le, Pardo and Hu, 1998; see also Dove, Pollman, Schubert, Wiggins and von Cramon, 2000) found that switching between two targets activated, amongst other areas, the superior parietal lobe, cuneus and precuneus. The importance of the precuneus in shifting attention in space is well established (e.g., Corbetta, Miezin, Shulman and Petersen, 1993; Posner, Walker, Friedrich and Rafal, 1984). However, recent studies

have also implicated this area in attention to object features (e.g., Fink, Halligan, Marshall, Frith, Frackowiak and Dolan, 1997; Nagahama, Okada, Katsumi, Hayashi, Yamauchi, Sawamoto, Toma, Nakamura, Nanakawa, Konishi, Fukuyama and Shibakasi, 1999). Sohn et al., (Sohn, Ursu, Anderson, Stenger and Carter, 2000) proposed a functional dissociation between posterior and inferior parietal cortex, with the former being involved in endogenous preparation for a task set and the latter involved in stimulus-driven aspects. However, Kübler et al. (2003) found widespread parietal activation that was sensitive to switching demands but not task specific, which the authors interpreted as suggesting that the parietal lobe contributions may go beyond those of stimulus representation, storage and maintenance. Our results support this interpretation, as they suggest that the parietal lobes are key in the execution of the switching operation.

Another aspect in which our results replicate those of Wager and colleagues (2004) is the absence of activations in the dorsolateral prefrontal cortex (DLPFC), contrary to what has been frequently reported in the literature (e.g., MacDonald et al., 2000). In many studies, this area has been ascribed the role of retrieving and maintaining rules for behaviour (e.g., Miller and Cohen, 2001; Wallis, Anderson and Miller, 2001). In light of this, given that only in the rule switching condition is a continuous retrieval and updating of two different rules needed, activation in the DLPFC should have been expected, but this was not observed. Therefore our results cast doubts on these interpretations of DLPFC function.

Activation in the left inferior frontal gyrus was higher for the rule switching condition. It is possible that this higher activation is due to subvocal rehearsal of the relevant rules (see Gohske, 2000; Luks, Simpson and Feiwell., 2002), given that only in the rule switching condition do participants need to switch between rules. Nevertheless,

it should be noted that this inferior frontal activation is bilateral, while it is widely agreed that language processes are lateralized to the left side of the brain. Bilateral inferior frontal gyrus activation has also been reported in other attention-switching studies (e.g. Dove et al., 2000; Garavan et al., 2000; Konishi, Nakajima, Uchida, Kameyama, Nakahara, Sekihara and Miyashita, 1998a; Loose, Kaufmann, Tucha, Auer and Lange, 2006), thus, the exact role played by this region remains to be elucidated. In this regard, it is worth mentioning the work by Konishi et al. (Konishi, Nakajima, Uchida, Sekihara and Miyashita, 1998b, see also Aaron and Poldrack, 2006; Shafritz, Kartheiser and Belger, 2005) who found activation in the right inferior frontal region in no-go trials but not on go trials, thus ascribing this region a role in the inhibition of inappropriate sets or responses. In this study, the right inferior frontal gyrus shows differential activations for type 1 and type 2 response mode shifts and no effect of switching condition, consistent with this region playing a role in task switching. However, no activation in either the right or left inferior frontal gyrii is observed on the conjunction map, suggesting that the involvement of this area is not related to the switching operation itself, an observation that fits better with the account by Konishi et al. on the role of this region.

Middle frontal and cingulate activations have often been reported in studies of attention switching (e.g., Garavan et al, 2000; Kübler et al., 2003), and in tasks involving, for example, manipulation of items in WM (e.g., Collette, Salmon, van der Linden, Chicherio, Belleville, Degueldre, Delfiore and Franck, 1999) and selective attention and performance monitoring (e.g., Badgaiyan and Posner, 1998; Venn et al., 2001) respectively. Both these areas show differential effects of shift type for the response mode switching condition, but no effect of switching condition which suggests a general involvement in the switching operation. Premotor and pre-SMA activations

have been previously reported for WM tasks (e.g., D'Esposito et al., 1998), as well as for attention switching (e.g., Garavan et al., 2000). There is evidence that the main role of this area in attention switching is the inhibition of prepotent motor responses in shift trials (Sylvester, Wager, Lacey, Hernández, Nichols, Smith and Jonides, 2003; see also Loose et al., 2006).

Finally, the observation of main effects of switching condition in the occipital activations might be explained by the different visual requirements of rule and response mode switching, as only the former requires participants to associate the relevant rule with the colour of the triple.

It could be argued that it is not sound to predict no differences in neural responses between two cognitive tasks. However we have also shown that although the areas for both kinds of switching are similar, they differ in that only one of them exhibits asymmetrical effects.

Brain correlates of asymmetrical switching costs

Asymmetries in behavioral switching costs are only observed in the response mode switching condition, and consistent with this observation, differences in activation between type 1 and type 2 shifts are only observed for the response mode condition. Lower levels of activation were observed for type 2 response mode shifts in the inferior frontal gyrus bilaterally, left cingulate gyrus, right middle frontal gyrus, right middle temporal gyrus and left postcentral gyrus. Furthermore, activation in the left cingulate gyrus correlated with the behavioral type 2 switching cost. In line with Logothetis (1999, 2000), this direct relations between behavioural and BOLD activation is a further proof that this signal is an adequate measure for cognitive activity. The fact that asymmetrical activation is observed mainly in the frontal lobe and not in the parietal

lobe, where coinciding clusters of activation for both types of switching were mostly observed, suggests that the asymmetry is not a property of the switching operation itself. Asymmetries could be therefore interpreted as a property of the tasks over which the switching mechanism operates. This question deserves further research.

Even though asymmetrical switching costs have received extensive attention in the task-switching literature (e.g., Allport et al., 1994; Meuter and Allport, 1999; Yeung and Monsell, 2003) this issue has seldom been addressed in event-related fMRI studies. One exception is a recent study by Wang and colleagues (Wang, Xue, Chen, Xue and Dong, 2007), in which Chinese college students who were learning English had to switch between naming pictures in their first and second languages. Similar to what has been found in this study, they found that the expected asymmetrical RT switching costs were reflected in higher activation for the lower cost, in this case switching into the more dominant language (Chinese), but they did not further interpret this result. The fact that activation is higher when the RT cost is lower might seem counterintuitive at first sight, as it happens with the phenomenon of asymmetrical switching costs itself. If switching trials produce a higher activity in the same areas activated in repeating trials (e.g., Kimberg, Aguirre and D'Esposito, 2000), then why are asymmetrical switching costs reflected in higher activation for the lower cost? It could be argued though that higher activation is needed in order to switch into the harder task, that is, the fact of having to switch into a harder task calls for more resources (e.g., Allport et al., 1994; Allport and Wylie, 1999), which is reflected both in larger activation and lower RTs. It is worth noting that in the left precentral gyrus higher activation was observed for type 2 response mode shifts instead, perhaps consistent with the role attributed to this area in the inhibition of prepotent responses (see above).

Cingular activity and rejection

Despite the caveats, higher levels of rejection produced higher activation levels in two bilateral locations in the cingulate gyrus, consistent with the view that rejection is a different and more demanding operation than rejection. The cingulate gyrus has been ascribed a main role in selective attention and performance monitoring (e.g., Badgaiyan and Posner, 1998; Cohen et al., 2004; Duncan and Owen, 2000; McDonald et al., 2000; Venn et al., 2001). Cingulate activity has been observed as well when tasks require divided attention, novel responses and overcoming of a prepotent response (Dehane, Posner and Tucker, 1994; Carter et al., 1998). Many authors agree that the anterior cingulate monitors the demands for cognitive control by detecting conflict situations and implementing this control when needed (e.g., Cohen et al., 2004; McDonald et al., 1998). Our results are consistent with this view.

The fact that a significantly greater deactivation is observed for “high rejection” blocks in the posterior cingulate is consistent with a pattern of activations which concerns a well known functional subdivision in the cingulate gyrus. While, as said before, the anterior part of the cingulate cortex has been considered by many as part of an attentional network in charge of executive functioning, the posterior cingulate cortex is ascribed a key role in assessing emotional and motivational information as well as regulating emotional responses (e.g., Drevets and Raichle, 1998; Bush, Lu and Posner, 2000). This pattern involves a reciprocal suppression of the affective subdivision during cognitive tasks, reflected in deactivations in the posterior cingulate when participants perform more demanding operations, which is precisely what we observe.

Conclusions

The present findings constitute a piece of convergent evidence in favour of the view that attention switching is achieved by means of a broadly distributed set of cerebral regions (e.g., Garavan et al., 1999; Kübler et al., 2000). Similar activation maps were observed for two different kinds of attention switching, that is, switching of rules and of response modes. However, it is mostly in parietal and premotor areas where these coincidences are observed. These results are in line with the observation that parietal activations are more consistent across studies of attention switching. It has also been shown that switching costs displaying an asymmetrical pattern are reflected in asymmetrical activations mainly located in frontal areas, with the higher levels of activation being related to the lower behavioral costs. Thus, the frontal lobe would appear to reflect patterns of activation more dependant on the particular demands of the currently relevant task. Finally, our results provide convergent evidence that switching into the stronger task is a more demanding operation than its opposite, as often reported in the task-switching literature by means of behavioral measures (e.g., Allport et al., 1994; Allport and Wylie, 1999; Yeung and Monsell, 2003).

Although the results of the block analysis have to be considered with caution, it is interesting that the only area to show differential effects of the number of confirmations is, as predicted, the cingulate gyrus. This area shows greater activity in blocks where the number of rejections is higher, consistent with our proposal that confirmation is a less demanding operation than rejection.

These findings, together with those in chapter 3, are strong evidence in favour of the proposal that response mode and rule switching costs share a switching component and, therefore, that switching between confirmation and rejection is equivalent to switching between a dominant and a weaker task.

Chapter 5 – General Discussion

This thesis has presented a body of research that provides convergent evidence for some of the main tenets of the mental model theory of reasoning, namely, the principles of truth and of parsimony. This evidence is convergent in that it is provided by measures and methods seldom employed in reasoning tasks. From a methodological point of view, this thesis has thus proved the adequacy of RT measures and the attention-switching paradigm for the study of the mental representation and operations that mediate human thinking. This marriage of attention methods and reasoning theories has also produced relevant results with respect to the study of attention-switching, thus highlighting the benefits of building “methodological bridges” between the various research fields that comprise cognitive science, in addition to the usual “divide-and-conquer” strategy.

The pattern of switching costs observed in chapter 2 is consistent with the mental model theory’s hypothesis that a richer initial representation is constructed for counterfactual conditionals; with the hypothesis that only for indicative MT inferences is the fleshing out process required. The empirical findings of chapter 3 indicated that people are faster to confirm than to reject that a given triple of numbers follows a previously specified rule; and that confirmation and rejection are indeed different operations, with confirmation being the more dominant one. Chapter 3 also produced relevant results with respect to attention-switching, by showing that object and task switching costs might reflect the functioning of a common switching mechanism and by

providing further evidence of the relation between asymmetrical costs and task dominance. Chapter 4 provided neuroimaging evidence pointing to a common neural basis for object and task switching costs, and of the direct relation between switching asymmetries and brain activation.

Summary of the thesis findings

In chapter 2 we employed the task-switching paradigm in order to provide a more direct test of the nature of the mental representations constructed for factual and counterfactual conditionals. Participants switched between modus ponens (MP) and modus tollens (MT) conditional inferences, and between the subjunctive and indicative moods. An asymmetrical pattern of inference switching costs was observed in all three experiments. In every case, switching costs were observed only for MP inference. We interpreted the results as showing that MP inferences were a stronger task than MT inferences, as they do not require the fleshing out of all of the possibilities consistent with the conditional statement. Another result predicted by the mental model theory is that the advantage for MP over MT inferences was not significant when these inferences were made from counterfactual conditionals (see Byrne and Tasso, 1999).

An asymmetrical pattern was observed as well for mood switching costs, with counterfactuals being in this case the stronger task when compared to factual conditionals, consistent with the proposal that inferences from counterfactuals do not require the fleshing out process. This interpretation was further supported by the interaction between Inference, Mood and Mood Switching, reflecting an asymmetrical pattern of mood switching costs for MT inferences, but a symmetrical one for MP inferences. This result suggests that counterfactual and indicative MP inferences are different tasks, as are counterfactual and indicative MT inferences. However, only for

MT inferences is the mood switching cost asymmetrical, with the cost being only significant for counterfactual conditionals, which fits well with the proposal that MT indicative inferences require the fleshing out process. On the other hand counterfactual MT inferences do not require the fleshing out, consistent with the proposal that the relevant possibility is already represented. This fits well with the principle of parsimony of the mental model theory, but what about the tendency to represent true possibilities? Is it possible that it also reflects a fundamental feature of the cognitive system, as the tendency to represent few possibilities does? These are the questions that motivated the empirical studies presented in the third chapter.

In chapter 3 we presented a series of studies in which participants had to categorize instances as confirming or rejecting a previously specified rule. Over three experiments, with three different rules, and regardless of rule switching being required, all results pointed to an advantage for confirmation over rejection. Furthermore, an asymmetrical cost of switching response mode was observed in experiments 4 and 6, suggesting that confirmation was a stronger task than rejection. Rule switching costs were observed in experiments 5, 6 and 7, consistent with the view that only one rule can be held at a given moment in the focus of attention. In experiment 6, these rule-switching costs showed a pattern of interaction and a correlation with the response mode switching costs consistent with a common switching mechanism that operates simultaneously upon both rules and response modes. This result was, in turn, a further piece of evidence that the response mode switching cost reflects the functioning of a switching process operating between tasks of differing dominance. The results of experiment 7 suggested that a likely cause of the dual-goal manipulation success in improving participants' performance in the 2-4-6 task is that it does not require participants to reject hypotheses.

In chapter 4 we presented an study which employed fMRI to investigate the neural basis of the attention switching costs and the tendency to think about true possibilities, aimed to provide convergent evidence to the results reported in chapter 3. The behavioral results replicated previous findings. Moreover, it was found that rule and response mode switching costs share a common neural basis and that asymmetrical costs were reflected in asymmetrical activation in the frontal lobes. This observation provided further support to the proposal that confirmation and rejection are different tasks, with the former being the more dominant. In the next section we will discuss the implications of all of the empirical findings reported in chapters 2 to 4 for the fields of attention and reasoning.

Implications for reasoning

This thesis has provided convergent evidence for two of the main tenets of the mental model theory of reasoning, the principles of truth and parsimony; and for the hypothesis of this theory regarding the representations and processes involved in making inferences from counterfactual and indicative conditionals. Moreover, some of the results reported are relevant as well to hypothesis-testing research. As previously mentioned, this evidence is convergent in that is based on RT measures and an attention-switching paradigm, which supports our proposal about the utility of building methodological bridges between the different research field that compose modern cognitive psychology.

The experiments presented in chapter 2 were consistent with the proposal that indicative MP inferences involve the representation of a single possibility, counterfactual inferences involve the initial representation of two possibilities, and MT indicative inferences require the fleshing out process in order to represent at least two possibilities. This was inferred from: the observation of asymmetrical mood switching

costs, which suggested that inferences from counterfactual conditionals are a stronger task than inferences from indicative conditionals; from the observation of asymmetrical inference switching costs, which suggested that MP inferences are a stronger task than MT inferences; and finally, from the observation that asymmetrical mood switching costs are only observed for MT inferences, with counterfactual MT inferences being the stronger task. No other theory of reasoning can predict or explain these effects, Evans and Over (2004) criticized the mental model theory account on counterfactual reasoning for proposing that people keep two possibilities in mind. However the observation that counterfactual MT inferences do not require the fleshing out process suggests that the possibility participants need to consider in order to carry the inference is already represented, we will return to this issue in the next section. As noted in chapter 1, two of the main theories of reasoning, mental models and formal rules, agree in the role played by capacity limitations in reasoning. However, only the mental model theory proposes that people tend to focus on what they consider to be true as well. We hypothesized that this tendency could indeed be a basic cognitive phenomenon (as capacity limitations are), and are able to be observed in simpler tasks than those employed in conditional reasoning research. Testing this proposal motivated the set of studies presented in chapters 3 and 4.

An advantage for confirmation over rejection was observed in experiments 4 to 6 in chapter 3, and in the response mode switching condition in the fMRI study presented in chapter 4. Moreover, asymmetrical response mode switching costs, suggesting that confirmation is a stronger operation than rejection, were observed in three different experiments (experiments 4 and 6 in chapter 3, and the response mode switching conditions in chapter 4). Behavioral data from experiment 6 in chapter 3 and the neuroimaging data presented in chapter 4 provided evidence of a common basis for

rule and response mode switching costs. Finally, chapter 4 presented neuroimaging evidence that the asymmetrical cost of switching response mode was directly related to asymmetrical activations in several areas of the frontal lobes, and that blocks with higher proportion of rejections elicit higher cingular activity. These results were taken as evidence of a general cognitive tendency towards positive information, consistent with the tendency to think about true possibilities postulated by the mental model theory. The only other theory of reasoning that may predict a similar pattern of results is Evans's dual-process theory (e.g., Evans, 2003). However, as noted in chapter 3, this theory proposes a pre-attentional selective process, the positivity-bias. The observation that the advantage for confirmations interacted with the difficulty of the rule tested would thus be difficult to explain from this perspective.

Results from experiments 5, 6 and 7 reported in chapter 3 are of relevance to hypothesis-testing research, specifically, with respect to the effect of the dual-goal manipulation of the 2-4-6 task. Results from these experiments indicated that a single representation of two rules is not easily constructed and, therefore, that the success of the dual-goal manipulation cannot be explained by a single mental representation for the two rules (pace Gorman et al., 1987). Results from experiment 7 suggested that the DAX-MED manipulation forces participants to consider all of the triples that do not fit the original rule (DAX) as instances of an explicitly represented alternative rule (MED), thus increasing the search space by forcing participants to search for examples of MEDs. In addition, it was proposed that this manipulation makes all information positive, that is, no rejections are needed, which fits well with our observation that people find confirmation easier than rejection.

The ultimate cause of this advantage for confirming responses, and of the tendency to think about true possibilities in general, remains to be elucidated. It could

be argued that, as people tend to think about true possibilities (Johnson-Laird and Byrne, 2005), these possibilities display a lower threshold of activation, in other words, they would be primed (e.g. Meyer and Schvaneveldt, 1977). In this regard, it is worth referring to a recent study by Prado and Noveck (2006), in which they use RT measures to examine the *matching bias* effect. This effect consists in a decrease in performance when elements mentioned in a rule do not correspond with those in a test item (see Evans, 1998 for a review). These authors (Prado and Noveck, 2006) had their participants verifying conditional rules, finding that the latency of responses was larger in trials where the test item did not match the elements mentioned in the rule. This result resemble the advantage for confirming responses reported in chapters 3 and 4. Rejecting trials could be seen as *mismatches* between the rule and the tested triple. However, it should be noted that the matching bias can be overcome by the use of explicit negations (e.g. Evans and Handley, 1999), as Johnson-Laird and Byrne point out, "*The model theory can rely on the simple unadorned principle of truth*" (2002, p. 669), of which the matching bias would be just another example.

The advantage for confirmations could also be explained in terms of higher processing demands for negations (e.g., Carpenter and Just, 1975; McDonald and Just, 1989; Oaksford, 2002; Prado and Noveck, 2006). But again the cause for these higher demands is a matter of argument. Some consider that negation directs attention away from the negated constituent and towards the contrast class, that is, the set of elements consistent with the negation (e.g., McDonald and Just, 1989; Oaksford, 2002; Schroeyens, Verschueren, Schaeken and d'Ydewalle, 2000; Yama, 2001). On the other hand, there are those who hold a narrower view of negation, considering that negation only denies information (e.g. Evans and Over, 2004). Prado and Noveck (2006) propose an intermediate account in which negation is first narrow and if needed can prompt the

contrast class (see also, Santamaría and Espino, 2000, 2006). These are only tentative explanations, and perhaps they are not mutually exclusive, but the aim of this experiment was to establish that this tendency to think about true possibilities is a basic cognitive phenomenon which can be observed in more basic tasks than those typically employed in thinking paradigms. In this sense, it is worth noting that whereas capacity limitations are a well established phenomenon for cognitive psychology, the ultimate cause for these limitations remains as well a disputed matter (see Meyer and Kieras, 1997a, 1997b for an approach in terms of strategies rather than limitations).

In summary, this thesis has provided convergent evidence for two of the main tenets of the mental model theory of reasoning, the principles of parsimony and truth, by means of latency of responses measures and a task-switching paradigm. Results from chapter 2 support the principle of parsimony and the hypotheses of the mental model theory on the representations and processes involved in the correct performance of inferences from indicative and counterfactual conditionals. Results from chapters 3 and 4 support the principle of truth and suggest that this principle may reflect a basic cognitive phenomenon that can be observed in tasks that are far removed from the complexity of hypothesis testing or conditional reasoning.

Implications for attention

Some of the empirical results reported in this thesis are of direct relevance to the field of attention in general and attention-switching research in particular. Thus, we have shown that different types of attention switching seem to be mediated by a common mechanism with a common neural substrate; we have provided convergent evidence on the relation between switching asymmetries and task-strength; and we have unveiled a similarity effect for object-switching costs.

Switching costs have been observed in each of the 8 experiments reported in this thesis. The robustness of this effect is further demonstrated by the presence of sizeable costs with intervals of various seconds between tasks in experiment 3. As already mentioned, object and task switching costs are interpreted to reflect the fact that only one object or operation can be held at a given moment in WM (e.g., Garavan, 1998; Mayr and Kliegl, 2000; Oberauer, 2002; Rubinstein et al., 2000; but see Cowan, 2001). An question that may arise is, what happens when both the operation and the object of this operation need to be switched? Is there a separate switching process for objects and tasks or a single one? We predicted that a common switching mechanism capable of operating simultaneously different elements should be reflected in a correlation and an underadditive interaction between the correspondent switching costs. A correlation and an underadditive interaction between rule (object) and response mode (task) switching costs was indeed observed in the third of the experiments reported in chapter 3. The interaction reflected that switching both elements was easier than switching each of them separately, which strongly suggested that a common switching mechanism was operating simultaneously over both rules and response modes. This interpretation received further support from neuroimaging results reported in chapter 4. These results pointed to a common neural basis for rule and response mode switching. Moreover, the imaging data indicated that the parietal lobes played a key role in the execution of the switching operation, consistent with a recent meta-analysis by Wager et al. (2004). Therefore, we provide evidence which can be interpreted to reflect the existence of a switching executive process that operates upon different types of content.

Switching asymmetries (e.g. Allport et al., 1994; Allport and Wylie, 1999; Yeung and Monsell, 2003) play an important role in the interpretation of most of our results. In this regard, the argument that one task is more dominant than the other

because they show an asymmetrical pattern of switching costs could be deemed as circular. However, we predicted this pattern for tasks that have never been used for this purpose, based on the hypotheses of the mental model theory of reasoning, which, to this point, has not been concerned with issues such as switching costs or switching asymmetries. Thus, those tasks which the mental model theory identifies as stronger did not show significant effects of switching, as has been reported in the task-switching literature for Stroop switching (Allport, 1994; Allport and Wylie, 1999, 2000), language switching (Meuter and Allport, 1999) and other studies involving switching between tasks differing in strength (De Jong, 1995; Yeung and Monsell, 2003). The asymmetry pattern has been often interpreted as evidence of proactive interference from the previous task set, the so-called “task-set inertia” (e.g., Allport, 1994), as well as episodic priming from the stimuli involved in the tasks (e.g. Waszak, Hommel and Allport, 1999). At first it was argued that the asymmetrical cost was not related to the operation of a switching mechanism (Allport, 1994), but later it was generally accepted that the switch cost reflected both proactive interference and the operation of a switching process (e.g., Mayr and Kliegl, 2000; Meiran, 2000; Rubinstein et al., 2000; Tornay and Milán, 2001). In this regard, we have provided one of the first pieces of evidence relating asymmetrical costs with asymmetrical brain activation. Interestingly, asymmetries were related to activations in the frontal lobes that were not directly implicated in the switching operation, which was rather reflected mainly in parietal activations. This result was interpreted to suggest that asymmetries are a property of the tasks over which the switching process operates, consistent with the view that the switching cost reflects more than the switching operation itself, that is, the cost also reflects interference effects from the tasks and stimuli that are being switched to.

The different rules used in the experiments presented in chapter 3 were classified according to their similarity, and it was found that rule switching costs were smaller when the rules were more similar. This is consistent with an object similarity effect, equivalent to the task-similarity effect reported in the task-switching literature (Arrington et al., 2000; Kleinsorge, 1999). In chapter 2 such a similarity effect was not observed either for mood or inference switching costs, but as mentioned, it could be argued that the manipulations in that experiment favoured the possibility of observing asymmetries rather than similarity effects. In this regard, similar costs were observed for all conditions except indicative MT inferences, whose proposed main difference with respect to the other inferences is precisely the need of engaging in an extra demanding process. It could thus be possible to observe “inference similarity” effects by controlling the demand for extra processing (i.e. using only inferences that do not require the fleshing out process).

In summary, this thesis has provided some novel findings that are directly relevant for research on attention-switching. First, different types of switching seem to be mediated by a common switching process, as revealed both by behavioral and neuroimaging results reported in chapters 3 and 4. In other words, we provide strong evidence of the existence of a switching executive process capable of operating over different types of elements (i.e. objects and tasks). The relation between different types of switching costs is further highlighted by the observation of object-similarity effects, akin to those reported in the task switching literature (experiments 5, 6 and 7, chapter 3). Finally, we have provided one of the first pieces of evidence relating switching asymmetries to asymmetric brain activation. Asymmetries seemed to be a feature of the tasks over which the switching process operated. This latter result is evidence of the complex nature of the switching cost which, as said above, is often interpreted to reflect

more than the switching operation itself. Moreover, we have uncovered a new set of tasks in which task asymmetries can be observed, namely, conditional inferences.

Both the fields of attention and reasoning have thus benefited from the strategy of employing the methods of the former to answer the questions of the latter. Indeed, this strategy has provided useful insights into some the main determinants of human reasoning, and into the functioning of a key executive process. But what do our results tell us about the relation between attention and reasoning?

Attention and Reasoning

The marriage between the methods of attentional research and the mental model theory's hypothesis on reasoning is a fruitful one. In the last two sections we have provided a set of empirical results that are of interest both to the fields of attention and reasoning. However, these results were only relevant to one or the other field. That is, we have shed light on the functioning of the switching process and we have shown how it can be used to investigate reasoning, but we do not provide a direct demonstration of how attention switching mediates reasoning. For example, consider the experiments presented in chapter 2. These experiments showed that an attention-switching task is well suited to test hypotheses on the way people tackle conditional inferences, but we do not show how this mechanism operates in carrying the inference itself. In other words, we have people switching between inferences not within the inference. Despite the latter, the studies reported here constitute a first step towards an interesting line of research that should tackle some further fruitful questions. A question that, we believe, our approach is particularly well suited to answer is: what is meant by saying that people keep more than one possibility in mind?

As said above, Evans and Over (2004) criticised the mental model theory's account of counterfactual reasoning for proposing that people represent more than one possibility at a time. The mental model theory does in fact identify many situations in which people represent initially two, three or even four possibilities (see Johnson-Laird and Byrne, 2002). Moreover there is evidence that both the conjecture and the presupposed fact are primed by a counterfactual conditional (e.g. Fillenbaum, 1974; Galinsky and Moscovitz, 2000; Santamaría, Espino and Byrne, 2005). Again, the results presented in this thesis do not allow a definite answer to the mentioned question, but we may speculate on the answers. The construct of WM makes reference to both storage and processing of information. In this regard, some have argued that the interplay of storage and processing is mediated by selective attention to the elements in WM that must be processed at any given moment (e.g. Engle et al., 1999; Garavan, 1998). As introduced in chapter 1, there are models of WM that distinguish between an activated part of long term memory and a focus of attention, with only the latter being limited in capacity. There is evidence that this capacity is usually restricted to a single object (Garavan, 1998; McElree and Doshier, 1989), which is consistent with the rule switching costs reported in this thesis. Some models, like the one proposed by Oberauer (2002) further assume a third functionally distinct region in WM, a region of direct access that holds information that has to be available in an ongoing cognitive process, but that is not in the focus of attention. Therefore, models that assume distinct functional regions within WM could easily accommodate situations in which various possibilities have to be represented. Results reported in chapter 2 suggested that no fleshing out is needed in order to carry counterfactual MT inferences, consistent with the alternative possibility being already represented. It could thus be argued that people represent various possibilities simply by assuming that they keep one in the focus of

attention and the other ones on an activated state allowing direct access to the focus, this proposal leads to many interesting research questions. In the introductory chapter, various approaches to the study of the relation between WM and reasoning were introduced. The main conclusion was that reasoning is mainly related to the activity of the central executive, which is understood as a general pool of resources for controlled processing. However, inasmuch as models of WM differ in their understanding of the interplay between storage and processing, theories of reasoning should commit to specific models of WM when hypothesizing about the processes and representations involved in solving reasoning problems. In this regard, the methodological approach put forward in this thesis might prove adequate to investigate which particular model of WM fits better with the hypotheses of the mental model theory. Reasoning and WM are not the same construct, but it is evident that our understanding of the former depends critically on how we conceptualize the latter.

Capacity limitations appear thus to be key in explaining reasoning performance, as stated by the mental model theory's principle of parsimony. But in this thesis we have also provided evidence of a basic tendency to consider true possibilities. The interplay between both principles is also an interesting matter for further research. A result reported in chapter 2 provides us the opportunity to speculate on this question. Indicative and counterfactual MP inferences differ in only one aspect, the former entails the representation of a single possibility and the latter entails the representation of two possibilities. The fact that a symmetrical cost is observed when switching between these two inferences suggests that people are equally practiced in representing one or two possibilities. In light of models of WM that assume a further focus of attention, it could be argued that it is not inherently harder to represent a second possibility, but we tend to keep only a true one on the focus of attention. This focusing on true possibilities is not a

problem if the alternative possibility needed to solve an inference is already in WM (i.e. counterfactual conditionals), but if a search for alternatives is required, further processing resources are needed. The hypothesized tendency to consider true possibilities could thus be another key explanatory factor of reasoning performance, and cognition in general and, perhaps, not as tightly related to capacity limitations as previously supposed. This interpretation is highly speculative, but it leads to some interesting research questions, some of which motivated the studies chapter 3, which further supported the view of a basic cognitive tendency towards positive information.

Concluding comments

This thesis has introduced a somewhat novel approach to the study of attention and reasoning that is critically based on the assumptions that permeate both research fields, for example, the existence of “switching mechanisms”, or the very existence of mental representations. Despite the latter, this thesis: provides strong convergent evidence that the mental model theory’s principles of parsimony and truth reflect basic cognitive phenomena; demonstrates the close relationship between different types of attention switching by showing that object and task switching costs by are operated by a common switching mechanism; unveils a direct relation between asymmetric costs and brain activation and an object similarity effect for object switching; and finally, highlights the fact that hypotheses about the representations and processes involved in reasoning should consider specific models of WM.

As noted in the introductory chapter, this thesis has tried to build a bridge between the fields of attention and reasoning by showing how the methods of the former can help to answer the questions posed by the latter. We believe that this effort was a worthwhile one, not only in that it provided convergent evidence to some of the

proposals put forward in both fields, but also in producing further interesting questions for future research. Indeed, a task as simple as the one used in chapters 3 and 4 has produced a fruitful set of results. It is safe to assume that building bridges between other disciplines of cognitive science, in addition to the most usual “*divide and conquer strategy*”, should provide many interesting answers to our questions about human cognition. The present work has only scratched the surface.

References

- Aaron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *Journal of Neuroscience*, 26(9), 2424-2433.
- Ackerman, P. L., Beier, M. E. & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, 131 (1), 30-60.
- Allport, A. and Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. M. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience* (pp. 273-296). Oxford, England: Oxford University Press.
- Allport, A. And Wylie, G. (2000). "Task-switching", stimulus-response bindings, and negative priming. In: S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.
- Allport, A., Styles, E. A. & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In: Umiltá, C. y Moscovitch, M. (eds.). *Conscious and unconscious information processing. Attention and Performance XV*. Cambridge, MA: MIT Press.
- Arrington, C. M., Altmann, E. M., & Carr, T. H. (2003). Tasks of a feather flock together: Similarity effects in task switching. *Memory & Cognition*, 31(5), 781-789.
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

- Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49a (1), 5-28.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4 (11), 417-423.
- Baddeley, A. (2002). Is working memory still working? *American Psychologist*, 56 (11), 851-881.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In: Bower, G. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. VIII (pp. 47-90). New York: Academic Press.
- Baddeley, A. D. & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). NY: Cambridge University Press.
- Badgaiyan, R. D., & Posner, M. I. (1998). Mapping the cingulate cortex in response selection and monitoring. *Neuroimage*, 7, 255-260.
- Barrouillet, P., Grosset, N. & Lecas, J. -F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition*, 75, 237-266.
- Barrouillet, P. & Lecas, J. -F. (1999). Mental models in conditional reasoning and working memory. *Thinking and Reasoning*, 5(4), 289-302.
- Braine, M. D. S. & O'Brien, D. P. (eds.) (1998). *Mental Logic*. London: L. Erlbaum.
- Braine, M. D. S. & O'Brien, D. P. (1991). A theory of If : A lexical entry reasoning program and pragmatic principles. *Psychological Review*, 85, 1-21.
- Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In: Medin, D. (ed.). *The Psychology of Learning and Motivation, Advances in Research and Theory*. Academic Press.

- Byrne, R. M. J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, 6(10), 426-431.
- Byrne, R. M. J. (2005). *The Rational Imagination*. Cambridge MA: MIT Press.
- Byrne, R. M. J. & Tasso, A. (1999). Deductive reasoning with factual, possible and counterfactual conditionals. *Memory and Cognition*, 27, 726-740.
- Bunge, S. A., Klingberg, T., Jacobsen, R. B., & Gabrieli, J. D. E. (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 3573-3578.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215-222.
- Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking and Reasoning*, 9(3), 203-244.
- Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 82(1), 45-73.
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. (1998). Anterior cingulate cortex, error detection and the online monitoring of performance. *Science*, 280(5364), 2331-2340.
- Cohen, J. D., Aston-Jones, G., & Gilzerant, M. S. (2004). A systems-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring and exploitation versus exploration. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention*. New York: Guilford Press.
- Collette, F., Salmon, E., van der Linden, M., Chicherio, C. Belleville, S., Degueldre, C., Delfiore, G., & Franck, G. (1999). Regional brain activity during tasks devoted to the central executive of working memory. *Cognitive Brain Research*, 7(3), 411-417.

- Conway, A. R. A. & Engle, R. W. (1994). Working-memory capacity as long-term memory activation: a resource-dependent inhibition model. *Journal of Experimental Psychology: General*, 123, 354-373.
- Conway, A. R. A., Kane, M. J. & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547-551.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.
- Corbetta, M., Miezin, F. M., Shulman, G. L. & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience*, 13(3), 1202-1226.
- Cowan, N. (1995). *Attention and Memory: An integrated framework*. New York: Oxford University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computer and Biomedical Research*, 29(3), 162-173.
- Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.
- D'Esposito, M., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., & Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research*, 7, 1-13.
- De Jong, R. (1995). Strategic determinants of compatibility effects with task uncertainty. *Acta Psychologica*, 88, 332-361.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error-detection and compensation. *Psychological Science*, 5(5), 303-305.

- Dove, A., Pollman, S., Schubert, T., Wiggins, C. J. & von Cramon, Y. D. (2000). Prefrontal cortex activation in task switching: and event-related fMRI study. *Cognitive Brain Research*, 9, 103-109.
- Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral flow during emotional versus higher cognitive processes: implications for interactions between emotion and cognition. *Cognition and Emotion*, 12, 353-385.
- Duncan, J. & Owen, A. M. (2000). Common regions of the frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23, 475-483.
- Engle, R. W. & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 145-199). NY: Academic Press.
- Engle, R. W., Kane, M. J. & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134).
- Engle, R. W., Tuholski, S. W., Laughlin, J. E. & Conway, A. R. A. (1999). Working memory, short term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128, 309-331.
- Evans, J. St. B. T. (1989). *Biases in Human Reasoning: Causes and Consequences*. Hove: Lawrence Erlbaum Associates Ltd.
- Evans, J. St. B. T. (1993). The cognitive psychology of reasoning: An introduction. *Quarterly Journal of Experimental Psychology*, 46A (4), 561-567.

- Evans, J. St. B. T. (1998). Matching bias in conditional reasoning. Do we understand it after 25 years? *Thinking and Reasoning*, 4, 45-110.
- Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454-458.
- Evans, J. St. B. T., & Handley, S. J. (1999). The role of negation in conditional inference. *Quarterly Journal of Experimental Psychology*, 52, 739-770.
- Evans, J. St. B. T. & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Eysenck, M W. & Keane, M T. (2005). *Cognitive psychology: a student's handbook*. Hove: Psychology Press.
- Fangmeier, Knauff, Ruff and Sloutsky (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, 18(3), 320-334.
- Fiddick, L., Cosmides, L. & Tooby, J. (2000). No interpretation without representation: the role of domain specific representations and inferences in the Wason selection task. *Cognition*, 77, 1-79.
- Fillenbaum, S. (1974). Information amplified: memory for counterfactual conditionals. *Journal of Experimental Psychology*, 102, 44-49.
- Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain*, 120(10), 1779-1791.
- Fodor, J. A. (1983) *The modularity of mind: an essay on faculty psychology*. MIT Press.
- Gale, M & Ball, M. J. (2002). Does positivity bias explain the patterns of performance on Wason's selection task? *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*.

- Gale, M. & Ball, M. J. (2003). Facilitated rule discovery in Wason's 2-4-6 task: The role of negative triples. *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*.
- Galinsky, A. D & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation of heuristics and consideration of alternatives. *Journal of Experimental Social Psychology*, 36, 384-409.
- Garavan, H. (1998). Serial attention within working memory. *Memory and Cognition*, 26, 263-276.
- Garavan, H., Ross, T. J., Li, S. J. & Stein, E. A. (2000). A parametric manipulation of central executive functioning. *Cerebral Cortex*, 10, 585-592.
- García-Madruga, J.A., Gutiérrez, F, Carriedo, N., Moreno, S. and Johnson-Laird, P.N. (2002). Mental models in deductive reasoning. *The Spanish Journal of Psychology*, 5, 125-140.
- García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Luzón, J. M. & Vila, J. O. (2004). Working memory and propositional reasoning: Searching for new working memory tests.
- Gilhooly, K.J., (1998). Working memory, strategies, and reasoning task. In R.H. Logie, and K.J., Gilhooly. *Working Memory and Thinking*. (23-45). Hove, East Sussex: Psychology Press.
- Gilhooly, K.J., Logie, R.H., Wetherick, N. E. and Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Journal of Memory and Cognition*, 21(1), 115-124.
- Gilhooly, K. J. (1998). Working memory, strategies and reasoning tasks. In: R. H. Logie & K. J. Gilhooly (Eds.). *Working memory and Thinking*. Psychology Press.

- Gilhooly, K.J., Logie, R.H., and Wynn, V. (1999). Syllogistic reasoning tasks and working memory: evidence from sequential presentation of premises. *European Journal of Cognitive Psychology*, 11 (4), 473-498.
- Giroto, V., Mazzocco, A & Tasso, A. (1997). The effect of premise order in conditional reasoning: a test of the mental model theory. *Cognition*, 63, 1-28.
- Glenberg, A. M. (1997). What memory is for? *Behavioural & Brain Science*, 20, 1-55.
- Glenberg, A. M., Meyer, M. & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory & Learning*, 26, 69-83.
- González, A., Milán, E. G., Pereda, A. and Hochel, M. (2005). The response-cued completion hypothesis and the nature of residual cost in regular shift. *Acta Psychologica*, 120, 327-341.
- Gorman, M. E., Stafford, A. & Gorman, M. E. (1987). Disconfirmation and dual hypotheses on a more difficult version of Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology*, 39A, 1-28.
- Goshke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In: S. Monsell and J. Driver (Eds). *Control of cognitive processes: Attention and performance XVIII*. Cambridge MA: MIT Press..
- Hitch, G. J. & Baddeley, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, 28, 603-621.
- James, W. (1890). *The Principles of Psychology*. New York: Dover Publications.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.

- Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction*. Hove: Lawrence Erlbaum Associates Ltd.
- Johnson-Laird, P. N. (2000). Deductive Reasoning. *Annual Review of Psychology*, 50, 109-135.
- Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics and inference. *Psychological Review*, 109(4), 646-678.
- Johnson-Laird, P. N., Byrne, R. M. J. & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99(3), 418-439.
- Johnson-Laird, P. N., Byrne, R. M. J. & Schaeken, W. (1994). Why models rather than rules give a better account of propositional reasoning: A reply to Bonatti and to O'Brien, Braine and Yang. *Psychological Review*, 101(4), 734-739.
- Jonides, J., Schumacher, E. H., Smith, E. E., Koeppe, R. A., Awh, E., Reuter-Lorenz, P. A., Marshuetz, C., & Willis, C. R. (1998). The role of parietal cortex in verbal working memory. *Journal of Neuroscience*, 18, 5026-5034.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kane, M. J. & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal-neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47-70.
- Kane, M. J., Hambrick, D. Z. & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier and Boyle (2005). *Psychological Bulletin*, 131 (1), 66-71.
- Kane, M. J., Poole, B. J., Tuholski, S. W. & Engle, R. W. (2003, November). *Working memory capacity and executive control in search and switching*. Paper presented

- at the annual meeting of the Psychonomic Society, Vancouver, British Columbia, Canada.
- Kareev, Y. & Halberstadt, N. (1993). Evaluating negative tests and refutations in a rule discovery task. *Quarterly Journal of Experimental Psychology*, 46A (4), 715-727.
- Kimberg, D. Y., Aguirre, G. K., & D'Esposito, M. (2000). Modulation of task related neural activity in task switching: An fMRI study. *Cognitive Brain Research*, 10, 189-196.
- Klauer, K.C., Stegmaier, R., and Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and reasoning*, 3, 9-48.
- Klayman, J. & Ha, Y-W. (1987). Confirmation, disconfirmation and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Klayman, J. & Ha, Y-W. (1989). Hypothesis testing in rule discovery : Strategy, structure and content. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 596-604.
- Kleinsorge, T. (1999). Hierarchical switching in a multi-dimensional task space. *Psychological Research*, 62(4), 300-312.
- Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K., & Miyashita, Y. (1998a). Transient activation of inferior frontal cortex during cognitive set shifting. *Nature Neuroscience*, 1, 80-84.
- Konishi, S., Nakajima, K., Uchida, I., Sekihara, & Miyashita, Y. (1998b). No-go dominant brain activity in human inferior prefrontal cortex revealed by fMRI. *European Journal of Neuroscience*, 10(3), 1209-1213.
- Kyllonen, P. C. & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14, 389-433.

- Le, T. H., Pardo, J. V., & Hu, X. (1998). 4 T-fMRI study of nonspatial shifting of selective attention: parietal and cerebellar contributions. *Journal of Neurophysiology*, 79(3), 1535-1548.
- Loose, R., Kaufmann, C., Tucha, O., Auer, D. P. & Lange, K. W. (2006). Neural network of response shifting: Influence of task speed and stimulus material. *Brain Research*, 1090, 146-155.
- Legrenzi, P., Girotto, V. & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision making. *Cognition*, 49 (1-2), 37-66.
- Logie, R. H. and Gilhooly, K. J. (1998). *Working Memory and Thinking*. Hove, East Sussex: Psychology Press.
- Luks, T. L., Simpson, G. V., Feiwell, R. J., & Miller, W. J. (2002). Evidence for anterior cingulate cortex involvement in monitoring preparatory attentional set. *Neuroimage*, 17(2), 792-802.
- Markovits, H., Doyon, C. and Simoneau, M. (2002). Individual differences in working memory and conditional reasoning with concrete and abstract content. *Thinking and reasoning*, 8, 97-107.
- Mayr, U. & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(5), 1124-1140.
- McDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(4), 633-642.
- McDonald, E. W., Cohen, J. D., Stenger, A. V., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal cortex and anterior cingulate in cognitive control. *Science*, 288(9), 1835-1838.

- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory. The time course of recognition. *Journal of Experimental Psychology: General*, 118, 251-269.
- Meuter, R. F. I., & Allport, D. A. (1999). Bilingual language switching in naming : Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25-40.
- Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 22-234.
- Meyer, D. E. y Kieras, D. E. (1997a). EPIC. A computational theory of executive cognitive processes and multiple task performance. Part 1. Basic Mechanisms. *Psychological Review*, 104, 3-65.
- Meyer, D. E. y Kieras, D. E. (1997b). EPIC. A computational theory of executive cognitive processes and multiple task performance. Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, 104, 749-791.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Monsell, S. (2003). Task Switching. *Trends in Cognitive Neurosciences*, 7(3), 134-140.
- Moreno-Ríos, S. & García-Madruga, J. (2002). Priming in deduction: A spatial arrangement task. *Memory & Cognition*, 30 (7), 1118-1127.

- Murphy, K. And Garavan, H. (2005). Deriving the optimal number of events for an event-related fMRI study based on the spatial extent of activation. *Neuroimage*, 27, 771-777.
- Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Sawamoto, N., Toma, K., Nakamura, K., Hanakawa, T., Konishi, J., Fukuyama H., & Shibakasi, H. (1999). Transient neural activity in the medial superior frontal gyrus and precuneus time locked with attention shift between object features. *Neuroimage*, 10, 193-199.
- Oaksford, M. R. (2002). Contrasts classes and matching bias as explanations of the effects of negation on conditional reasoning. *Thinking and Reasoning*, 8, 135-151.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(3), 411-421.
- Oberauer, K., Süß, H., Schulze, R., Wilhelm, O. & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision and coordination. *Intelligence*, 31, 167-193.
- Oberauer, K., Schulze, R., Wilhelm, O & Süß, H. (2005). Working memory and intelligence-Their correlation and their relation: Comment on Ackerman, Beier and Boyle (2005). *Psychological Bulletin*, 131 (1), 61-65.
- Owen, A. M., Evans, A. C., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cerebral Cortex*, 6, 31-38.
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362(6418), 342-345.

- Pollard, P. (1982). Human reasoning: Some possible effects of availability. *Cognition*, 12, 65-69.
- Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1994). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*, 4(7), 1863-1874.
- Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's progressive matrices test. *Cognitive Psychology*, 33, 43-63.
- Prado, J., & Noveck, I. A. (2006). How reaction time measures elucidate the matching bias and the way negations are processed. *Thinking and Reasoning*, 12(3), 309-328.
- Pruessman, K. P., Weiger, M., Scheidegger, M. B., & Boesiger, P. *Magnetic Resonance Medicine*, 42, 952..
- Quelhas, A. C. & Byrne, R. M. J. (in press). Reasoning with deontic and counterfactual conditionals. *Thinking and Reasoning*.
- Rips, L. J. (1994). *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: MIT Press.
- Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90, 38-71.
- Rubinstein, J. S., Meyer, D. E. & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27 (4), 763-797.

- Ruff, C. C., Knauff, M., Fangmeier, T. & Spreer, J. (2003). Reasoning and working memory: common and distinct neuronal processes. *Neuropsychologia*, 41, 1241-1253.
- Santamaría, C., Espino, O. & Byrne, R. M. J. (2005). Counterfactual and semifactual conditionals prime the consideration of alternatives. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31(5), 1149-1154.
- Santamaría, C. & Espino, O. (2000). Truth and falsity in propositional reasoning: the negation heuristic. In: J. A. García-Madruga, N. Carriedo and M. J. González Labra. *Mental Models in reasoning*. Madrid. UNED.
- Santamaría, C., & Espino, O. (2006). Thinking the truth and selecting the false. *Psicológica*, 27, 195-206.
- Schaeken, W. S., Johnson-Laird, P. N. & d'Ydewalle, G. (1996). Tense, aspect and temporal reasoning. *Thinking and Reasoning*, 2, 309-327.
- Schneider, W., Eschman, A. & Zucolotto, A. (2002). *E-Prime User's Guide*. Psychology Software Tools Inc., Learning Research and Development Centre, University of Pittsburgh.
- Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.
- Schroeyens, W., Schaeken, W., Verschueren, N. & D'Ydewalle (2000). Conditional reasoning with negations: Implicit and explicit affirmation or denial and the role of contrast classes. *Thinking and Reasoning*, 6, 221-252.
- Shafritz, K. M., Kartheiser, P., & Belger, A. (2005). Dissociation of cognitive systems mediating shifts in behavioral response and cognitive set. *Neuroimage*, 25, 600-606.

References

- Shah, P. & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125(1), 4-27.
- Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190.
- Smith, E. E., Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283, 1657-1661.
- Sohn, M. H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). Inaugural article: the role of prefrontal cortex and posterior parietal cortex in task switching. *Proceedings of the National Academy of Sciences U. S. A.*, 97(24), 13448-13453.
- Sperber, D., Cara, F. & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 52, 3-39.
- Sperber, D. & Wilson, D. (1995). *Relevance: communication and cognition* (Rev. ed.). Oxford, England: Blackwell.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders method. *Acta Psychologica*, 30, 276-315.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernández, L., Nichols, T. Smith, E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, 41(3), 357-370.

- Talairach, T. & Tournoux, P. (1998). *Co-planar Stereotaxic Atlas of the Human Brain*. New York: Thieme.
- Thompson, V. A. & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(6), 1154-1170.
- Toms, M., Morris, N. & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology*, 46A (4), 679-699.
- Tornay, F. J., and Milan, E. G. (2001). A more complete task-set reconfiguration in random than in predictable task switch. *The Quarterly Journal of Experimental Psychology A*, 54, 785-803.
- Tukey, D. D. (1986). A philosophical and empirical analysis of subject's modes of inquiry in Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology*, 38A, 5-33.
- Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R. Gross, K. A. & Arrkelin, D. L. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology*, 32, 109-123.
- van Venn, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage*, 14, 1302-1308.
- Wager, T. D., Jonides, J. & Reading, S. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *Neuroimage*, 22, 1679-1693.
- Wallis, J. D., Anderson, J. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840), 953-956.

- Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural basis of asymmetrical language switching in second-language learners: An ER-fMRI study. *Neuroimage*, 35, 862-870.
- Ward, B., Garavan, H., Ross, T. J., Bloom, A., Cox, R. W., and Stein, E. A. (1998). Nonlinear regression for fMRI time series analysis. *Neuroimage*, 7, 57-67.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.
- Waszak, F., Hommel, B. y Allport, A. Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, 46, 361-413.
- Yama, H. (2001). Matching versus optimal data selection in the Wason selection task. *Thinking and Reasoning*, 7, 295-311.
- Yeung, N. & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 455-469.

Appendix Instructions for the experiments presented in chapter 2

Instructions for Experiment 1

This is a task of everyday reasoning. It is not a test of intelligence, but there are wrong and right questions so please think carefully about your answers.

You will be reminded of this sentence at the beginning of each block:

“If there is an vowel letter, there is an even number”

after that you will be given a series of sentences like:

“A”

your task is to judge: what if anything follows? And you will do that by evaluating the conclusion that will be given afterwards:

•
• • (the three dots mean “therefore”)
“2”

If you judge that the conclusion must be true, that is, it is a *valid conclusion*, press the “valid” key. But if you judge that the conclusion need not be true, that is, it is an *invalid conclusion*, press the “invalid” key.

For this example, given “If vowel then even” and “A”, the conclusion “2” must be true, so the response must be “valid”. However if the conclusion had been “3”, the response should have been “invalid”.

Instructions for Experiment 2 and 3

This is a task of everyday reasoning. It is not a test of intelligence, but there are wrong and right questions so please think carefully about your answers.

You will be given sentences like the this:

“If there was an A then there was a C”

after that you will be given:

“A”

your task is to judge: what if anything follows? And you will do that by evaluating the conclusion that will be given afterwards:

.
 . . (the three dots mean “therefore”)
“C”

If you judge that the conclusion must be true, that is, it is a *valid conclusion*, press the “valid” key. But if you judge that the conclusion need not be true, that is, it is an *invalid conclusion*, press the “invalid” key.

For this example, given “If A then C” and “A”, the conclusion “C” must be true, so the response must be “valid”. However if the conclusion had been “not C”, the response should have been “invalid”.