

The development of reasoning heuristics in autism and  
in typical development.

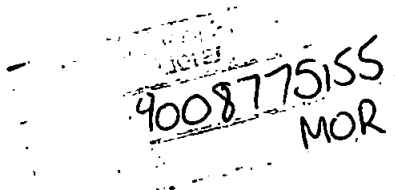
*Kinga Ella Morsanyi*



**Doctor of Philosophy**

**School of Psychology**

**2010**



## **Abstract**

**Kinga Ella Morsanyi: The development of reasoning heuristics in autism and in typical development.**

Reasoning and judgment under uncertainty are often based on a limited number of simplifying heuristics rather than formal logic or rule-based argumentation. Heuristics are low-effort mental shortcuts, which save time and effort, and usually result in accurate judgment, but they can also lead to systematic errors and biases when applied inappropriately. In the past 40 years hundreds of papers have been published on the topic of heuristics and biases in judgment and decision making. However, we still know surprisingly little about the development and the cognitive underpinnings of heuristics and biases.

The main aim of my thesis is to examine these questions. Another aim is to evaluate the applicability of dual-process theories of reasoning to the development of reasoning. Dual-process theories claim that there are two types of process underlying higher order reasoning: fast, automatic, and effortless (Type 1) processes (which are usually associated with the use of reasoning heuristics), and slow, conscious and effortful (Type 2) processes (which are usually associated with rule-based reasoning).

This thesis presents eight experiments which investigated the development of reasoning heuristics in three different populations: typically developing children and adolescents between the age of 5 and 16, adolescents with autism, and university students. Although heuristic reasoning is supposed to be basic, simple, and effortless, we have found evidence that responses that are usually attributed to heuristic processes are positively correlated with cognitive capacity in the case of young children (even after controlling for the effects of age). Moreover, we have found that adolescents with autism are less susceptible to a number of reasoning heuristics than typically developing children. Finally,

our experiments with university students provided evidence that education in statistics increases the likelihood of the inappropriate use of a certain heuristic (the equiprobability bias). These results offer a novel insight into the development of reasoning heuristics. Additionally, they have interesting implications for dual-process theories of reasoning, and they can also inform the debates about the rationality of reasoning heuristics and biases.

## **Contents**

<b>Chapter 1. Introduction to the heuristics and biases research program and dual-process theories of reasoning.</b>	<b>1</b>
1.0. Introduction	1
1.1. What are heuristics?	2
1.1.1. An introduction to the heuristics and biases research program.	2
1.1.2. Where do heuristics come from?	9
1.2. Dual-process theories of reasoning	15
1.2.1. Distinguishing between two types of reasoning processes	15
1.2.2. Evidence for the existence of Type 1 and Type 2 processes	19
1.3. New additions to dual-process theories and alternative approaches	22
1.3.1. The effect of knowledge on reasoning, and the revised heuristic-analytic theory	22
1.3.2. How many systems, and how do they differ?	25
1.3.3. Gigerenzer's critique of dual-process theories	27
1.4. Questions to be answered and the outline of the chapters to follow.	30
<b>Chapter 2. The development of reasoning skills.</b>	
2.0. Introduction	35
2.1. Developmental dual-process accounts, and the expected developmental trajectories of reasoning performance	36
2.1.1. Research into the development of children's reasoning using heuristics and biases tasks	36
2.1.2. Developmental dual-process theories	43
2.1.3. The links between the development of memory and reasoning: The fuzzy-trace theory.	48
2.1.4. An alternative (single process) approach to the development of reasoning heuristics: the intuitive rules research program	51
2.2. The role of memory retrieval and contextualisation in logical reasoning: The case of conditional inferences	57
2.2.1. The role of the activation and integration of real-life knowledge in everyday conditional reasoning	57
2.2.2. Developmental research on conditional reasoning.	60
2.3. Research on the development of children's reasoning abilities - Summary	64
<b>Chapter 3. The development of heuristic reasoning in childhood and adolescence (Experiments 1-3.)</b>	<b>68</b>
3.0. Introduction to Chapter 3.	68

3.1.	Experiment 1: Changes in reasoning performance on some typical heuristics and biases tasks between the age of 5 and 11.	69
3.2.	Experiment 2: The development of heuristic reasoning in adolescence and its relationship to executive functioning	88
3.3.	Experiment 3. The development of heuristic reasoning in adolescence, and its relationship with analogical reasoning ability.	103
3.4.	General discussion (Experiments 1-3.)	127
<b>Chapter 4. Cognitive theories of autism, and research on reasoning in autism.</b>		131
4.0.	Introduction to Chapter 4.	131
4.1.	Autism and the cognitive theories of autism	131
4.1.1.	The theory of mind deficit hypothesis of autism	132
4.1.2.	The weak central coherence theory	136
4.1.3.	The executive dysfunction theory of autism	139
4.1.4.	The extreme male brain theory of autism	142
4.2.	Reasoning and autism	146
4.2.1.	Analogical reasoning in autism	146
4.2.2.	Memory retrieval and everyday conditional reasoning in autism	148
4.2.3.	Heuristics and biases in autism	152
4.3.	Discussion	156
<b>Chapter 5. Heuristic reasoning in autism (Experiments 4 and 5)</b>		159
5.0.	Introduction to Chapter 5	159
5.1.	Experiment 4: The development of heuristic reasoning in autism and its relationship to cognitive abilities and executive functioning.	162
5.2.	Experiment 5: Susceptibility to heuristics, and analogical reasoning ability in autism.	172
5.3.	General discussion	197
<b>Chapter 6: The role of education and knowledge in heuristic reasoning (Experiments 6-8)</b>		206
6.0.	Introduction to Chapter 6.	206
6.1.	Heuristics in probabilistic reasoning and their relationship with cognitive ability, thinking styles and educational background.	208
6.2.	Experiment 6. The role of education, discipline, cognitive abilities and cognitive styles in heuristic reasoning.	217

6.3.	Experiment 7. The replication of Experiment 6 with Italian psychology students.	227
6.4.	Experiment 8. The effect of instructions on susceptibility to the representativeness heuristic and to the equiprobability bias.	233
6.5.	General discussion	240
Chapter 7: A review and interpretation of the experimental findings, and recommendations for future research.		245
7.0.	Introduction to Chapter 7.	245
7.1.	Experimental findings regarding the development of heuristics and biases in autism and in typical development, and the role of education in heuristic reasoning.	246
7.1.1.	A summary of the predictions of dual-process theories regarding the development of heuristic reasoning.	246
7.1.2.	Contrasting our experimental findings with the predictions of dual-process theories.	250
7.1.3.	Our findings regarding heuristics and biases, and executive functioning in autism	255
7.1.4.	The role of knowledge and thinking dispositions in reasoning.	258
7.2.	Are two systems better than one?	262
7.2.1.	Are the results of our studies in line with the intuitive rules approach?	262
7.2.2.	Some theoretical issues regarding dual-process theories of reasoning.	266
7.3.	Concluding comments and possible future directions	270
7.3.1.	Is “rational thinking” more rational than heuristics and biases?	270
7.3.2.	Possible future lines of research and conclusions	273
References		277
Appendices		
	APPENDIX A: Task used in Experiments 1, 2, and 4.	310
	APPENDIX B: Judgment tasks used in Experiments 3 and 5.	315
	APPENDIX C: Syllogistic reasoning problems used in Experiment 3	328
	APPENDIX D: Instructions for the analogical reasoning problems	333
	APPENDIX E: The tasks administered in Experiments 6-8.	335
	APPENDIX S1: Supplementary tables for Experiment 1.	341
	APPENDIX S2: Supplementary tables for Experiment 2.	342

<b>APPENDIX S3: Supplementary tables for Experiment 3.</b>	<b>345</b>
<b>APPENDIX S4: Supplementary tables for Experiment 4</b>	<b>347</b>
<b>APPENDIX S5: Supplementary tables for Experiment 5.</b>	<b>350</b>
<b>APPENDIX S6: Supplementary tables for Experiment 6.</b>	<b>355</b>
<b>APPENDIX S7: Supplementary tables for Experiment 7.</b>	<b>357</b>
<b>APPENDIX S8: Supplementary tables for Experiment 8.</b>	<b>359</b>

## Tables

### Chapter 2

- |      |                                                                                                                              |    |
|------|------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1. | Stanovich et al.'s (2008) taxonomy of heuristics and biases.                                                                 | 45 |
| 2.2. | Developmental trajectories predicted by developmental dual-process theories, and the cognitive foundations of these changes. | 66 |

### Chapter 3

- |       |                                                                                                                                                                                         |     |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1.  | The mean proportion of heuristic, analytic and "other" responses on the different tasks across the three age groups.                                                                    | 81  |
| 3.2.  | Correlations between the individual differences measures                                                                                                                                | 82  |
| 3.3.  | Correlations between age, cognitive ability and response types across tasks.                                                                                                            | 84  |
| 3.4.  | Intercorrelations between heuristic responses on the reasoning and decision-making tasks.                                                                                               | 85  |
| 3.5.  | Proportion of heuristic, analytic and "other" responses on the different types of reasoning tasks across age groups.                                                                    | 94  |
| 3.6.  | Correlations between age and the cognitive ability measures.                                                                                                                            | 96  |
| 3.7.  | Correlations between the individual differences measures and heuristic responding on the different tasks.                                                                               | 97  |
| 3.8.  | The proportion of participants choosing each type of response across age groups and conflict conditions on the engineers and lawyers problem with a representative description.         | 115 |
| 3.9.  | The proportion of children choosing each type of response across age groups and conflict/non-conflict tasks on the engineers and lawyers problem with a non-representative description. | 116 |
| 3.10. | The proportion of participants giving normative/non-normative responses on the positive/negative recency problems across age groups and conflict/non-conflict tasks.                    | 117 |
| 3.11. | The proportion of participants choosing each type of response on the conjunction fallacy problems across age groups and conflict/non-conflict tasks.                                    | 117 |
| 3.12. | The proportion of belief-based and logical responses across problem types and age groups on the syllogistic reasoning problems.                                                         | 118 |
| 3.13. | The proportion of correct responses on the distractor and non-distractor picture analogy problems across age groups.                                                                    | 119 |
| 3.14. | The proportion of different types of errors on the picture analogy problems.                                                                                                            | 119 |
| 3.15. | The proportion of correct responses on the scene analogy problems across age groups, distraction conditions and relational complexity.                                                  | 120 |



3.16.	The proportion of perceptual, relational and other errors on the scene analogy problems across age groups.	121
3.17.	Correlations between the different measures of analogical reasoning.	121
3.18.	Correlations between age, working memory, general intelligence, and analogical reasoning ability.	122
3.19.	Correlations between susceptibility to different biases.	123
3.20.	Correlations between susceptibility to different biases and age and the measures of cognitive capacity.	124
<b>Chapter 5</b>		
5.1.	Expectations about heuristic use in autism based on some prominent theoretical approaches.	163
5.2.	Comparisons between the autistic and the control group on the measures of cognitive abilities and executive functions.	165
5.3.	Proportion of heuristic, normative and “other” responses on the different types of reasoning and judgment tasks across participant groups (SDs in parentheses)	167
5.4.	Correlations between the cognitive ability measures for the autistic sample (the results for the typically developing sample in Experiment 2 are displayed in brackets).	168
5.5.	Correlations between the individual differences measures and heuristic responding on the different tasks in the autistic group (results for the typically developing sample from Experiment 2 are displayed in brackets).	169
5.6.	Comparisons between the autistic and the control group on the measures of cognitive capacity and executive functioning.	178
5.7.	The proportion of participants choosing each type of response across groups, content and conflict conditions on the engineers and lawyers problem with a representative description.	181
5.8.	The proportion of children choosing each type of response across groups, content and conflict/non-conflict tasks on the engineers and lawyers problem with a non-representative description.	182
5.9.	The proportion of participants choosing each type of response on the positive/negative recency problems across groups, content and conflict/non-conflict tasks.	183
5.10.	The proportion of participants choosing each type of response on the conjunction fallacy problems across groups, content and conflict/non-conflict tasks.	184
5.11.	Proportion of correct responses on the picture analogy tasks.	185
5.12.	Proportion of perceptual, semantic and other errors across groups on the picture analogy tasks.	186
5.13.	Proportion of correct responses on the scene analogy tasks.	187
5.14.	Proportion of perceptual, relational and other errors across groups on the scene analogy tasks.	187
5.15.	Correlations between age and the cognitive ability measures in the autistic group (the correlations for the typically developing sample in Experiment 3 are displayed in brackets).	188

5.16.	Correlations between susceptibility to different biases in the autistic group (correlations for the typically developing group from Experiment 3 are displayed in brackets).	190
5.17.	Correlations between age, the cognitive capacity measures and susceptibility to different biases in the autistic group. (Correlations that were significant at the $p < .10$ level are marked with †, correlations for the typically developing group from Experiment 3 are displayed in brackets)	191
5.18.	Summary of the results of Experiments 4 and 5.	199
Chapter 6.		
6.1.	The means and standard deviations (in brackets) of the representativeness and equiprobability responses across groups and problem types as a percentage of all responses.	224
6.2.	The means and standard deviations (in brackets) of the normative responses across groups and contents as a percentage of all responses.	225
6.3.	The means and standard deviations (in brackets) of the representativeness and equiprobability responses across groups and problem types as a percentage of all responses.	229
6.4.	The means and standard deviations (in brackets) of the normative responses across groups and problem types as a percentage of all responses.	230
6.5.	The means and standard deviations (in brackets) of the representativeness and equiprobability responses across normative potential (NP) groups and problem types as a percentage of all responses.	236
6.6.	The means and standard deviations (in brackets) of the normative responses across normative potential (NP) groups and problem types as a percentage of all responses.	237
Chapter 7.		
7.1.	Developmental trajectories predicted by developmental dual-process theories, and the cognitive foundations of these changes (a reproduction of Table 2.2.)	250

# Figures

## Chapter 1

- 1.1 The hot hand fallacy and the gambler's fallacy (also known as positive and negative recency effects). 3
- 1.2 The engineers and lawyers problem with a representative and with a non-representative description. 4
- 1.3 The four types of syllogistic reasoning problems that are used in the belief bias paradigm. 7
- 1.4 A summary of Stanovich and West's (2008) model 24

## Chapter 2

- 2.1. Lost ticket and lost money scenarios (Tversky & Kahneman, 1981) 38
- 2.2. Examples for Brecher's (2005) ratio bias task 53
- 2.3. Illustration of Mendel's (1998) rectangle task 55

## Chapter 3

- 3.1. Examples for the conflict and non-conflict versions of the conjunction fallacy task, the two types of engineers and lawyers task, and the positive/negative recency problem. 104
- 3.2. Examples for the different types of syllogistic reasoning problems. 111
- 3.3. Non-distractor and distractor versions of a picture analogy problem (A: non-distractor; B: distractor). 112
- 3.4. Different versions of a scene analogy problem: (A) one relation/non-distractor; (B) one relation/distractor; (C) two relations/non-distractor; (D) two relations/distractor. 114

## Chapter 5

- 5.1. An illustration of the content and conflict manipulations used in Experiment 5. 174

## Chapter 7

- 7.1. Age-related changes in the overall proportion of heuristic, normative and "other" responses on four typical heuristics and biases tasks between the age of 5 and 16 (based on Experiments 1 and 2.) 251

## **Acknowledgments**

I would like to thank Simon Handley and Jonathan Evans for all the interesting discussions, their comments on my work and for all their help and advice. I am also grateful for the help of Rebecca McKenzie with recruiting participants. I would also like to thank the Experimental Psychology Society for making it possible for me to visit Keith Holyoak at UCLA and to design a study with him. Keith Holyoak has been a great person to collaborate with and I very much appreciate his expertise, his being always available for discussions, and his great contribution to our project. I would also like to thank Caterina Primi and Francesca Chiesi at the University of Florence for their interest in my work, and for getting involved in some of my studies. Finally, I would also like to thank the pupils and staff of Brunel Primary and Nursery School, Callington Community College, Chelfham Senior School, Compton C.E. Primary School, Launceston College, Plymstock School, Ridgeway School, Tamarside Community College, and Tavistock College.

## Declaration

I declare that this thesis was composed by myself. The individual differences measures used to match participants in Experiments 4-5 were administered in collaboration with Rebecca McKenzie, a PhD student at the University of Plymouth. The data for Experiments 7-8 were collected by Caterina Primi and Francesca Chiesi at the University of Florence. All other data was collected solely by myself. The conceptual design and analysis of the studies contained in this thesis are my own (including Experiments 7-8), and this work has not been submitted for any other degree or professional qualification.

Data from this thesis are included in the following published papers:

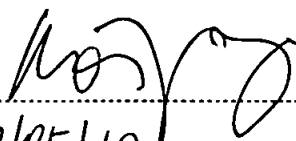
- Morsanyi, K., Handley, S.J. & Evans J.S.B.T. (*in press*). Decontextualised minds: Adolescents with autism are less susceptible to the conjunction fallacy than typically developing adolescents. *Journal of Autism and Developmental Disorders*.
- Morsanyi, K & Holyoak, K.J. (*in press*). Analogical reasoning ability in autistic and typically-developing children. *Developmental Science*.
- Morsanyi, K., Primi, C., Chiesi, F., Handley, S.J. (2009). The effects and side-effects of statistics education. Psychology students' (mis-)conceptions of probability. *Contemporary Educational Psychology*, 34, 210-220.
- Morsanyi, K., Handley, S.J. & Evans J.S.B.T. (2009). Heuristics and biases in autism: Less biased but not more logical. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 75-80). Austin, TX: Cognitive Science Society.
- Morsanyi, K., Handley, S.J. (2008). How smart do you need to be to get it wrong? The role of cognitive capacity in the development of heuristic-based judgment. *Journal of Experimental Child Psychology*, 99, 18-36.

Copies of these papers are attached to the thesis.

Word count of main body of thesis: 82 583.

Signed

Date

  
19/05/10

# Chapter 1: Introduction to the heuristics and biases research programme and dual-process theories of reasoning.

## *1.0. Introduction*

The main aim of my thesis is to examine the development of reasoning heuristics and the relationship between heuristics, cognitive abilities (especially working memory, but also some other aspects of executive functioning, such as inhibition), and relevant knowledge and education. In the first chapter I provide a review of the literature on reasoning heuristics, introduce dual-process theories, and also outline some issues that need further investigation.

In Section 1.1. I introduce the heuristics and biases research programme, including some examples of the typical tasks used in this line of research. I also give an overview of the mechanisms that have been put forward to account for heuristic reasoning, for example the attribute substitution model. I also briefly summarize the ideas of evolutionary psychologists, and especially the concept of the modularity of the mind.

Section 1. 2. introduces dual-process theories of reasoning, the methods used by dual-process theorists, and the evidence for dual processes.

This is followed by a review of some new additions to dual-process theories in Section 1.3., for example the idea that besides cognitive capacity and cognitive styles, relevant knowledge (“mindware”) is also important for analytical reasoning. This section also raises some issues about dual-process theories - for example, whether either heuristic or analytic processes are really uniform, or whether they just share some common properties without necessarily having a common underlying mechanism behind them. I finish this section with a brief description of Gigerenzer’s adaptive toolbox model, and I also give a summary of Gigerenzer’s critique of dual-process theories, and the heuristics and biases program.

The focus of Section 1.4. is on the motivation for my research, and it also gives an outline of the subsequent chapters.

### *1.1. What are heuristics?*

This section introduces the heuristics and biases research program, including some examples of the typical tasks used in this line of research. I also give an overview of the mechanisms that have been put forward to account for heuristic reasoning, and I briefly summarize the ideas of evolutionary psychologists, and especially the concept of the modularity of the mind.

#### *1.1.1. An introduction to the heuristics and biases research program.*

Reasoning and judgment under uncertainty are often based on a limited number of simplifying heuristics rather than formal logic, or rule-based argumentation. Heuristics are low-effort “mental shortcuts” which are useful, even essential, for people who live in an uncertain world and have to come to decisions within reasonable time limits using only restricted information and limited cognitive capacity. Heuristics save time and effort, and usually result in accurate judgment, but sometimes fail utterly and lead to systematic errors (i.e., biases) when presented with data outside of their “domain of expertise”.

The first surprising evidence for the inappropriate use of heuristics came from an investigation of the statistical intuitions of experts (Tversky & Kahneman, 1971). The results showed a systematic bias in their judgments: an overconfidence in the replicability of results from small samples. This tendency is also present in the general population, and it can be illustrated by the *hot hand fallacy* and the *gambler's fallacy* (they are also called positive and negative recency effects, respectively).

The *gambler's fallacy* (see Figure 1.1.) is the belief that, for random events, a run of a particular outcome (e.g., a series of heads on the toss of a coin) will be balanced by a

tendency for the opposite outcome (i.e., tails). Kahneman and Tversky (1972) explained the gambler's fallacy in terms of the operation of the representativeness heuristic. They argued that people expect the essential characteristics of a chance process to be represented in small samples as well. As people expect random sequences to include an equal amount of all possible outcomes, they perceive long runs of the same outcome as non-representative. Consequently, participants will expect runs of the same outcome to be less likely than they are.

Hot hand fallacy	Gambler's fallacy
<p>A basketball player has scored five times in a row, and now he's preparing to shoot again. What is most likely?</p> <p>a) The player will score again. (<i>heuristic response</i>)</p> <p>b) The player won't score this time.</p> <p>c) Both are equally likely. (<i>normative response</i>)</p>	<p>A fair coin is flipped five times, each time landing with tails up; TTTTT. What is the most likely outcome if a coin is flipped a sixth time?</p> <p>a. Tails</p> <p>b. Heads (<i>heuristic response</i>)</p> <p>c. Tails and a heads are equally likely (<i>normative response</i>)</p>

Figure 1.1. *The hot hand fallacy and the gambler's fallacy (also known as positive and negative recency effects).*

The *hot hand fallacy* (see Figure 1.1.) is the exact opposite of the gambler's fallacy. In this case, people have the incorrect expectation that a run of the same outcome will continue, although they are observing a random sequence (Gilovich, Vallone & Tversky, 1985). For example, most basketball fans believe that a player who has just scored several times in a row is now more likely to score - because he or she is "hot." However, in actual fact, if anything, players who have had a run of successful scoring attempts are somewhat less likely to score next time. Although in this case people observe the same sequence of events (i.e., the statistical properties of the sample are identical), they come to the exact opposite conclusion than in the case of the gambler's fallacy. Nevertheless, Gilovich et al. (1985) explained this illusion by the notion of representativeness as well. Namely, when people see runs in players' performance, they interpret this as an evidence for a pattern, and refute the notion of randomness, although the sample is too small to make reliable predictions. As Ayton and Fischer (2004; see also Gronchi & Sloman, 2008) point out,



these explanations only make sense if we assume that people use their prior beliefs as a guide to interpreting random outcomes. Thus, they expect people's performance to reflect their skills and not randomness. On the other hand, they do not expect coins to become "hot".

The observation of such systematic biases in reasoning led Tversky and Kahneman to the idea that intuition is governed by different principles than deliberate, rule-based reasoning. More specifically, intuitive judgment is viewed as an extension of perception to judgment (Kahneman & Frederick, 2005) where automatic inferences can lead to cognitive illusions, and the framing of problems can have a great effect on the perception of problems (see more on *framing effects* below).

<p>A panel of psychologists have interviewed and administered personality tests to 30 engineers and 70 lawyers, all successful in their respective fields. On the basis of this information, thumbnail descriptions of the 30 engineers and 70 lawyers have been written.... For each description, please indicate the probability that the person described is an engineer, on a scale of 0 to 100.</p>	
<i>Representative description:</i>	<i>Non-representative description:</i>
<p>Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.</p> <p>The probability that Jack is one of the 30 engineers in the sample of 100 is _____ %.</p>	<p>Dick is a 30 year-old man. He is married with no children. A man of high ability and high motivation he promises to be quite successful in his field. He is well liked by his colleagues.</p> <p>The probability that Dick is one of the 30 engineers in the sample of 100 is _____ %.</p>

Figure 1.2. *The engineers and lawyers problem with a representative and with a non-representative description.*

Another striking illustration of the neglect of sample sizes, and the effect of prior beliefs is the *engineers and lawyers problem* (Kahneman & Tversky, 1973 – see Figure 1.2.). In the classic study, participants were told that descriptions have been prepared of 30 engineers and 70 lawyers (the base rates were reversed in another condition). Then participants were shown a description of a person from this sample, and they had to decide whether it referred to an engineer or a lawyer. Participants' judgments were mostly based

on the description, whereas base rates had a significant but very small effect on responses. When participants were given a non-representative description (which was not characteristic of either a lawyer or an engineer) the average rating of the likelihood of the person being a lawyer/engineer was 50%. This indicated that participants tended to disregard base rates (which were readily available) even when there was no other information provided. Instead, if an individual's description was neither characteristic of an engineer or a lawyer they concluded that their group membership could not be determined.

Base rates, however, are just one example of the statistical properties that people readily ignore. Probably the most famous example for a task that triggers a reasoning heuristic is the Linda problem (Tversky & Kahneman, 1983), a demonstration of the *conjunction fallacy*. The conjunction fallacy violates a fundamental rule of probability, that the likelihood of two independent events occurring at the same time (in "conjunction") should always be less than, or equal to the probability of either one occurring alone ( $P(A) \geq P(A \& B)$ ). People who commit the conjunction fallacy assign a higher probability to a conjunction than to one or the other of its constituents. The most famous demonstration in the literature is the Linda problem.

*Linda is 31 years old single, outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice and also participated in anti-nuclear demonstrations.*

*Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.*

- a) Linda is a teacher in a primary school.*
- b) Linda works in a bookstore and takes Yoga classes.*
- c) Linda is an active feminist.*
- d) Linda is a psychiatric social worker.*

e) *Linda is a member of Women Against Rape.*

f) *Linda is a bank teller.*

g) *Linda is an insurance salesperson.*

h) *Linda is a bank teller and is an active feminist.*

People usually rank the statement (h) "*Linda is a bank teller and is an active feminist*" above the statement (f) "*Linda is a bank teller*", thus committing the fallacy. As the Linda-problem (and other reasoning heuristics and biases tasks) demonstrate, real-life knowledge has a great influence on people's judgments. Although there are many different reasoning biases most of them are based on the same effect: the contextualisation of problems, that is, the tendency to rely on prior knowledge when solving a reasoning task, and to evaluate arguments on the basis of plausibility or soundness (Thompson, 1996). This is true even when a person is explicitly instructed to ignore their real-world knowledge.

A famous paradigm for investigating this phenomenon is the *belief bias* effect in syllogistic reasoning. The belief bias in reasoning is a non-logical tendency to accept conclusions that are compatible with beliefs more frequently than conclusions that contradict beliefs. Consider the following examples in Figure 1.3. (used by Evans, Barston, & Pollard, 1983). In these examples the believability and validity of the conclusions are systematically manipulated to result in four possible combinations: valid believable, valid unbelievable, invalid believable, and invalid unbelievable syllogisms. In general, participants tend to accept more believable than unbelievable conclusions (in the above study acceptance rates for believable and unbelievable conclusions were 80% and 33%, respectively). Participants also showed evidence of logical reasoning, as they accepted more valid (73%) than invalid conclusions (41%). Interestingly, the effect of beliefs is less pronounced on valid than on invalid problems, giving rise to a logic by belief interaction in acceptance rates. With the inclusion of neutral materials for comparison, Evans and Pollard

(1990) demonstrated that, in general, belief bias is associated with the rejection of unbelievable conclusions and is hence a negative or “debiasing” effect. In this study, acceptance rates for believable and neutral conclusions were similar. However, there was a significant reduction in the acceptance rates of unbelievable conclusions, especially for invalid problems. Positive belief bias (increased acceptance of believable conclusions as valid) was also demonstrated in another study, but it was restricted to certain types of problem (Evans, Handley & Harper, 2001). These results suggest that belief bias is based on two different tendencies. The strongest effect consists of a tendency to withhold a fallacious conclusion that would typically be made in cases when the conclusion is unbelievable, presumably because participants engage in a motivated search for counterexamples. A weaker effect is responsible for a tendency to accept believable conclusions regardless of their logical validity (for similar findings, see also Klauer, Musch & Naumer, 2000).

Valid believable	Invalid believable
<i>No police dogs are vicious. Some highly trained dogs are vicious. Therefore, some highly trained dogs are not police dogs.</i>	<i>No addictive things are inexpensive. Some cigarettes are inexpensive. Therefore, some addictive things are not cigarettes.</i>
Valid unbelievable	Invalid unbelievable
<i>No nutritional things are inexpensive. Some vitamin tablets are inexpensive. Therefore, some vitamin tablets are not nutritional.</i>	<i>No millionaires are hard workers. Some rich people are hard workers. Therefore, some millionaires are not rich people.</i>

Figure 1.3. *The four types of syllogistic reasoning problems that are used in the belief bias paradigm.*

As these examples demonstrate, the tendency to contextualize problems with prior knowledge seems both ubiquitous and automatic, and thus has been termed the *fundamental computational bias* (Stanovich, 2003). In addition to contextualisation, people also tend to “socialize” problems, and rely on pragmatic cues and inferences even in impersonal situations; they are looking for deliberative design and pattern in randomly

generated events; and they try to create a coherent narrative of the events of their lives. Of course, relying on real-life knowledge and trying to make sense of patterns when solving problems is usually the best thing to do. However, being able to disregard previous knowledge and to override instinctive tendencies can be very important in certain situations, such as when we are learning about science, or when we have to choose between mortgage lenders or savings accounts.

Reasoning performance is also affected by the way problems are presented (i.e., *framing*). A famous example for this is Tversky and Kahneman's (1981) Asian disease problem.

*Problem 1:* Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. There are two alternative programs. If Program A is adopted, 200 people will be saved. If Program B is adopted, there is a one-third probability that all 600 people will be saved and a two-thirds probability that no people will be saved.

Which do you prefer, Program A or Program B?

*Problem 2:* Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. There are two alternative programs. If Program A is adopted, 400 people will die. If Program B is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.

Which do you prefer, Program A or Program B?

Although both versions are identical in terms of the possible outcomes, in the case of Problem 1, 72% of the participants preferred Program A, while in the second framing only

22% preferred Program A. According to Kahneman and Tversky (1981) this is because people underweight possibilities that are merely probable as compared to possibilities that are certain (this is called the *certainty effect*). This contributes to the tendency for risk aversion in the case of sure gains, and risk seeking in the case of sure losses. Overall, the effect is based on a passive acceptance of the formulation given, that is, participants focus on gains in Problem 1, and losses in Problem 2, instead of considering and integrating the information about both gains and losses in both frames which would require more effort (Kahneman, 2003).

Even if a person does not have any relevant knowledge, they might *anchor* their judgment on, or *adjust* it to information provided with the problem, regardless of the fact that the information might be unreliable or even deceptive (Kahneman & Tversky, 1973). When anchoring a judgment, people start with an implicitly suggested reference point (the "anchor") and give estimates which are close to it. For example, when asked to guess the percentage of African nations which are members of the United Nations, people who were first asked "Is it more or less than 45%?" guessed lower values than those who had been asked if it was more or less than 65% (Tversky & Kahneman, 1974). This result was replicated in other experiments for a wide variety of different subjects of estimation.

In summary, reasoning can be affected by many aspects of problems, including believability, plausibility, ease of memory retrieval, the context of the problem, etc. These problem attributes can divert attention from the underlying logical structure of the problem, and trigger heuristic responses.

### *1.1.2. Where do heuristics come from?*

According to the most recent conceptualisation of heuristics and biases (Kahneman & Frederick, 2002) heuristics are based on attribute substitution. Instead of answering a difficult question, people are inclined to answer an easier one. How easy a question is

depends on the accessibility of the concepts it involves, that is, how easy it is to activate and retrieve these concepts from memory in a given context. The intent to judge a target attribute initiates a search for a reasonable value. Sometimes the search ends quickly, because people have a stored memory of the required response, or they can rely on their current experience. In other cases, accessing a relevant response is more complicated. In these cases the selection of a particular response crucially depends on its accessibility (i.e., how easily it comes to mind). This, in turn, may depend on stimulus salience, associative activation, priming, specific training, similarity, cognitive fluency, surprisingness, affective valence, etc.

In this view, in the case of the Linda problem the representativeness of the statements is more accessible to participants than the probability of the statements. As a result, participants tend to respond to the easier “How representative is the statement to Linda?” question, instead of answering the more difficult “How likely it is that the statement is true about Linda?” question. In fact, when a group of participants were asked to rate the statements according to how representative they are to Linda, and another group of participants were asked to rate them according to their probability, the correlation between these rankings was almost perfect (.99 – Tversky & Kahneman, 1982). In contrast to this interpretation, the conjunction fallacy has been attributed to pragmatic effects stemming from the linguistic ambiguities that are inherent in the task (see e.g., Hertwig, Benz, & Krauss, 2008). However, removing the linguistic ambiguity is not enough to eliminate the effect (Tentori, Bonini & Osherson, 2004).

Evolutionary psychologists (e.g., Barkow, Cosmides, & Tooby, 1992; Buss, 1999, 2000; Pinker, 1997; Tooby & Cosmides, 1992) claim that the fundamental computational biases are part of the human inferential machinery because they provide a perfect fit with the demands of our natural environment. The famous metaphor pictures the brain as a Swiss army knife (Tooby & Cosmides, 1992) with different tools to solve different tasks

(i.e., specialised modules to solve specific problems) rather than a general purpose “axe” which is powerful, but might be unsuitable to solve certain unique problems. The evolutionary psychologists’ other claim is a reformulation of Chomsky’s (1980) argument on the “poverty of the stimulus”. The original argument is based on the observation that young children learn their native language very quickly, and they are able to use it flexibly despite the lack of any attempt from their parents’ part to teach them the rules of the language in any systematic way. Similarly, young children seem to learn about all important aspects of their environment in a very limited time, and without much guidance or instruction. These theorists argue that this would be impossible without hard-wired, domain-specific learning mechanisms. Finally, the evolutionary psychologists’ framing argument states that general-purpose learning devices are too slow to enable the quick decision-making that was necessary for survival in the wild, and for passing on genes for the future generations. In this view, our mental modules make us capable of reasoning quickly and effectively.

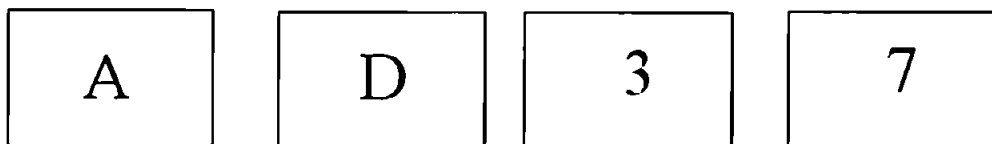
According to Stanovich (2003), what evolutionary psychologists fail to take into account is that our brains and their specialised modules are “frozen in time”, adapted to circumstances that existed up to 10, 000 years ago when humans lived in the so-called environment of evolutionary adaptedness that existed throughout the Pleistocene (Buss, 1999). As biological evolution takes place at a much slower pace than cultural development, it is no wonder that our brains lag behind the challenges of our days in some respect. Many evolutionary psychologists (e.g., Tooby & Cosmides, 1992) advocate a massive modularity approach where they claim that all inferential processes of our brains are adapted to specific functions, and there is no general purpose mechanism that could deal with a variety of problems. However, our modern technological society poses many challenges that content-specific cognitive modules are unable to deal with, as they require



decontextualised representations. To achieve decontextualisation, we need some sort of context-independent, domain-general cognitive process.

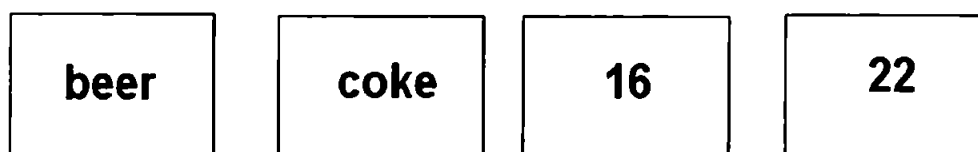
In fact, one serious problem with the massive modularity claim is the flexibility of human cognition (e.g., Machery, 2008). We can distinguish between three main types of flexibility: stimulus independence (the fact that our cognitive processes are very much independent of our current perceptions); content flexibility (our capacity to combine concepts that are supposed to belong to different modules); and the flexibility of reasoning (our ability to change our views). These are hard to explain if human cognition is based solely on rigid modular processing.

Another problem is the way evolutionary claims have been tested. In a series of papers, Cosmides and colleagues (Cosmides, 1989; Cosmides & Tooby 1994; Fiddick et al., 2000) have used the *Wason selection task* (Wason, 1966) to test whether humans possess a psychological adaptation for cheater detection. Consider this common form of the problem. Cards are labelled "A" or "D" on one side and "3" or "7" on the other, and you have the following four cards:



A rule says that "if there is an A on one side of the card then there is a 3 on the other side". The task is to decide which cards need to be turned over to know whether the rule is true or false. The common responses are "A and 3", or "A" alone. The correct response is "A and 7" (this is the only combination of cards which could refute the rule if it is not true). Overall, about 90% of participants give an incorrect response.

In contrast to the poor performance on the abstract version of the selection task, participants solve thematic (deontic) versions of the selection task very easily (the first demonstration came from Wason and Shapiro in 1971, and was followed by many others). For example, consider the following version of the problem (Griggs & Cox, 1982). On one side of each card is the name of a drink; on the other side is the age of the drinker. What card(s) must be turned over to determine whether the rule is being violated? *If a person is drinking beer, then the person is over 19-years-old.*



In Cosmides and colleagues' studies the problem is presented as a social contract infringement. For example, the rule is that "If a man eats cassava root, then he must have a tattoo on his chest". In this case a cheater would be somebody eating cassava root without a tattoo on their chest. According to Cosmides, the fact that people solve these versions of the selection task very easily points to the operation of a content-specific mental algorithm devoted to social contract problems. However, one problem with the use of the selection task for investigating deductive reasoning is that in most cases very little reasoning is involved when people solve the task (see e.g., Evans & Over, 1996, 2004; Sperber, Cara & Girotto, 1995). Moreover, some of the variants of the Wason task that Cosmides used were not actually Wason tasks at all. In addition, they were not even appropriate for the purpose of discovering a competence for social exchange, because they did not ask participants about the truth or falsity of a conditional rule, but merely asked for the selection of cards that defined a category (such as "cheater" in the above example), and so were trivially easy (Sperber & Girotto, 2002). Sperber and Girotto concluded that Cosmides' hypothesis has not been properly tested experimentally, least of all by Cosmides herself.

Finally, the idea of modularity gave a boost to research into developmental disorders starting in the late 1980s. For example, a specific deficit of the “theory of mind module” had been put forward to account for the core symptoms of autism (i.e., deficits in social behaviour, social cognition, and communication – see e.g., Baron-Cohen, Leslie & Frith, 1985). However, theory of mind in young children is consistently found to be highly correlated with executive functions, even after controlling for age and/or receptive vocabulary (e.g., Frye, Zelazo, & Palfai, 1995; Perner, Lang, & Kloo, 2002). Apperly, Riggs, Simpson, Samson, and Chiavarino (2006), using reaction time data, also demonstrated that adults do not ascribe beliefs to agents automatically, and yet another study (McKinnon & Moscovitch, 2007) reported that both older adults and younger adults, under conditions of divided attention, showed performance decrements on theory of mind tasks. The idea that developmental disorders can be characterized by a distinct pattern of intact and impaired cognitive modules has also been challenged on a theoretical basis, as it disregards the process of ontogenetic development, and the plasticity of the brain (Karmiloff-Smith, 1998; Karmiloff-Smith, Scerif & Ansari, 2003).

Nevertheless, there is evidence for innate reasoning systems. Research with infants suggests that human cognition is founded, in part, on four systems. These systems have evolved for representing 1.) objects and their mechanical interactions; 2.) agents and their goal-directed actions; 3.) sets and their numerical properties (together with addition and subtraction); and 4.) space and geometric relationships (Spelke & Kinzler, 2007). These systems are shared with non-human primates; they appear in the first months of development, and they continue to shape the mental lives of adults. These systems, however, have their gaps and inaccuracies which make both adults and children prone to errors in reasoning about properties of object mechanics, or non-Euclidean geometry (e.g., McCloskey, 1983; Randall, 2005). Although these concepts are resilient, they can be overcome by explicit instruction and experience. Moreover, conceptual change does not

only occur in academic science. Preschool children change their conceptions of numbers when they learn to count (Spelke, 2000), and they change their conceptions of agents when they learn about biological processes like eating and breathing (Carey, 2001). Thus, in this approach innate learning algorithms form the basis of human cognition. However, these innate schemas are reasonably flexible, and they can be modified by learning and experiences.

The ideas of modularity (or innate reasoning algorithms – e.g., Cosmides & Tooby, 1994), and associative processes (e.g., Kahneman & Frederick, 2005) have also been incorporated in reasoning theorists' conceptualisation of intuitive inference. Stanovich (2004) described The Autonomous Set of Systems (TASS), which is claimed to form the basis of heuristic processing, as including implicit learning, overlearned associations, and the modular processes for solving adaptive problems. In the next section I introduce dual-process theories which claim that human reasoning is based on the interaction of quick, heuristic, and slow, effortful processes.

## *1.2. Dual-process theories of reasoning*

This section introduces dual-process theories of reasoning, the methods used by dual-process theorists, and the evidence for dual processes.

### *1.2.1. Distinguishing between two types of reasoning processes*

In many cases people disregard statistical information or the logical structure of problems and rely on their immediate impressions or relevant knowledge when they make judgments and decisions. This can happen even when they have knowledge of the relevant statistical rule, such as in the case of experts (Tversky & Kahneman, 1971). To explain this contrast between people's normative knowledge and their often non-normative decisions, it has been suggested that these different outputs represent the operations of two separate

reasoning systems: a quick and intuitive, and a slow, more deliberate system (e.g., Kahneman & Frederick, 2002). In this section I introduce the ideas of dual-process theories.

Dual-process theories (Evans and Over, 1996; Sloman, 1996; Stanovich, 1999) are based on the idea that besides automatic inferential processes human reasoning also relies on general purpose problem solving mechanisms. Thus, they presuppose two systems (or more recently, two types of processes – Evans, 2008) that form the basis of human reasoning. The original theories pictured the two systems as “two minds in one brain” (Evans, 2003). They distinguished between System 1 which is fast, automatic, effortless, independent of cognitive abilities, contextually cued, phylogenetically older and shared with other animals, and System 2 which is slow, effortful, related to cognitive abilities and dispositions, context-independent and uniquely human.

System 1 processes happen rapidly and mandatorily when their specific triggering stimuli are present, they can be executed in parallel with other processes without much mental effort, and only their outputs are available for consciousness. In this view the main purpose of System 2 (the conscious and effortful system) is to override the default System 1 responses through the process of decoupling (e.g., Stanovich, 1999, 2006). Decoupling is a mechanism that enables people to maintain and manipulate internal representations in the presence of distractive interfering stimuli (i.e., it is basically the decontextualisation of representations), which is essential for hypothetical thinking.

A more recent publication (Evans, 2008) does not commit to a two systems approach but proposes the name of Type 1 and 2 processes. The main basis for discriminating between these processes is that Type 2 processes require access to a single, central working memory resource, while Type 1 processes do not. This implies that Type 2 processes are slow, sequential and capacity limited, their functioning correlates with individual differences in cognitive ability and they are disrupted by concurrent working memory load.

They are also associated with consciousness and intentional, higher order control. In contrast to the original idea of System 2 these processes are not considered to be uniquely human, and are not necessarily associated with decontextualised, logical reasoning.

Type 1 processes in this approach are a collection of implicit processes that can operate automatically without occupying working memory space. These are supposed to include innate cognitive modules (e.g., for perception, attention, and language processing) as well as an associative and implicit learning system that acquires knowledge of the world which cannot be retrieved as explicit knowledge but which can directly affect our behaviour. There are also habitual and automated behaviour patterns that once required conscious Type 2 effort, but that have become Type 1 with practice and experience. Finally, pragmatic processes that rapidly identify and retrieve explicit knowledge for conscious processing are also Type 1.

An important question is how Type 1 and Type 2 processes interact, and how we can tell which one will provide a response to a particular question. In *default-interventionist* models (Evans, 2006; Kahneman & Frederick, 2002), Type 2 processing is based on the continuous stream of relevant content delivered into working memory by Type 1 processes. Thus, the basis of Type 2 processes is contextualised representations of problems, which also cue intuitive answers. Quick, knowledge-based, context-dependent responses are usually very useful. However, in some cases they might be in conflict with the person's aims, or they can be inappropriate for other reasons. In these situations these quick heuristic responses can be overridden by Type 2 processing given that the reasoner is aware of the conflict, has sufficient cognitive capacity, and possesses certain mental dispositions (Stanovich & West, 2008). This overriding process requires conscious effort, and people with the highest working memory capacity will be more likely to succeed (e.g., De Neys, Schaeken, & d'Ydewalle, 2005; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Markovits & Doyon, 2004; Moutier, Plagne-Cayeux, Melot, & Houdé, 2006). Such a

belief inhibition or decontextualisation process is the basis of decoupling and hypothetical thinking (e.g., Stanovich & West, 2000). On the other hand, when there is no conflict between beliefs and logic, performance will be unrelated to cognitive capacity (e.g., Newstead, Handley, Harley, Wright, & Farrelly, 2004; Stanovich & West, 2000). In other cases, however, the role of Type 2 processes is merely to rationalize responses cued by Type 1 processes (Evans, 1995). Stanovich (2009) used the term *serial associative cognition with a focal bias* to describe a reasoning process which is not purely associative and effortless (that is, not completely Type 1), but consists of the justification of a heuristically cued response. This involves conscious, but low-effort processing which lacks a real attempt to consider alternatives, instead it is looking for easy ways of justifying the most obvious answer.

Besides *default-interventionist* models, Evans (2007) also identified so-called *parallel-competitive* models. These models assume that each system operates in parallel to deliver a putative response, resulting sometimes in conflict which then needs to be resolved. Examples of these kinds of models are Epstein's (1994) rational-experiential theory or Sloman's (1996) dual-process account. An interesting feature of these models is that they suppose that because both systems deliver a response, when these responses are in conflict people will believe two contradictory things simultaneously. This also implies that reasoners will always take analytic considerations into account, and they will always be aware when there is a conflict between the responses offered by the two systems. In this view errors arise because people fail to inhibit their prepotent heuristic beliefs. This is in contrast with the *default-interventionist* view (e.g., Evans, 1984; Kahneman & Frederick, 2002) where Type 1 processes always come first, and they offer a default solution which is very often accepted without any further thought, or without people ever becoming aware of the relevance of normative considerations.

In recent publications, De Neys (De Neys & Glumicic, 2008; De Neys, Vartanian & Goel, 2008) argued that people are able to detect the conflict between heuristically cued responses and logic, although they might not be aware of this, or the processing demands of the task might prevent them from choosing the normatively correct response. Nevertheless, when selecting a heuristically cued response which is in conflict with standard logic, participants take longer to respond (De Neys & Glumicic, 2008). An fMRI study also showed activation of the anterior cingulate cortex (which is involved in the detection of competing responses) when people selected a response that violated a normative rule (De Neys et al., 2008). Ball, Phillips, Wade and Quayle (2006) using latency and eye-tracking methods also demonstrated that people took longer to respond, and spent more time inspecting problems where logic and belief conflicted, as compared to non-conflict problems.

### *1.2.2. Evidence for the existence of Type 1 and Type 2 processes*

As I described in the previous section, the fundamental claim of dual-process theories is that human reasoning is based on two distinct types of processes: quick, contextualised, and effortless (Type 1) ones, and slow, sequential and effortful (Type 2) ones. The main basis for differentiating between these processes is whether they require working memory capacity or not (Evans, 2008). The idea is that reasoning processes which rely on working memory capacity are slow, sequential and they need conscious attention, whereas automatic, Type 1 processes are rapid, can be executed in parallel, and they do not require mental effort. In order to demonstrate the difference between these processes, a number of different methods have been implemented.

In *speeded tasks*, participants have to give a response to a reasoning problem in a short period of time determined by the experimenter. Under such conditions the effect of certain reasoning biases increases (e.g., the matching bias on the Wason selection task -



Roberts & Newton, 2001; and the belief bias in syllogistic reasoning - Evans & Curtis-Holmes, 2005). Giving instructions for fast responding also increases the rate of endorsement of fallacies in conditional reasoning (Schroyens, Schaeken, & Handley, 2003; although see Evans, Handley & Bacon, 2009).

*Non-verbal process tracing methods* such as asking participants to indicate the current focus of their attention by mouse-pointing, or using eye-tracking methods can also give insight into participants' cognitive processes. For example, owing to these methods we know that on the Wason selection task, people spend quite a long time thinking about cards that they will end up selecting (Ball, Lucas, Miles, & Gale, 2003; Evans, 1996). These are often the ones corresponding to well-known heuristics (see Evans, 1998; Evans & Over, 2004). Verbal protocol data also indicate that participants are more likely to rationalise their initial choices (Wason & Evans, 1975) than to try and consider multiple possibilities.

Looking at *individual differences in cognitive ability* (usually working memory) and *cognitive style* (i.e., how much an individual is inclined to engage in effortful cognitive operations) and how these correlate with participants' responses has also been used in many studies. A series of experiments by Stanovich and West (reviewed by Stanovich, 1999; Stanovich & West, 2000) used this methodology, as well as a number of studies by other authors (for example, Capon, Handley, & Dennis, 2003; De Neys, 2006; Klaczynski, 2000; Klaczynski & Daniel, 2005; Klaczynski & Gordon, 1996; Newstead, Handley, Harley, Wright, & Farelly, 2004). These studies show that when people have to solve tasks where there is a straightforward, heuristic response, together with a normative, but less straightforward option, people higher in cognitive ability, and people who are inclined to think hard about problems are more likely to choose the normatively correct answer. However, when a contextualised (knowledge-based) representation of the problem leads to

a normative solution cognitive ability is not related to participants' reasoning performance (Stanovich & West, 1998).

The influence of cognitive ability on reasoning performance was also demonstrated using *dual-task manipulations*. In a study looking at the effects of beliefs in syllogistic reasoning (De Neys, 2006) participants' working memory capacity was burdened with a secondary task. When the believability of a conclusion conflicted with its logical validity, participants showed worse logical performance. However, there was no effect of cognitive load when beliefs and logic cued the same response. Although participants with high working memory spans performed better than those with lower spans when there was a conflict between logic and beliefs, all reasoners showed similar effects of cognitive load, indicating that the difference between high and lower ability participants was quantitative rather than qualitative. De Neys, Schaeken, and d'Ydewalle (2005) also found that burdening working memory with a secondary task decreased the efficiency of counterexample retrieval in conditional reasoning, although the effect was less pronounced for the most strongly associated counterexamples. This indicated that in addition to an automatic search component, counterexample retrieval draws on working memory resources.

*Instructing participants to reason intuitively/logically* can also have an effect on their reasoning performance. According to previous studies (e.g., Epstein et al., 1992; Ferreira et al. 2006; Klaczynski, 2001) instructing participants to reason logically (e.g., "take the point of view of a perfectly logical person" – Denes-Raj & Epstein, 1994) increases the mental effort that they invest in the task, and as a result, participants give more normative responses. Although the exact mechanism is not known, a possible explanation for this effect is that logical instructions sensitise participants to the potential conflict between logic and intuitions and/or encourage participants to rely more on Type 2 reasoning processes (Stanovich & West, 2008). More specifically, it has been proposed that Type 2

or analytic processes can intervene to inhibit reasoning biases and replace them with normatively correct reasoning (see Evans, Handley, Neilens, Bacon & Over, *in press*). Evans et al. (*in press*) also found that participants with higher cognitive ability comply with instructions more readily than lower ability participants.

In short, a large body of experimental evidence is pointing to the existence of dual processes, making a distinction between quick, pre-conscious, effortless, and slow, conscious and effortful processes seem reasonable. It also seems that slow, effortful processing is more likely to result in a normative solution, and people with higher cognitive capacity (i.e., people who are able to invest more effort) are more likely to reason normatively.

### *1.3. New additions to dual-process theories and alternative approaches*

This section reviews some of the new additions to dual-process theories, for example the idea that besides cognitive capacity and cognitive styles, relevant knowledge (“mindware”) is also important for analytical reasoning. This section also raises some issues about dual-process theories - for example, whether either heuristic or analytic processes are really uniform, or whether they just share some common properties without necessarily having a common underlying mechanism behind them. I finish this section with a brief description of Gigerenzer’s adaptive toolbox model, and I also give a summary of Gigerenzer’s critique of dual-process theories, and the heuristics and biases program.

#### *1.3.1. The effect of knowledge on reasoning, and the revised heuristic-analytic theory*

In contrast to the findings that demonstrate a link between cognitive capacity and the ability to override heuristic responses, recently it emerged that a large number of thinking biases (including, the conjunction effect, framing effects, anchoring effects, outcome bias, base-rate neglect, “less is more” effects, affect biases, omission bias, myside bias, sunk-

cost effect, and certainty effects) are uncorrelated with cognitive ability when they are presented in a between subjects design. Moreover, people with higher cognitive ability are no less susceptible to the effects of beliefs (Klaczynski, 2000; Newstead et al., 2004; Torrens, Thompson, & Cramer, 1999), and they are also prone to egocentric thinking and myside bias (Stanovich & West, 2007, 2008).

To account for these findings Stanovich and West (2008) have suggested that cognitive capacity will only predict reasoning performance if the following conditions are satisfied. Participants have to have the necessary “mindware” (i.e., relevant knowledge), they have to detect the need to override the default heuristic response, and sustained inhibition or decoupling has to be necessary to solve the problem. If any of these is not present then there will be no relationship between people’s performance on a task, and their cognitive abilities. In addition, it is more likely that people will detect the conflict between heuristics and normative reasoning if they are given instructions to think logically, if there are cues in the task that make the conflict salient (e.g., when a within-subjects design is used as opposed to a between-subjects design), and when people are predisposed to reason carefully, and to invest mental effort into solving problems (see Figure 1.4. for a summary of the model).

Another recent development in the dual-process field is the revised heuristic-analytic theory of Evans (2006a; Evans, Handley & Over, 2003). According to this theory, the three basic principles of hypothetical thinking are: 1. the *singularity principle* (people consider only one possibility/one mental model at a time); 2. the *relevance principle* (mental models are shaped by preconscious heuristic processes that contextualise problems in a way that makes them maximally relevant to current goals); finally 3. the *satisficing principle* (analytic processes evaluate the models, but they tend to accept them unless there is good reason to reject them, in which case the cycle starts again from the beginning, with the consideration of another possibility).

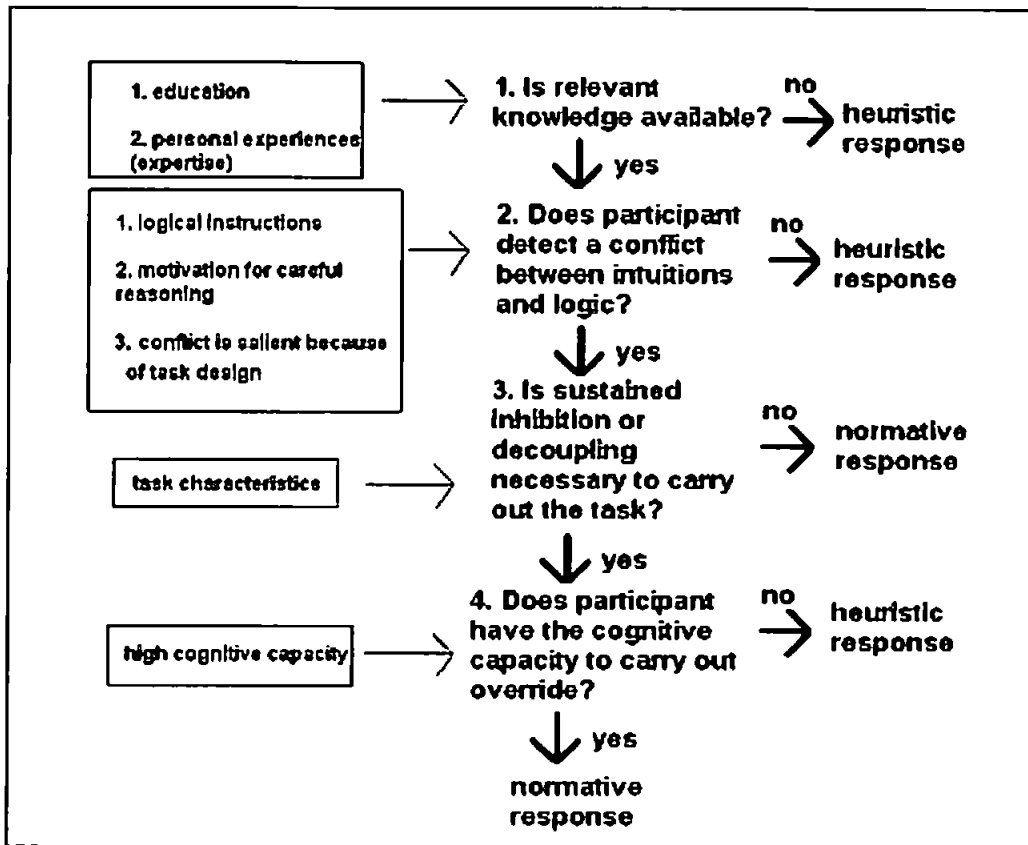


Figure 1.4. A summary of Stanovich and West's (2008) model.

In this conceptualisation, biases are the result of logically relevant information (such as base rates, for example) being omitted or logically irrelevant information (such as frames or some other salient aspects of the representation) included at the heuristic stage to create a believable or plausible model. Other biases, however, are attributed to the operation and properties of the analytic system. One important factor is the limited capacity of the analytic system. For this reason, people consider only one model at a time, and consider evidence in relation to the current model, although they are capable of revising or rejecting the model in the light of new evidence. For example, in the case of syllogistic reasoning people construct a model of the premises which is motivated by the believability of the conclusion, and tend to accept conclusions that could be true given the premises, but do not have to be true (i.e., they do not necessary follow in a logical sense - e.g., Evans, Handley, Harper & Johnson-Laird, 1999).

Both Stanovich and West's (2008) and Evans's (2006a) model offers a more complex idea of human cognition than their initial conceptualisations of dual processes (Evans & Over, 1996; Stanovich, 1999). One aspect of human cognition that both models emphasize is that cognitive effort does not always guarantee a normative solution, or even an abstract, logical representation of the task. Hypothetical thinking, and mental simulation of alternative outcomes might not be possible without Type 2 thinking, but Type 2 thinking is also no guarantee for an abstract, logical representation (see Evans, 2006a; Stanovich, 2009).

### *1.3.2. How many systems, and how do they differ?*

Recently, Evans (2006b) also published a paper where he identified a couple of problems concerning dual-process theories. One problem is the claim that System 1 is ancient, and shared with other animals, whereas System 2 is modern, and uniquely human. Probably the most interesting evidence about this issue concerns the belief bias effect. Neurological evidence (Goel & Dolan, 2003) indicates that the belief bias arises in frontal brain areas associated with semantic memory, thus it is obviously not related to the functioning of an ancient part of the brain. Another popular idea in dual-process theories is associating System 2 with controlled, and System 1 with automatic processes. As the heuristic-analytic theory of Evans (2006a) implies, although analytic processes might override heuristic response tendencies, most of the time analytic processing consists of a superficial rationalisation process that justifies heuristically cued responses. In these cases it is clearly System 1 that is in control of our behaviour. For similar reasons, it is inappropriate to think about analytic processes as necessarily abstract and logical. Finally, as already acknowledged by Stanovich (2004) in his idea of The Autonomous Set of Systems (TASS), heuristics are based on a number of separate cognitive processes with

different origins, and probably different neural bases as well, making the idea of a unitary heuristic system implausible.

Even more recent developments in the field also question the unitary nature of the analytic system. For example Stanovich (2009) proposed a distinction between the algorithmic and autonomous minds, corresponding to the independent effects of fluid intelligence (or working memory), and thinking dispositions (as indexed by measures of need for cognition, and actively open-minded thinking, for example) on reasoning performance. Another possible distinction is between the effects of working memory and inhibition. In Stanovich's (e.g., Stanovich, 1999, 2006) conceptualisation decoupling is one of the most important roles of the analytic system. Decoupling is a mechanism that enables people to maintain and manipulate internal representations in the presence of distracting interfering stimuli (i.e., it is basically the decontextualisation of representations), which is essential for hypothetical thinking. However, psychometric evidence (e.g., Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006) indicates that the maintenance of information, and resisting interference can be attributed to two closely related, but separable executive function components: updating (which is closely related to working memory), and inhibition. Although the two components are highly correlated (in the above study there was a correlation of  $r=.62$  between the two), they are clearly not the same.

In fact, we have evidence from reasoning research that working memory and inhibition play a different role in analytic reasoning. In a study with children that investigated the belief bias effect (Handley, Capon, Beveridge, Dennis & Evans, 2004) where the content was manipulated in such a way that the conclusions were either congruent, neutral, or incongruent with beliefs, and either logically valid or logically invalid. Participants also received a measure of working memory capacity (the counting span task) and a measure of inhibitory control (the stop signal task). On belief-based problems, belief bias and logical reasoning were predicted independently by both working

memory and inhibition. In contrast, logical reasoning on belief neutral problems was predicted by working memory alone. These results indicate that whereas logical reasoning in general depends on working memory capacity, inhibition might only be necessary to resolve the conflict between beliefs and logic.

Overall, these results suggest that analytical reasoning performance depends on separable aspects of executive functioning, as well as dispositional factors, which makes the idea of a unitary analytic system seem less plausible. A similar argument has been made about the heuristic system, and it is reflected for example in Stanovich's (2004) concept of The Autonomous Set of Systems (TASS). Although at the behavioural level we can distinguish between Type 1 and Type 2 processes, it seems unlikely that we can find corresponding structures at the level of brain architecture or even of cognitive mechanisms.

### *1.3.3. Gigerenzer's critique of dual-process theories*

One of the main opponents and critics of the heuristics and biases research programme is Gerd Gigerenzer who proposes a very different approach to the study of heuristics. The fast and frugal heuristics approach (e.g., Gigerenzer et al., 1999) makes similar claims as evolutionary psychologists in that they propose that our brains consist of specialized modules or evolved capacities (such as recognition and recall memory, or imitation), which serve as the building blocks for heuristics. These building blocks, which can be flexibly combined with each other, form the basis of what they call the adaptive toolbox. These theorists claim that our brains are adapted to the demands of our days (not to the demands of the Pleistocene) which is evident from the fact that most decisions we make lead to satisfactory outcomes. They claim that heuristics outperform (or they are at least as effective as) complex statistical models in situations where people need to act fast, the probabilities or utilities are unknown, and there are multiple goals and ill-defined problems (Gigerenzer, 2008). Although fast and frugal heuristics might not lead to optimal



outcomes from the point of view of strict logic or probability theory, they still offer fast and good-enough solutions to problems (i.e., they are ecological, rather than logical).

So, how do we select a heuristic from the adaptive toolbox, or how do we construct a new one from the existing building blocks? Heuristic selection can be based on a) reinforcement learning; b) social learning (learning from tutors, for example); c) evolutionary learning (Hutchinson & Gigerenzer, 2005). Heuristics can be applied consciously or unconsciously, and people tend to check the ecological rationality of heuristics from trial to trial, ensuring a good fit with the environment. As only a small set of heuristics is potentially applicable in any given situation, this makes the search and selection process quick and easy.

An interesting statement of the fast and frugal heuristics approach is that the more unpredictable a situation is the more information needs to be ignored, even if the information costs nothing. For example, the “take the best” heuristic (Gigerenzer & Goldstein, 1996) is based on the idea that when people search for the best solution, they should search according to just one criterion (which is the most important) instead of weighing up different types of evidence. Decision trees based on this rule have been successfully used in medical settings, for example, to decide coronary care unit allocation (Green & Mehr, 1997) and macrolide prescription (Fischer, Steiner, Zucol, Berger, Martignon & Bossart, 2002).

In a recent publication, Gigerenzer (2009) criticised the heuristics and biases approach, and dual-process theories, highlighting some important shortcomings. One problem he noted (which he called one-word explanations) was the way heuristic processes were defined in the heuristics and biases approach. For example, the term “availability” has been used to denote the “number” of instances that come to mind, the “ease” with which instances come to mind, as well as recency, salience, memorability, and vividness. However, when ease and number were actually measured, they were found to differ and,

most importantly, to *not* correlate with the frequency judgments that the term *availability* purportedly explains (Sedlmeier, Hertwig & Gigerenzer, 1998). Similarly, the term “representativeness” has been used to stand for two opposite phenomena: the gambler’s fallacy and the hot hand fallacy (see also Section 1.1.1.). The main problem here is that although the heuristics and biases approach identified and gave labels to many common errors, they failed to uncover the underlying cognitive mechanisms of heuristic reasoning.

Another issue that Gigerenzer raised was what he called circular restatements. For example, from the observation that people are influenced both by the logical form of a syllogism and the believability of its conclusion, it does not necessarily follow that these effects are based on two reasoning systems: a logical one, and a belief-based one. Similarly, framing effects were “explained” by the fact that one of the frames makes the solution more “transparent” or “salient.” The fact that a representation makes a solution transparent is hardly an explanation, however, but rather what needs to be explained. Finally, according to Gigerenzer dual-process theories are no more than a “yin-yang list of dichotomies” (e.g., associations versus rules, intuitive versus rational processing, etc.) which taken together can account for all possible phenomena, again, without specifying the underlying processes.

In sum, the main problem seems to be that although research on heuristics and biases identified some important characteristics (mostly weaknesses) of human cognition, and gave labels to them, it has not been very successful in uncovering the underlying cognitive mechanisms. Unfortunately, the fast and frugal heuristics approach (i.e., Gigerenzer’s own theory) suffers from very similar problems. One important issue is that, although this approach has been successful in identifying some heuristics that people use, and that can lead to good decisions in a fast and frugal way, Gigerenzer and colleagues (e.g., Gigerenzer, Hoffrage & Goldstein, 2008) emphasize that the heuristics that people will use in a given situation will largely depend on their individual learning history. Thus, it is hard

to predict what sort of heuristic a given individual will use in a particular situation. In addition, even if they use a particular heuristic, this can lead to different responses depending on the person's individual learning history.

For example, in the case of the *take the best* heuristic, people can rely on evolutionary learning (which could play a role, for example, in partner choice). However, individual learning can also have an effect - although this type of learning is very slow, and it can be dangerous too, or even impossible in the case of rare events or when feedback is absent or unreliable. Finally, social learning is probably the fastest and most widely used type of learning. Social learning can be supported by the "imitate the successful" heuristic (e.g., Garcia-Retamero, Takezawa & Gigerenzer, 2006), that is, to imitate the behaviour of the most successful member in a social situation, although many other mechanisms can play a role too. As this example illustrates, based on the fast and frugal heuristics approach, it is very hard to make predictions about people's future behaviour in any given situation, as this will be affected by such a large number of (both individual and environmental) factors.

#### 1.4. *Questions to be answered and the outline of the chapters to follow.*

In Section 1.2 I reviewed evidence suggesting that a distinction between fast and effortless, and slow and effortful reasoning processes describes many aspect of human reasoning in a meaningful way, and the two types of processes can be distinguished between in experiments. However, as I described in Section 1.3., the problem with dual-process theories is that the underlying processes of reasoning, and especially heuristic reasoning, are mostly unknown. In my thesis I am going to investigate three broad topics which are all aimed at answering some questions about how heuristics develop.

According to Stanovich (2004) heuristics can be based on implicit learning, overlearned associations, and the modular processes for solving adaptive problems.

Theoretically, it should be possible to distinguish between these different types of heuristic by looking at developmental samples. We can expect innate heuristics to appear very early in the course of development, and they probably correspond to the mechanisms described by Spelke and Kinzler (2007; see Section 1.1.2.). We should also be able to distinguish between heuristics that are the result of implicit versus explicit learning. Heuristics based on implicit (i.e., associative) learning should be effortless from the start, and they are likely to be inaccessible to conscious awareness. By contrast, heuristics that are based on explicit learning (e.g., the equiprobability bias), or the ones that require the integration of presented information and retrieved knowledge (e.g., the conjunction fallacy) might be effortful to begin with, and they should deliver output which is available for consciousness. They might also emerge at a later stage of development.

The prediction that follows from this is that if we look at a sample of young children, some typical heuristic responses should be positively correlated with cognitive ability (that is, children of higher cognitive capacity should give more heuristic responses). Moreover, older children should use these types of heuristics more than younger children. Later on, as these heuristics become “overlearned” (i.e., automatized) the pattern should reverse (as at this stage children become able to suppress heuristics voluntarily), and older children, and children of higher cognitive ability should give these responses less. In order to be able to examine this question, in Chapter 3 I present some experiments where we looked at a wide range of heuristics, and children from a broad age range (between the age of 5 and 16). However, before doing this I review the literature on the development of heuristic reasoning, and developmental dual-process theories in Chapter 2.

Heuristics are very similar across individuals. At least some typical heuristics (such as the matching response in the case of abstract versions of the Wason selection task) are produced by a large majority of people. The reason for this is probably that heuristics have high adaptive value, as they provide us with appropriate, fast and effortless responses

in most everyday situations. An interesting question is whether we can also observe the same heuristics in atypical populations, for example in children with a developmental disorder. By definition, children with a developmental disorder are less well adapted to their environment than typically developing children. Thus, it seems possible that children with a developmental disorder produce fewer heuristic responses than typically developing children, and as a result in the minority of cases, when heuristic use is maladaptive, they should actually give more normatively correct responses than their typically developing peers. In order to test this intriguing hypothesis, I decided to investigate heuristic reasoning in autism.

Autism is an interesting disorder from the point of view of reasoning research. One of the cognitive theories of autism, the mind-blindness hypothesis which was first proposed by Baron-Cohen, Leslie and Frith (1985) states that some of the symptoms of autism can be explained by a deficit of the theory of mind (ToM) module. As I described in Section 1.1.2. it is unclear whether a ToM module actually exists or not. However, autistic people's difficulty with ToM reasoning seems to suggest that some of the quick and effortless (or low-effort) reasoning processes are effortful for autistic people.

Another relevant cognitive theory of autism is the Weak Central Coherence (WCC) theory (Frith & Happé, 1994; Happé, 1999). The WCC theory proposes that typically-developing individuals tend to engage in global processing, building up a gist-based representation, whereas autistic individuals engage in more detailed, local or piecemeal processing. As a result, autistic people are less able to process information in context (e.g., Jolliffe & Baron-Cohen, 1999). From this it seems to follow that autistic children should be able to avoid some reasoning heuristics. According to Stanovich (1999, 2006) decoupling or decontextualisation is probably the most important function of the analytic system. If contextualisation does not take place in autism (or it takes place to much less extent than in

typical development) we can expect that autistic people will perform better in situations when a decontextualised representation is required for a logical/normative answer.

Interestingly, the popular image of autism is that autistic people are very objective, unbiased, factual and logical. This is even reflected in yet another theory of autism, the “extreme male brain theory” (Baron-Cohen, 2002) which proposes that autistic people are very good at analysing and constructing systems, whereas at the same time they are less able to empathize with other people. Although the cognitive theories of autism are quite diverse, they all seem to predict that autistic people will use reasoning heuristics less, and show more normative performance in cases where logic and heuristics conflict. In Chapter 4 I present a review of the cognitive theories of autism. In Chapter 5 I describe two studies that investigated heuristic use in children with autism.

Finally, the third topic that I investigate in my thesis is the role of knowledge (and education) in giving heuristic/analytic responses. I look at this question by comparing the performance of students at different stages of their education, and students studying different disciplines at a university level. According to Stanovich and West’s (2008) model, relevant knowledge (“mindware”) should have an effect on reasoning. More specifically it should decrease heuristic use. By contrast, educational theorists (e.g., Fischbein, 1997) propose that education can actually increase heuristic use in some cases. Because of these conflicting predictions, and because the role of knowledge and education in heuristic use has been largely ignored in reasoning research so far (apart from some studies on the effect of training) it seems an interesting question to look at. A corollary of the idea that education can increase heuristic use is that some heuristics might appear at later stages of development, well beyond childhood. To test for this possibility I use a sample of university students in the experiments in Chapter 6. In this chapter I also review the literature on the role of education and knowledge in reasoning heuristics.

Chapter 7 is a summary of the experimental results, and an attempt to evaluate the implications of my findings for dual-process theories, and for the broader field of reasoning research. In the last chapter I also outline some possible future directions of research.

## Chapter 2. The development of reasoning skills.

### *2.0. Introduction*

In the previous chapter I introduced dual-process theories of reasoning which is the general framework for my investigation. I reviewed evidence suggesting that a distinction between fast and effortless, and slow and effortful reasoning processes describes many aspect of human reasoning in a meaningful way, and the two types of process can be distinguished between in experiments. However, as I noted earlier, the problem with dual-process theories is that the underlying processes of reasoning, and especially heuristic reasoning, are mostly unknown. Dual-process theories make claims about the origins of heuristic processes (e.g., Stanovich, 2004). However, these claims have rarely been tested.

The bulk of research into the development of reasoning processes has been done with children above the age of 10. If we accept that heuristics are basic processes that develop at a young age, we can expect that by early adolescence they have already appeared, making it hard to identify the cognitive processes that played a role in the development of heuristics. In order to be able to examine this question, in Chapter 3 I present three experiments where we looked at a wide range of heuristics, with the participation of children from a broad age range (between the ages of 5 and 16). However, before doing this, in this chapter I review the existing literature on the development of reasoning abilities.

In Section 2.1. I describe some heuristics and biases tasks that have been used to examine children's reasoning. I also describe three developmental dual-process accounts, and also the developmental changes that these theories predict for both logical competence and heuristic use. I also highlight the similarities and differences between the predictions of these dual-process accounts, and identify the factors (e.g., cognitive ability, metacognition) that they consider as the main driving forces of the developmental changes



in reasoning abilities. Finally, I describe an alternative (single process) theory of the development of reasoning heuristics.

Section 2.2. summarizes the literature on conditional reasoning. In contrast to dual-process theorists, who consider contextualisation as a quick, and effortless process which usually hinders normative reasoning, the researchers of conditional reasoning propose that the retrieval and integration of relevant knowledge during reasoning is often effortful, and it is also necessary for logical reasoning, and for hypothetical thinking. Researchers in this field also emphasize the importance of the maturation of inhibitory skills in the development of children's reasoning abilities.

Finally, in Section 2.3. I give a summary of the literature on the development of reasoning skills.

### *2.1. Developmental dual-process accounts, and the expected developmental trajectories of reasoning performance*

In this section I describe some heuristics and biases tasks that have been used to examine children's reasoning. I also describe three developmental dual-process accounts, and also the developmental changes that these theories predict for both logical competence and heuristic use. I also highlight the similarities and differences between the predictions of these dual-process accounts, and identify the factors (e.g., cognitive ability, metacognition) that they consider as the main driving forces of the developmental changes in reasoning abilities.

#### *2.1.1. Research into the development of children's reasoning using heuristics and biases tasks*

Dual-process theories (Evans & Over, 1996; Sloman, 1996; Stanovich, 1999) presuppose two systems (or more recently, two types of process – Evans, 2008) that form

the basis of human reasoning. Type 1 (i.e., heuristic) processes are fast, automatic, effortless, and they can be executed parallel with other processes. They are independent of cognitive abilities, and contextually cued. By contrast, Type 2 (i.e., analytic) processes are slow, sequential, effortful, and related to cognitive abilities and thinking dispositions. An important function of Type 2 processes is to decontextualise (or “decouple”) representations, in order to enable hypothetical thinking (e.g., Stanovich, 1999, 2006). However, most of the time Type 2 processes are merely used for the justification of heuristically cued, contextualised responses (Evans, 2006).

Developmental research on reasoning has been relatively sparse. One possible reason for this might be the implicit assumption that, if adults perform poorly on reasoning and decision-making tasks, children’s performance must be even worse (Klaczynski, 2009). Another general assumption is that heuristic reasoning is simple and it appears early in the course of development, thus most theorists do not expect the heuristic system to change considerably with age (although see e.g., Reyna & Farley, 2006; and Klaczynski, 2009, for exceptions). On the other hand, most theorists agree that explicit cognition develops with age, and this assumption has been tested in numerous studies. As a result of the almost exclusive focus on logical reasoning in developmental studies, there is an illusion that the development of reasoning proceeds from relatively illogical to relatively logical (e.g., Piaget, 1976).

These ideas are also reflected in developmental dual-process theories. In general, these accounts assume that (in some form) both Type 1 and Type 2 processes are available at all points of development — at least after an early age (e.g., Jacobs & Klaczynski, 2002; Kokis, Macpherson, Toplak, West, & Stanovich, 2002). They also agree that the efficiency and prevalence of Type 2 processing increases with development, and Type 2 processing is also associated with increases in cognitive capacity within a given age group (e.g., Handley, Capon, Beveridge, Dennis, & Evans,

2004; Stanovich & West, 2000). However, the evidence concerning age-related changes in Type 1 and Type 2 responding is mixed. Although most evidence indicates age increases in normative responses, age can also be associated with increases in non-normative responses and violations of formal rules of inference. Some studies have found that normative responding increased with age on certain tasks, but this increase was not apparent on other tasks (e.g., Klaczynski, 2001a; Kokis, et al., 2002). Although heuristic responding is expected to decrease with age, some studies have reported the opposite pattern. For example, Webley and Plaisier (1998) using a problem similar to Tversky and Kahneman's (1981) "lost ticket scenario" (see Figure 2.1.) found that older children (between the age of 8 and 12) were increasingly influenced by past investments whereas young children (age 5-6) were not.

Lost money scenario	Lost ticket scenario
<p>Imagine that you have decided to see a play where admission is \$10 per ticket. As you enter the theater you discover that you have lost a \$10 bill.</p> <p>Would you still pay \$10 for a ticket for the play?</p> <p>Yes                      No</p>	<p>Imagine that you have decided to see a play and paid the admission price of \$10 per ticket. As you enter the theater you discover that you have lost the ticket. The seat was not marked and the ticket cannot be recovered.</p> <p>Would you pay \$10 for another ticket?</p> <p>Yes                      No</p>

Figure 2.1. *Lost ticket and lost money scenarios (Tversky & Kahneman, 1981)*

According to traditional economic analysis, past investments should not influence present economic decisions. In Tversky and Kahneman's (1981) *lost money/ticket task*, there are two scenarios which are equivalent in terms of the expected utility and required financial investment. However, in one scenario (the lost ticket scenario) people associate a past investment with the ticket. In the lost money scenario, the past investment is not associated with the ticket. This gives people the impression that the actual price of the ticket is \$10 in the lost money scenario, whereas it is \$20 in the lost ticket scenario. As a result, 88% of participants in the original study were willing to pay for the ticket in the former, as opposed to only 46% in the latter case. The effect is based on making links

between events, that is a kind of contextualisation of the present situation, which seems to be absent in the case of young children (in Webley & Plaisier's, 1997, study 80% of the children indicated that they would buy the ticket in the lost ticket, and 70% indicated that they would buy the ticket in the lost money scenario). Interestingly, on the other end of the age scale, older adults are less influenced by past investments as indicated by their decreased susceptibility to the sunk cost effect as compared to young adults (Strough, Mehta, McFall, & Schuller, 2008).

An increase in heuristic responding with age has been found in a number of other studies as well. Davidson (1995) reported that susceptibility to the conjunction fallacy increased during the elementary school years. Another study with elementary school children (Jacobs & Potenza, 1991) found that children increasingly relied on their own experience and anecdotal evidence as opposed to probabilistic information in making decisions about social situations, although in non-social situations the trend was the opposite (they increasingly favoured probabilistic information with age). Reyna and Ellis (1994) reported that the framing effect (i.e., people are more risky in the domain of losses, than in the domain of gains – see Section 1.1.1.) does not emerge until roughly 10 years of age. Age increases in the tendency to ignore denominators on ratio problems (see below – e.g., Brainerd, 1981), and to draw non-logical "transitive" inferences (e.g., "A is a friend of B. B is a friend of C. Therefore, A and C are friends"; Markovits and Dumas, 1999) have also been reported.

Klaczynski (2001) investigated early and middle adolescents' and adults' performance on three types of task: problems measuring the *if-only* fallacy (Denes-Raj & Epstein, 1994), the denominator neglect (Reyna, Lloyd & Brainerd, 2001), and the sunk cost fallacy (Frisch, 1993). The *if-only fallacy* occurs when behaviours are judged more negatively when in hindsight it seems that a negative consequence could have been easily

anticipated, and therefore avoided, in one of two logically identical, equally unpredictable situations. For example:

*When parking his new car in a half-empty parking, Tom's wife asked him to park in a spot close to where she wanted to shop. Instead, he parked in a spot closer to where he wanted to shop. When he backed out after shopping, the car behind him backed out at the same time, and both cars sustained about \$1000 worth of damage. Robert parked his car in the same parking lot when there was only one parking place available. When he backed out after shopping, the car behind him backed out at the same time, and both cars sustained about \$1000 worth of damage.*

Despite the fact that in these stories Tom and Robert are arguably equally responsible for the accidents, participants typically attribute more foolish behaviour to Tom, because the circumstances make it look like as if his conscious decision “caused” the accident (e.g., Denes-Raj & Epstein, 1994).

An example for the *denominator neglect* task used in the study is the following:

*You are playing a lottery in which you can win \$1000. There are two jars from which you can select a winning ticket. In the first jar, there are only 10 tickets, and 1 of these is the winning ticket. In the second jar, there are 100 tickets and 10 winning tickets.*

*Which jar, if either, would you select from to have a better chance of winning the lottery?*

- a. *The jar with 1 winning ticket*
- b. *The jar with 10 winning tickets*
- c. *It would not matter to me*

Although the proportions of targets in options “a” and “b” are identical, both children and adults often fail to consider differences in denominators and favour the option (“b”) with the greatest absolute number of winning tickets. Interestingly, most participants choosing the heuristic response are aware that this does not make sense statistically (e.g., Pacini & Epstein, 1999).

Finally, the examples below illustrate the *sunk cost* effect.

- A. *You are staying in a hotel room on vacation. You paid \$10.95 to see a movie on pay TV. After 5 minutes, you are bored and the movie seems pretty bad. How much longer would you continue to watch the movie?*
  
- B. *You are staying in a hotel room on vacation. You turn on the TV and there is a movie on. After 5 minutes, you are bored and the movie seems pretty bad. How much longer would you continue to watch the movie?*

Response options for both scenarios were the following:

- a) *stop watching entirely*
- b) *watch for 10 more minutes*
- c) *watch for 20 more minutes*
- d) *watch for 30 more minutes*

e) *watch until the end*

As is the case of the lost ticket scenario (see above), past investments should not influence present economic decisions. Because sunk costs (that is, the money paid for the movie, in this case) are irretrievable, they should be ignored. Thus, decisions in the two situations should be the same. However, participants tend to continue investing in a worthless case (i.e., watch the movie longer), when they invested resources in it already (i.e., when they paid for watching the movie).

In Klaczynski's (2001) study participants were given two problems of each type: two if-only, two denominator neglect, and two sunk cost problems. Each problem had three response options: a normatively correct one, a heuristic one, and a third option which was neither heuristic nor normatively correct. This made it possible to measure heuristic and normative responding partly independently. After each of the six problems, two "framing" instructions were presented. The intent of one instruction ("Think about this situation as you normally would.") was to elicit participants' default manner of processing. The purpose of the other instruction ("Think about this situation from the perspective of a *perfectly logical person*.") was to elicit analytic processing. Instructing participants to think logically increases normative responding (e.g., Denes-Raj & Epstein, 1994; Ferreira, Garcia-Marques, Sherman & Sherman, 2006), presumably because the instructions increase the amount of mental effort that participants invest in solving the tasks. Another possible reason for the effect of instructions is that asking people to think logically sensitizes them to potential conflicts between logic and intuitions (Stanovich & West, 2008; see also Section 1.2.2.).

The main findings of Klaczynski's (2001) study were the following. Normatively correct responding increased with age on all tasks, although it was quite low generally (in the "think as you normally would" frame less than one third of the responses were normatively correct, even in the case of the older participants). Logical instructions

increased normative, and decreased heuristic responding in all age groups. Probably the most interesting finding of this study is that heuristic responding remained stable across age groups in the case of all of the tasks. The reason for this is that early adolescents often chose the neither heuristic nor normative response, but this tendency decreased with age.

In sum, although logical competence, rule use, and normatively correct responding generally increases with development, heuristic use can also increase or remain stable across age groups. This results in “messy” developmental trajectories in children’s performance on heuristics and biases tasks. In the next two sections I am going to describe three developmental dual-process accounts that have attempted to identify the driving forces behind these changes.

### *2.1.2. Developmental dual-process theories*

One dual-process account (e.g., Kokis et al., 2002; Stanovich & West, 2000) in line with traditional theories of cognitive development (e.g., Piaget, 1976) asserts that children’s reasoning becomes more analytical, complex, and abstract with age. Although these theorists acknowledge that some heuristic responses become more common with development, they propose that it might just be the by-product of task characteristics (e.g., studies using tasks that require specific knowledge, such as social stereotypes, that are not available for children under a certain age) rather than a reflection of actual developmental changes in children’s reasoning abilities. Similarly, they argue that if people are not aware of a relevant normative rule, they will not be able to use it, even if they have the necessary cognitive capacity to inhibit heuristic responses (Stanovich & West, 2008). Although these theorists acknowledge that sometimes acquired beliefs can lead to errors and biases, this is restricted to certain types of self-serving “memes” or “memeplexes” (e.g., Blackmore, 1999, Dawkins, 1993), which are also referred to as “contaminated mindware” (Stanovich et al., 2008). These include social rules and stereotypes that members of certain



communities are expected to identify with without questioning or rationally analysing them.

According to these theorists, in addition to relevant knowledge, the connection between age and reasoning performance is mediated by cognitive abilities and thinking dispositions (i.e., how much people are inclined to think analytically and open-mindedly about problems regardless of their cognitive abilities), and normative responding replaces heuristic responses given enough cognitive capacity. They also propose that Type 2 processes can become automatized with practice, leading to a greater coincidence of normative and heuristic responses as people get older. In this view, one of the most critical functions of Type 2 processes is to override Type 1 processing. This is based on two (possibly related) capabilities of Type 2 processing. One of these is to interrupt Type 1 processing, and to suppress Type 1 response tendencies (Stanovich, Toplak & West, 2008). The other one is decoupling (e.g., Stanovich, 1999, 2006). Decoupling is a mechanism that enables people to maintain and manipulate internal representations in the presence of distracting interfering stimuli, which is essential for hypothetical thinking (although Type 2 thinking might be possible without decoupling as well - see e.g. Stanovich & West, 2008).

Recently, Stanovich et al. (2008) suggested a taxonomy of heuristics and biases (see Table 2.1.), and reviewed the literature on reasoning development. The first category that they described was “cognitive miserliness”, which is the tendency to invest as little cognitive effort in reasoning and judgements as possible. This can happen by defaulting to the response options primed by Type 1 processing, such as relying on vivid information, or attribute substitution (i.e., “answering an easy question instead of a hard one” – Kahneman & Frederick, 2002). Another example for cognitive miserliness is serial associative cognition with a focal bias - that is, the tendency to deal only with the most easily constructed cognitive model of a problem (for example, through displaying framing

effects). The next type of bias is “override failure”, when sustained decoupling is not carried out. This type of bias requires the knowledge of the relevant rules and procedures. However, people fail to apply them properly due to the high capacity demands of tasks (e.g., in the case of the belief bias in syllogistic reasoning). When people do not have relevant knowledge available, the source of error is “mindware gaps”, rather than override failure. Mindware gaps are often responsible for fallacies in probabilistic reasoning (e.g., in the case of the conjunction fallacy) and critical thinking. Finally, if people apply biased or inappropriate rules to solve a problem, then the source of the error is “contaminated mindware”. “Contamination” can come from egocentric processing, or “unquestionable” belief systems of certain communities.

Table 2.1. *Stanovich et al.’s (2008) taxonomy of heuristics and biases.*

	Categories of bias	Examples
<b>1. The cognitive miser</b>	a) Default to Type 1 processing	Attribute substitution
	b) Focal bias	Framing effects
<b>2. Override failure</b>	Failure of sustained decoupling	Belief bias, denominator neglect
<b>3. Mindware gaps</b>	a) Missing probability knowledge	Conjunction errors
	b) Failure to consider alternative hypotheses	Wason selection task
<b>4. Contaminated mindware</b>	a) Evaluation disabling strategies	Confirmation bias
	b) Self and egocentric processing	Myside bias

This taxonomy of biases, however, does not imply some sort of developmental sequence. After reviewing the developmental literature on heuristics and biases, Stanovich et al. (2008) concluded that children show every one of the biases that have been identified in the adult literature. They also emphasize that age-related changes in

biases mirror the cognitive-ability related differences within a single age group. For example, there are developmental increases in the avoidance of belief bias, and analytic responding in the selection task, and individual differences in the performance on these tasks are related to cognitive abilities in adult samples. On the other hand, egocentric processing and framing effects do not show developmental changes, and these biases are also unrelated to cognitive ability. Stanovich et al. (2008) are aware of the developmental increase in certain heuristics reported in a number of studies. However, in their view the literature on developmental trends in the case of these tasks is “too inconsistent and sparse to warrant any conclusion at this point” (page 273). In addition, some biases (e.g., the conjunction fallacy) require the knowledge of social stereotypes that children might be unfamiliar with, thus they fail to display these biases due to their lack of relevant social knowledge. In sum, in this approach logical and analytical reasoning is expected to increase with age due to increases in cognitive capacity and knowledge of relevant normative rules and procedures, although some exceptions are possible.

Another line of research (e.g., Jacobs & Klaczynski, 2002; Klaczynski, 2001b; Klaczynski, 2009) proposes that cognitive capacity and abstract reasoning competence are not good predictors of reasoning and decision making in real-life situations. Social, motivational, and affective influences, as well as prior beliefs, greatly affect the way people reason about problems. They also emphasize that an increasing number of heuristics are acquired over the course of development, and the use of these heuristics becomes more prevalent with age. This phenomenon appears to be linked to increases in knowledge of stereotypes and belief systems (e.g., religion). On the other hand, Type 2 processing also develops. This approach, however, considers metacognitive abilities (i.e., monitoring, evaluating, and controlling information processing), rather than computational capacity *per se*, as the key determinant of normative performance. They propose that metacognition starts to shape reasoning around mid-adolescence (Klaczynski & Cottrell,

2004). Metacognition is mostly independent of cognitive abilities, and it is related to the tendency to inhibit heuristically cued responses, and also to consider alternatives (Klaczynski, 2005). However, this ability is not always used or fully developed even in adults. In fact, adults rely on heuristic processing most of the time. Given that people prefer cognitively economical strategies, the development of analytic competence must be accompanied by developments in tendencies to consciously utilize these competences (see also Stanovich and West, 1998, 2000; Stanovich, 1999). That is, analytic reasoning critically depends on the acquisition of dispositions to control impulsive actions, and on effortful processing (Klaczynski, 2009).

In this approach, normative reasoning does not necessarily increase with development. What is developing is cognitive flexibility, and the variability of heuristics and cognitive strategies. One reason for this is that the repertoire of heuristics becomes more diverse and more easily activated with age, as memories accrue and as conscious strategies are transformed into automatic procedures. That is, people acquire more heuristics with development, and these heuristics can be activated more easily with age. Additionally, children sometimes seem more rational than adults, because they follow rules more rigidly, whereas adolescents and adults integrate rule understanding with contextual considerations. Consequently, the decisions they make are more likely to deviate from the relevant normative rules (Klaczynski, 2007; Kuhn, 2001).

In summary, dual-process theorists claim that reasoning performance depends on a small number of factors, including cognitive capacity, metacognitive skills, dispositions (related to the tendency to invest cognitive effort), and knowledge/beliefs. In one approach (e.g., Kokis et al., 2002; Stanovich et al., 2008) the interaction between these factors is expected to result in an increase in normative responding with age, and they also predict a relationship between cognitive capacity and normative responding. According to another approach (e.g., Jacobs & Klaczynski, 2002) developmental patterns are less clear, and

cognitive capacity is not necessarily a good predictor of reasoning performance. However, normative responding is expected to increase after mid-adolescence due to the development of metacognitive abilities.

*2.1.3. The links between the development of memory and reasoning: The fuzzy-trace theory.*

A third developmental dual-process account is the fuzzy-trace theory (e.g., Brainerd & Reyna, 1992, 2001; Reyna & Ellis, 1994; Reyna & Farley, 2006). Fuzzy-trace theorists claim that the two processing modes are based on two different memory systems: one that contains precise “verbatim” representations (which preserves the surface information of the stimuli, and consequently it is more phenomenological – it corresponds to the actual, immediate experience of the stimuli), and one that consists of “fuzzy gists” which preserve the underlying meaning of a task (i.e., how we understand and interpret our experience). The latter is the primary (default) system which generates automatic inferences based on the relational and semantic properties of problems. This system flexibly changes with age, as individuals extract new meanings and structures all the time from the stimuli they encounter. As verbatim representations rapidly fade, judgment and decision making usually relies on gist representations (Reyna & Farley, 2006). The secondary system is only used for solving problems which demand the application of specific rules to precise details (e.g., as in mathematics). Importantly, not only the contents of long-term memory change with development, but also the reasoning processes based on these fuzzy traces. Because of this, reasoning performance cannot be reliably linked to cognitive capacity. Fuzzy-trace theory differs from other dual-process models in its focuses on levels of rationality and intuition as advanced reasoning (Reyna & Farley, 2006). Fuzzy-trace theory also emphasizes two types of “quick and easy” thinking: one, which is

similar to the heuristic processing as described by other dual-process theories, and another one, which is a form of higher-order reasoning based on gist.

The difference between the two types of quick and effortless reasoning lies in the ability of experts to quickly react to a small number of relevant cues (ignoring verbatim detail), whereas impulsive and inexperienced decision-makers react to misleading or irrelevant cues. For example, more knowledgeable physicians process fewer dimensions, and do this more qualitatively than do those with less knowledge and training. However, their decisions are more accurate (Reyna & Farley, 2006). Another example is developmental differences in risk-taking. Experimental evidence indicates that young children multiply probabilities with the number of prizes, that is, they quantitatively combine two dimensions (e.g., Schlottman, 2000). By contrast, older children and adults focus on a single dimension: outcomes (e.g., Reyna & Brainerd, 1995), making a more mature, risk-averse decision.

The dissociation of the two systems is demonstrated by: (a) parallel storage of verbatim and gist traces (that is, we can remember surface characteristics of stimuli that we encountered, for example, font size and colour of a sentence that we have read, and independently of this we can also recall the meaning of the sentence); (b) dissociated retrieval of verbatim and gist traces; (c) differential survival rates for verbatim and gist traces (i.e., verbatim representations diminish with time, whereas “false memories” based on the gist increase); (d) retrieval phenomenology (i.e., the retrieval of the actual stimuli together with their perceptual properties – *true recollection* vs. the retrieval of the meaning of those stimuli – *phantom recollection*); and (e) developmental variability in verbatim and gist memory. There is evidence that the accuracy of children’s memory for problem information and the accuracy of their solving the same problems are dissociated. In addition, age-related changes in one type of memory are independent of changes in the other type of performance. In the preschool-to-young-adult age range both verbatim and

gist retrieval increases. The general pattern is that initial improvements in verbatim memory, during the preschool and early elementary-school years, are more pronounced than subsequent improvements (see Reyna, 1996).

One way of studying the development of memory is to look at the spontaneous use of category clustering in free recall. Most children do not spontaneously do this before adolescence (e.g., Bjorklund & Jacobs, 1985). Another way of investigating verbatim and gist memory is by using the DRM paradigm (the DRM acronym is based on the names of Deese, 1959; and Roediger & McDermott, 1995, the developers of the procedure). It is a word learning paradigm in which subjects are presented with lists of words in which every word is an associate of a critical non-presented word. For instance people are presented with *bed, rest, awake, tired, dream, snooze, blanket, doze, slumber, snore, nap, peace*, but are never presented with the word that is related to all of these words *sleep*. The DRM paradigm produces high levels of false recognition and false recall of the critical non-presented word. Using the DRM paradigm, Brainerd et al., (2004) found that true recollection showed an increase between the age of 7 and 14 years, and the increase was more marked between 11 and 14 than between 7 and 11. By contrast, for phantom recollection, the increase was more marked between 7 and 11 than between 11 and 14. In the DRM paradigm, although the reported information is false, it is congruent with the gist of experience.

This is not to say that false memories in general are more common in adolescence and adulthood than in childhood. A repeated finding about spontaneous false memories (e.g., remembering drinking a coke rather than water in a restaurant) is that they become less common between early childhood and young adulthood (for a review, see e.g., Reyna, 1996). When people recall past experiences false and true memories might emerge mixed together and they can create equal feelings of confidence. On the other hand, clear verbatim memories can be used to suppress false memories. Thus, older children and

adults are more likely to reject false memories because of their superior retention of verbatim traces over time. To put it in another way, verbatim processing supports both the acceptance of true events and the rejection of false ones that preserve the meaning of experience, whereas gist processing supports the acceptance of both true and false events. The interactions between the two systems can lead to U-shaped and inverted U-shaped patterns of development (as discussed by Brainerd, 2004).

If we compare the claims of the fuzzy trace theory with other developmental dual-process theories, we can conclude that although these theories show marked differences, there is still some consistency in the claims they make (other than proposing that two separate processes form the basis of human reasoning). First of all, these theorists agree that both types of process are available at all points of development, although they go through changes, especially between early childhood and adolescence (in this review I did not discuss the changes in reasoning abilities at old age, but they more or less show the reverse of the changes in childhood – see e.g., DeNeys & Van Gelder, 2008). These theorists also agree that heuristics are independent of cognitive capacity, and that they work outside conscious control. In addition, they propose that Type 1 (i.e., quick, contextualised, automatic) processing is the default, but slow and effortful Type 2 processes can override or suppress them. Finally, they all agree that Type 2 reasoning requires the decontextualization of problems (by means of decoupling, metacognitive operations, or by relying on the actual stimuli, and suppressing automatic inferences and personal interpretations of the problem).

#### *2.1.4. An alternative (single process) approach to the development of reasoning heuristics: the intuitive rules research program*

In this section I am going to briefly summarize a single system account of reasoning development, the *intuitive rules research program* (Stavy & Tirosh, 2000). In two recent



publications Osman (2004) and Osman and Stavy (2006) criticized dual-process theories on the basis that they cannot account for typical patterns of reasoning performance in developmental samples. The intuitive rules approach deals with the development of certain heuristics (i.e., intuitive rules) which are defined as self-evident and self-consistent cognitions (based on Fischbein, 1987). These rules are retrieved without conscious intention and they are activated by certain aspects of tasks (note the similarity with the ideas of default-interventionist dual-process theories and the fuzzy-trace theory). A distinctive feature of this theory is a focus on the importance of saliency, and a distinction between bottom-up and top-down saliency. Salient stimuli are arousing, they capture attention, and behavioural resources are preferentially directed toward them (Osman & Stavy, 2006). *Bottom-up saliency* is usually based on the similarity between items in a task. The more an item differs from other elements of a task, and the more easily a response based on that stimulus can be generated, the more salient it is. By contrast, *top-down saliency* is based on the number of times a particular stimulus has been experienced in a learning situation. If a stimulus is strongly associated with a response, it will generate a response in novel situations, regardless of its relationship to other information provided in the task.

Another interesting aspect of the intuitive rules approach is the distinction between implicit, explicit and automatic forms of reasoning. In dual-process theories implicit and automatic reasoning would both be categorised as Type 1 processes. By contrast, according to the intuitive rules approach implicit and automatic processing correspond to the two extremes of the same continuum. In this view, implicit reasoning is based on weak representations that are not stable, and as a result they influence participants' behaviour without arousing the feelings of intention and conscious awareness. However, these representations are still capable of influencing explicit processes through priming. Explicit representations are stable, strong and distinctive, and as a result, they can be consciously

controlled and manipulated in working memory, and are available for declarative knowledge. Finally, automatic representations are also strong and stable, and thus accessible for consciousness. However, they became so strong through repeated activation that they are hard to control or modify (see Cleeremans & Jimenez, 2002; Osman & Stavy, 2006). Automatic representations are considered to be the final products of cognitive development, and they form the basis of skill-based reasoning, and mental flexibility. Although the representations change with individuals' experiences, the ways that they can be utilized remain stable with development, and for this reason implicit, explicit and automatic representations are attributed to a single system.

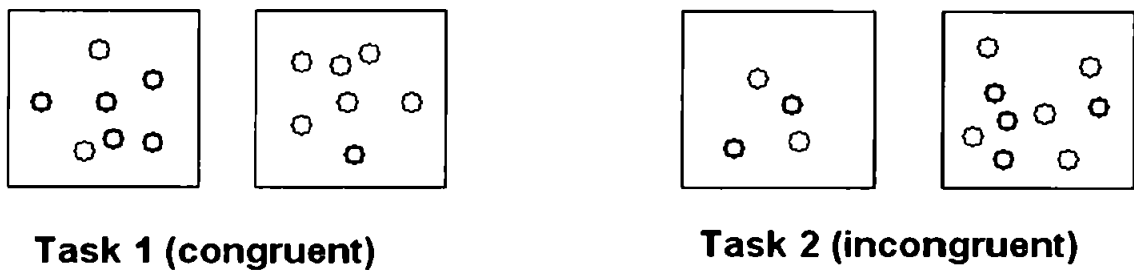


Figure 2.2. Examples for Brecher's (2005) ratio bias task.

An example for a task which elicits bottom-up saliency effects is the *ratio bias problem* (e.g., Brecher, 2005). In this task the intuitive rule “more A – more B” is invoked. In Brecher's experiment participants were presented with a probability task in which they had to decide which one out of two boxes would give them a better chance of drawing a black counter from a mix of black and white counters (see Figure 2.1.). In this task the salient feature is the number of black counters, which is also affected by the proportion of black and white counters. In congruent tasks the box that contains more black counters also contains black counters in a higher proportion than white counters. By contrast, in the case of incongruent tasks the box that contains more black items contains black counters in a lower proportion. Task difficulty was also manipulated based on the difference between the number of black and white counters in the two boxes.

Given that the salient feature is the number of black counters, this task activates the “more A - more B” intuitive rule. Babai, Brecher, Stavy & Tirosh (2006) found that participants responded to congruent tasks more quickly and more accurately than to incongruent tasks. However, response speed and accuracy were also affected by the perceptual discriminability between the number of black and white counters. This indicates that the status of bottom-up salient task stimuli is dependent on their relationship to other presented stimuli. In the case of bottom-up saliency participants are unaware of how the relevant intuitive rule got activated, and thus they are unable to control it. This is not to say that they are unaware of the perceptual stimuli that invoked the rule, as these are salient, and as such, are in the focus of attention.

By contrast, top-down salient stimuli activate relevant rules automatically, which means that individuals are consciously aware of them, and they also possess meta-knowledge about them. Osman and Stavy (2006) refer to the “same A same B” rule as an example for an intuitive rule based on top-down saliency. This is also an example for how the same rule can be applied on range of different tasks, sometimes leading to correct, other times to incorrect responses. For example, Mendel (1998) presented students with a problem in which two rectangles were shown. Students were told that the second rectangle is a modified version of the first one, where the length of the rectangle was decreased by 20% and the width was increased by 20% (see Figure 2.2.). Students were asked about the perimeters of the two rectangles (i.e., whether they were equal, or whether it was longer in the case of one of the rectangles). In this experiment participants had knowledge about the relevant rule (i.e., how to compute the perimeter of a rectangle), and the available perceptual information was also in line with the rule (i.e., that the perimeter of rectangle 1 was longer). However, instead of relying on either of these, over 70% of participants applied the “same A - same B” rule which is related to the acquisition and stabilization of the proportion scheme (see Fischbein & Schnarch, 1997). That is, they claimed that the

perimeters of the two rectangles were equal, because “adding 20% and removing 20% equals to no change”. The fact that the justification in itself makes sense explains why it is so hard to eliminate biases based on top-down saliency. On the other hand, young children tend to respond to similar questions correctly, because they do not possess relevant knowledge, and thus they rely on perceptual information (and apply the “more A – more B” rule in the present case). This example also shows how saliency is influenced by cognitive development and children’s knowledge. In addition, this also shows that salient features in themselves do not belong to a certain type of cognitive process, as it depends on the reasoners’ experiences what sort of response they will generate based on given stimuli.

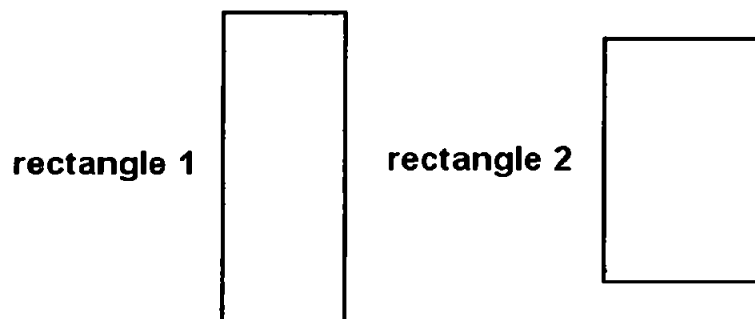


Figure 2.3. *Illustration of Mendel's (1998) rectangle task.*

In the example of the ratio bias task reducing the discriminability between salient and non-salient features leads to conflict at the perceptual level. However, it is also possible to create conflicts at the level of explicit rule use by presenting students with counterexamples. Finally, two intuitive rules can get activated simultaneously by different aspects of the same problem, and these can cue conflicting responses. As these examples show, the fact that children might experience conflict is, in itself, no evidence for the existence of two separate reasoning systems, or two separate types of process (Osman & Stavy, 2006). In the case of many tasks there are U-shaped developmental patterns, where young children and adults respond similarly, although for different reasons. U-shaped

developmental patterns are usually considered to provide evidence for the interaction between two systems (i.e., the gist and the verbatim system in the case of the fuzzy-trace theory, and automatic retrieval and effortful inhibition processes in the case of conditional reasoning – see Sections 2.1 and 2.2.). These patterns, however, are explained by changes in children’s perceptions of tasks, and in their tendency to apply certain intuitive rules with age.

The application of intuitive rules can be both positively and negatively related to cognitive capacity. For example Babai and Alon (2004) reported that the application of the “more A-more B” rule decreased with cognitive capacity. On the other hand, the use of the “same A-same B” rule (i.e., the equiprobability bias) and the “everything can be divided” rule increased with cognitive capacity. Given that the application of these rules can lead to both correct and incorrect responses depending on the task, there is no set relationship between cognitive capacity and normative responding. In addition, responding on a particular task can show U-shaped and inverted U-shaped patterns which is uninterpretable with regard to the monotonic increase of cognitive capacity throughout childhood and adolescence.

In summary, the intuitive rules research program claims that children’s reasoning performance is based on a single system. Representations gain strength with repeated exposure, and the procedures based on them can become conscious and eventually, automatic. With the strength of representations the awareness of them, as well as the ability to control them changes too. Finally, reasoning processes, as well as the intuitive rules that are implemented in a given situation change with children’s experiences, their explicit knowledge, and the way problems are presented. The intuitive rules program can give an explanation of some typical developmental patterns in heuristic use (i.e., inverted U-shaped patterns), and they can also explain the inconsistent relationships between heuristic use and cognitive capacity. On the other hand, there are still some unresolved issues. For example,

can these rules be generalized to problems outside of the domain of learning about science? Another question is whether we can predict children's performance on novel tasks. It seems likely that a person using an intuitive rule erroneously in one context will be able to use the same rule to generate a correct response in the case of another task, where it is appropriate to use it. However, whether children employ a particular rule in the case of a task, depends on the context of the task as well, and children's previous experience with the task. Thus, it seems hard to apply the intuitive rules approach outside of the school environment where children's knowledge and their exposure to certain problems is similar within groups.

## *2.2. The role of memory retrieval and contextualisation in logical reasoning: The case of conditional inferences*

This section gives a summary of the literature on conditional reasoning. In contrast to dual-process theorists, who consider contextualisation as a quick, and effortless process which usually hinders normative reasoning, the researchers of conditional reasoning propose that the retrieval and integration of relevant knowledge during reasoning is often effortful, and it is also necessary for logical reasoning, and for hypothetical thinking. Researchers in this field also emphasize the importance of the maturation of inhibitory skills in the development of children's reasoning abilities.

### *2.2.1 The role of the activation and integration of real-life knowledge in everyday conditional reasoning*

Although, as we have seen, dual-process theorists agree that contextualised reasoning (the retrieval and integration of relevant knowledge with the content of problems) is quick and effortless, this idea is not shared by all researchers of reasoning. Conditional reasoning is the ability to reason on the basis of "if-then" statements, and it

is considered to be one of the cornerstones of human cognition. However, the conditional inferences that we draw are not always logically appropriate (Evans & Over, 1996; Manktelow, 1999). One example is the modus ponens (MP) inference. Let's suppose that we are given the following statement: "If it rains, James gets wet." and the additional information that "It rains." The logical inference to draw is that in this case James gets wet. However, people sometimes fail to draw this conclusion (e.g., Byrne, 1989; Cummins, 1995; De Neys, Schaeken, & d'Ydewalle, 2002; Markovits & Quinn, 2002; Thompson, 1994). The reason for this is that they can think of *disabling conditions*, which prevent James from getting wet, such as that he might have an umbrella with him. In this example real-life knowledge blocked a valid inference. On the other hand, background knowledge can also be used to reject an invalid inference. For example, let's suppose again that "If it rains, James gets wet.", and also that we know that "James gets wet." We might draw the invalid (affirmation of the consequent – AC) inference that "It rains." However, we can again rely on our real-life knowledge and come up with an alternative cause, such as that "James has a bath."

The effects of retrieved counterexamples on adults' reasoning are well established (for a view see Politzer & Bourmaud, 2002). The outcome of the counterexample search directly determines the extent to which inferences will be drawn. The more counterexamples retrieved, the less likely that a conclusion will be accepted (De Neys, Schaeken, & d'Ydewalle, 2003; Liu, Lo, & Wu, 1996). Consequently, we could expect that people who can retrieve more counterexamples will draw less MP and AC inferences. It is also known that the effective retrieval of information from long-term memory is related to working memory capacity in the case of both children and adults (e.g., De Neys, Schaeken, & d'Ydewalle, 2002, 2005; Janveau-Brennan & Markovits, 1999). According to Barrouillet, Markovits and Quinn (2001) during reasoning the memory structures that are associated with the content of the problem are automatically accessed

and specific elements within this structure are activated (note the similarity with the assumptions of default-interventionist dual-process accounts, and the fuzzy-trace theory). Moreover, the less strongly a given element is associated with the context of the task, the more effort (i.e., cognitive resources) and time are necessary for them to be retrieved. In line with this, it has been also shown that higher ability people are more inclined to retrieve elements from their long-term memory that are more remotely associated with a given context (Verschueren, De Neys, Schaeken, & d'Ydewalle, 2002). In addition, the efficiency of counterexample retrieval suffers from dual-task loads, which also indicates that working memory is involved in this process (De Neys, 2003). This is presumably because using counterexamples requires both an active controlled search process, and the integration of the premises with the retrieved counterexamples. When working memory capacity is burdened by preload some reasoners do not engage in the demanding process of consulting their background knowledge (Verschueren, Schaeken, & d'Ydewalle, 2005).

If conditional reasoning only depended on the successful retrieval and integration of real-life knowledge, this would mean that higher ability people would be more “logical” on AC problems, but less “logical” on MP problems than lower ability people. This is not quite the case. Counterexamples can be selectively retrieved or the retrieval process can be blocked depending on the reasoner’s cognitive capacity, and whether they find the invited inference logically valid. Although the underlying processes are not fully understood, it has been found that for conditionals with many disablers it is more likely that disablers will be retrieved inappropriately. In addition, high ability adults (i.e., people with higher working memory capacity) are more able to prevent the inappropriate retrieval of counterexamples than lower ability adults (e.g., De Neys, Schaeken, & d'Ydewalle, 2005; Markovits & Doyon, 2004; Verschueren, Schaeken, & d'Ydewalle, 2005). This corresponds to the idea that the main purpose of working memory is to



maintain information in the presence of distracting stimuli (see e.g., Kane & Engle, 2002).

In sum, adults' conditional reasoning seems to depend on the ability to selectively retrieve counterexamples depending on the logical status of the invited inference (e.g., De Neys et al., 2005; Quinn & Markovits, 2002). High ability people are more able to retrieve counterexamples, and they are also more able to discriminate between situations when counterexamples are needed, and when they are not. This difference in the ability to selectively retrieve knowledge depending on people's cognitive resources results in different patterns of relationship between cognitive ability and endorsement rates for MP and AC. AC inferences can be successfully blocked by the retrieval of counterexamples, so there will be a linear relationship between cognitive ability and drawing AC inferences in the case of everyday conditionals. As higher ability people are more likely to successfully retrieve counterexamples, they will be more likely to reject the fallacious AC inference. For MP the pattern is more complex. Low ability people who are not very much able to generate counterexamples will be likely to (correctly) draw MP inferences. People with higher cognitive ability will be able to retrieve counterexamples, so they will be more likely to (incorrectly) block this inference. Finally, people with the highest cognitive ability will be able to recall counterexamples, but they will also be able to block the retrieval of counterexamples, and reason on the basis of logic. This results in a U-shaped relationship between cognitive ability and the endorsement of MP (De Neys, Schaeken & d'Ydewalle, 2003).

### *2.2.2. Developmental research on conditional reasoning.*

Everyday conditional reasoning in adults seems to be based on a complex interplay between real-life knowledge, cognitive capacity and the understanding of logical necessity. Developmental studies have found that the retrieval of background knowledge

plays a crucial role in children's conditional reasoning as well. Janveau-Brennan and Markovits (1999), for example, reported that the better elementary school children were at generating alternatives for a set of conditionals in a pretest, the less they tended to accept the AC inference in a subsequent reasoning task.

As cognitive development increases children's knowledge base and contributes to more efficient memory retrieval (e.g., Kail, 1992), we can expect that the effect of real-life knowledge will be more pronounced as children develop. In line with this prediction, a number of studies with preadolescents have found that when children reason with familiar everyday conditionals, acceptance rates of the AC and MP inferences show a steady decline between the age of 6 and 11 years of age (e.g., Janveau-Brennan & Markovits, 1999; Markovits, 2000; Markovits, Fleury, Quinn, & Venet, 1998; Markovits et al., 1996).

Markovits and Barrouillet (2002) suggested that the development of conditional reasoning is based on an interaction between retrieval and inhibition. However, these are not attributed to the same mechanism (i.e., working memory). The ability to retrieve counterexamples develops earlier, resulting in an initial decline in AC and MP inferences. However, after the onset of adolescence, children's inhibitory capacities become sufficiently strong to start overriding the impact of disablers. This leads to an increase in MP acceptance from early adolescence to late adolescence (e.g., Barrouillet, Markovits, & Quinn, 2001; Klaczynski & Narasimham, 1998; Markovits, 1995; Markovits & Vachon, 1989).

A direct comparison of elementary school children and adolescents (between the age of 10 and 18) showed a U-shaped MP acceptance trend from preadolescence to later adolescence; and a stable decrease in AC acceptance from preadolescence to later adolescence (De Neys & Everaerts, 2008). In this study the availability of counterexamples was systematically manipulated, which made it possible to demonstrate

that inference acceptance in the individual age groups depended on the ease with which alternatives could be retrieved. On the basis of this we can refute an alternative explanation for the U-shaped developmental pattern in drawing MP inference. From a Piagetian point of view it would be possible to argue that the reason why children start to inhibit the effect of real-life knowledge on MP inferences around the age of 12 is that this is the time when they enter the formal operational stage. However, if adolescents' reasoning would be determined purely by their ability to understand formal logic, the availability of disablers should not have an effect on their conclusions. Other studies have showed that young children (under the age of 12) are able to rely on logical rules during reasoning when they are explicitly instructed or trained to do so (e.g., Barrouillet et al., 2001; Markovits & Vachon, 1989; Morris, 2000). Thus, the developmental patterns in conditional reasoning are unlikely to be simply based on changes in children's ability to grasp the concept of logical necessity.

Although logical, hypothetical thinking might not be the Type 2 process that people most often engage in, it is still a very important ability that forms the basis of understanding modern science, for example. It also gives further insight into the role of retrieval and inhibition in reasoning, and also the development of inhibitory control. When we use empirically false (contrary to fact) premises, and ask people to reason on the basis of the premises, disregarding their beliefs, then they are invited to engage in hypothetical thinking. Let's suppose that "if a feather is thrown at a window, then the window will break" is given, accompanied by instructions to reason on the basis of the premise regardless of it is being true or false. When the MP inference "if a feather is thrown at a window, will the window break?" is made, only if we can disregard our real-life knowledge about feathers and windows, will we be able to respond with the logically correct inference "the window will break." However, if we rely on our knowledge, we will respond that "the window will not break". On the other hand, giving the correct

uncertainty response to the AC inference (“if a window is broken, was a feather thrown at it?”) requires access to information about alternative ways of breaking windows (e.g., throwing a brick). The results of Markovits (1995) and Simoneau and Markovits (2003) indicate that reducing the use of information that feathers do not break windows tends to be accompanied by a reduction in the use of information about other ways of breaking windows in both children and adolescents. Specifically, logical instructions with younger reasoners increase the proportion of logically correct responding to contrary to fact MP and MT forms and also increase the proportion of logically incorrect responses to the AC and DA forms. Markovits and Doyon (2004) explained this by suggesting that children use their ability to control interference in a global way, that is, they are unable to selectively activate relevant knowledge, whilst suppressing irrelevant associations. These results show that abstract reasoning can be facilitated by realistic context, that is, the development of abstract reasoning seems to be reliant on appropriate access to empirical knowledge (Venet & Markovits, 2001).

Similarly, studies with adults have shown that activating information retrieval processes after logic instructions increases the tendency to make empirically appropriate, but logically invalid inferences (Markovits & Potvin, 2001). This is due to a failure to control the inappropriate activation of information (i.e., because of susceptibility to interference). This is supported by the finding that those adults who were more susceptible to interference in contrary to fact conditional reasoning were also more susceptible to interference on a different (negative priming) task (Simoneau & Markovits, 2003). Although the instructions to reason logically do not explicitly indicate the necessity to inhibit real-life knowledge, the findings with both children and adults show that the inhibition of knowledge is the general response to such instructions. Another study also demonstrated the separate role of working memory and inhibition in reasoning. Handley et al. (2004), using both belief-neutral, believable and unbelievable conditional

reasoning problems, found that indices of belief bias and logical reasoning on belief-based problems were predicted independently by both working memory and inhibition. In contrast, logical reasoning on belief neutral problems was predicted by working memory alone. A comparison between the developmental and the adult data shows an increasing selectivity of inhibition with age. This is considered as a very important aspect of cognitive development, and also a very important prerequisite of sound reasoning (Markovits & Doyon, 2004).

In sum, conditional reasoning with both everyday and contrary to fact premises is based on an interplay between retrieval and inhibition (e.g., Handley et al. 2004; Janveau-Brennan & Markovits, 1999). Differences in the development of these two abilities (De Neys & Everaerts, 2008; Markovits & Barrouillet, 2002), as well as the relationship between susceptibility to the intrusion of real-life knowledge and performance on an inhibition measure (Simoneau & Markovits, 2003) indicates that retrieval and inhibition are separate abilities. This corresponds to the notion that executive functions (working memory, inhibition and set-shifting) are not a unitary construct, although they are highly related (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000).

### *2.3. Research on the development of children's reasoning abilities - Summary*

In this chapter I have described a number of tasks which (together with other tasks that I introduced in Chapter 1) have been used to investigate children's reasoning abilities. I have also described the developmental trajectories of children's performance on these tasks. Although traditional theories of cognitive development (e.g., Piaget, 1976) assume that children's reasoning proceeds from relatively illogical to relatively logical (which has been termed "the illusion of replacement"- e.g., Reyna & Ellis, 1994), the actual developmental patterns seem to be more complex. For example, some heuristics and biases

remain stable with development (see e.g., Klaczynski, 2001), or they even increase with age (e.g., Davidson, 1995; Jacobs & Potenza, 1991).

In Section 2.2. I have introduced three developmental dual-process accounts of reasoning. All of these theories propose that human reasoning is based on two separate systems: one which is effortless and automatic (Type 1), and another one which is rule-based and requires conscious attention (Type 2 processes). Moreover, all of these theorists agree that both types of process are available at any point of development, although these processes can also go through age-related changes (see Table 2.2. for the predicted developmental patterns, and the cognitive changes that elicit these developments). They also agree that the main function of Type 2 processing is to suppress the automatic response tendencies of Type 1 processes. According to one dual-process account (e.g., Stanovich et al., 2008) the main driving force behind age-related changes in reasoning is an increase in cognitive capacity which has a general effect on reasoning performance. By contrast, Klaczynski et al. (e.g., Klaczynski & Cottrell, 2004) emphasize the role of real-life experiences in reasoning. Klaczynski and colleagues also claim that different experiences in different domains of knowledge can result in uneven levels of performance across different types of task. Moreover, they predict that children's reasoning performance changes considerably during adolescence due to the emergence of metacognitive abilities, and generally cognition becomes more flexible and versatile with development. In contrast with the above two accounts, the fuzzy-trace theory (e.g., Brainerd & Reyna, 1992, 2001) proposes that the maturation of cognition is characterized by the increasing dominance of the primary (gist-based) system, which consists of a collection of people's relevant experiences, and makes it possible for people to reason quickly, effortlessly and normatively in an increasing range of situations. Although Type 2 reasoning can be used to suppress Type 1 processing, these theorists emphasize that this process is slow, often ineffective, and error-prone.

Table 2.2. *Developmental trajectories predicted by developmental dual-process theories, and the cognitive foundations of these changes.*

Developmental theories of reasoning	Normative responding	Heuristic responding	Development is driven by...
Stanovich et al.	Steadily increases with age, or remains stable	Decreases or remains stable with age	increases in cognitive ability although thinking dispositions and relevant knowledge play a role as well
Klaczynski et al.	Can increase, decrease or remain stable with age, although it is expected to increase during adolescence	Can increase, decrease or remain stable with age	the development of metacognitive abilities, and thinking dispositions, together with real-life experiences
Reyna et al.	Increases with age, or U-shaped pattern	inverted U-shaped pattern	changes in representations as a result of experiences; increasing reliance on gist processing

Although dual-process theorists agree that contextualised reasoning (the retrieval and integration of relevant knowledge with the content of problems) is quick and effortless, this idea is not shared by the researchers of conditional reasoning. In this view (at least in the case of certain types of problem) contextualisation is necessary for normative reasoning, it is positively related with cognitive capacity (especially working memory capacity), and contextualisation also plays an important role in the development of hypothetical thinking. On the other hand, they agree with dual-process theories that the retrieval of real-life knowledge has to be blocked in some cases, in order to reason in line with standard logic, and the ability to suppress the retrieval of counterexamples is related to cognitive capacity (specifically, inhibition). Similarly to Klaczynski (2009) they expect a change in reasoning performance during adolescence as a result of the maturation of inhibitory processes. They also share some ideas with the fuzzy-trace theory, in that they propose that when reasoning

performance is based on a retrieval and inhibition interaction, developmental patterns will show (inverted) U-shaped patterns.

In Section 2.1.4. I also introduced an alternative, single process account of the development of reasoning heuristics, the intuitive rules approach (e.g., Osman & Stavy, 2006). These theorists predict similar developmental patterns as the fuzzy trace theory. However, they claim that there is no need to presuppose two systems in order to be able to explain U-shaped and inverted U-shaped developmental patterns, or the fact that people sometimes experience a conflict when they reason about problems.

In Chapter 3 I am going to describe some studies which were done with children from a broad age range (between the age of 5 and 16). Besides a wide range of reasoning tasks (adapted from the heuristics and biases literature), in these studies we also administered a number of cognitive ability measures (i.e., measures of IQ, inhibition, verbal and non-verbal working memory, fluid intelligence, and set-shifting). The purpose of these studies was: 1. to investigate developmental patterns from age 5 to 16 on a range of typical heuristics and biases tasks; 2. to see how developmental patterns of the performance on these tasks are related to changes in cognitive abilities; and 3. to contrast the predictions of different theories and to see which one (if either) describes our empirical findings most appropriately. As the development of reasoning heuristics is an under-researched area, and the results of existing studies are inconsistent, the studies that I describe in Chapter 3 were conducted with an exploratory purpose. However, I will also contrast the findings of these studies with the predictions of existing theories of reasoning development.



## Chapter 3. The development of heuristic reasoning in childhood and adolescence (Experiments 1-3.)

### *3.0. Introduction to Chapter 3.*

In this chapter I describe three experiments which were carried out with children from a wide age range. The studies used typical heuristics and biases tasks (similar to the tasks introduced in Chapters 1 and 2). The purpose of these studies is: 1. to determine the developmental trajectories of some well-known heuristics and biases tasks between early childhood and late adolescence (i.e., between the age of 5 and 16); 2. to investigate the relationship between these changes, and some aspects of cognitive functioning (most importantly, general intelligence, working memory, and executive functioning); and 3. to compare the predictions of different developmental theories of reasoning, and to see how well they describe the empirical findings of these studies.

The first experiment investigates the performance of children between the age of 5 and 11 on a number of heuristics and biases tasks. In this experiment we also measured children's general intelligence (using a short form of the WISC), and their working memory. In the second experiment we used the same tasks, but the participants were a group of children between the age of 12 and 16, and in addition to measuring their general intelligence and working memory, we also measured their inhibition skills and set-shifting ability (the latter three tasks were designed to assess the three main aspects of executive functioning – see Miyake et al., 2000). We employed these new tasks because many reasoning theorists predict a sudden improvement in reasoning performance around adolescence, which is possibly related to the maturation of executive functioning.

Finally, the last experiment was also carried out with adolescents (between the age of 12 and 16), but we used a new set of tasks. Moreover, each type of task was presented in two different versions: one, where intuitions and normative considerations pointed to different responses (i.e., conflict tasks, or experimental tasks – similar to the tasks used in

Experiments 1 and 2), and another version, where there was no conflict between intuitions and normative rules (i.e., non-conflict, or control tasks). The purpose of using both conflict and non-conflict tasks was to get a clearer picture of how much adolescents are inclined to rely on certain normative rules (for example how much they consider base rate information) when these are not in conflict with some more salient information (for example, stereotypes activated by a person's description). In this experiment we also used a new set of individual differences measures. Apart from children's general intelligence and verbal working memory scores, we also measured their analogical reasoning ability (using the Raven test, picture analogy problems, and scene analogy problems).

*3.1. Experiment 1: Changes in reasoning performance on some typical heuristics and biases tasks between the age of 5 and 11.*

Dual-process theories of reasoning (e.g., Evans & Over, 1996; Kahneman & Frederick, 2002; Stanovich, 1999 – see also Chapter 1 for a more detailed review and evaluation) presuppose two distinct processing modes. System 1 processes are fast, automatic, effortless, independent of cognitive abilities, and contextually cued, whereas System 2 processes are slow, effortful, related to cognitive abilities and dispositions, and context independent. Most of the time, the two systems offer the same solution to problems (Klaczynski, 2001; Moshman, 2000; Stanovich, 1999). This is because heuristics are very useful and effective when used in the right situation. In addition, System 2 processes can become automatized through practice; thus, they can function in the manner of heuristics.

Nevertheless, there are so-called conflict problems where heuristic and normative responses are mutually exclusive (this is the case in the typical tasks used in the heuristics and biases literature). In these cases, the heuristic response is assumed to be the default because it is fast and automatic (Evans & Over, 1996). Moreover, heuristic responding remains predominant in all age groups (Klaczynski, 2001). According to default interventionist models (Evans, 2006; Kahneman & Frederick, 2002), the heuristic system

delivers contextualized representations of problems and cues intuitive responses to them. However, these quick heuristic responses can be overridden by analytic processing given sufficient cognitive capacity and certain mental dispositions (Stanovich & West, 2000). In another approach (e.g., Klaczynski, 2001, Sloman, 1996), the two systems function independently and deliver their own solutions to problems. Then metacognitive skills and inhibitory processes are needed to allow System 2 responses to shape behaviour instead of the quick and effortless processes of System 1. (See more on this in Section 1.2.1.).

The focus of this experiment is on the development of heuristic and analytic responding and specifically on the relationship between cognitive capacity and these types of responses among children. Although research in this area is limited, dual-process theories have nevertheless been applied to the development of children's reasoning (see also Sections 2.1.2 and 2.1.3.). In general, these accounts assume that (in some form) both processing modes are available at all points of development—at least after an early age (e.g., Jacobs & Klaczynski, 2002; Kokis et al., 2002). They also agree that the efficiency and prevalence of analytic processing increases with development and it is associated with increases in cognitive capacity within a given age group (e.g., Handley et al., 2004; Stanovich & West, 2000).

The evidence concerning age-related changes in heuristic and analytic responding is mixed. Some studies have found that analytic responding increases with age on certain tasks, but this increase is not apparent on other tasks (e.g., Klaczynski, 2001; Kokis et al., 2002). Although heuristic responding is expected to decrease with age, some studies have reported the opposite pattern. There are three mainstream developmental dual-process accounts that have attempted to explain these findings. One of these accounts (e.g., Kokis et al., 2002) assumes a reduction in heuristic responding with age driven by increases in cognitive ability. The other accounts (e.g., Klaczynski & Cottrell, 2004; Reyna & Ellis, 1994) propose that some heuristic responses will become more prevalent with age regardless of cognitive ability. However, all dual-process theorists agree that heuristics are

independent of cognitive capacity. In addition, they all propose that analytic reasoning requires the decontextualization of problems (by means of decoupling, metacognitive operations, or the extraction of gist representations).

Nevertheless, there is some reason to suppose that the contextualization of problems (the retrieval and application of relevant knowledge) might not be automatic and effortless (see Section 2.2. for a review and discussion of the evidence from the research on conditional reasoning). Working memory has been found to be linked to effective retrieval of information from long-term memory in the case of children (e.g., Janveau-Brennan & Markovits, 1999). According to Barrouillet et al. (2002), during reasoning the memory structures that are associated with the content of the problem are automatically accessed and specific elements within this structure are activated. The less strongly a given element is associated with the context of the task, the more effort (i.e., cognitive resources) and time are necessary for it to be retrieved. In addition, there is evidence that abstract reasoning is facilitated by realistic content (see Section 2.2.2.). Despite the abstract nature of the premises, the development of abstract reasoning has been found to be reliant on appropriate access to empirical knowledge (Venet & Markovits, 2001). Similarly, in adults working memory capacity is related to the successful activation of relevant background knowledge when people reason with familiar causal conditionals (De Neys, et al., 2005), and higher ability people are also more inclined to retrieve elements from their long-term memory that are more remotely associated with the given context (Verschueren, et al., 2002). Thus, there is evidence that cognitive abilities (i.e., System 2 processes) not only play a role in abstract, hypothetical, logical thinking but they are also needed for reasoning about contextualized problems.

Taking into consideration all of the above, there are three contrasting predictions that can be made about the relationship between cognitive capacity and children's reasoning. Dual-process theories claim that heuristic and analytic responses are the result of independent processing modes. One dual-process account (e.g., Kokis et al., 2002)

presupposes a negative relationship between cognitive capacity and heuristic responding, whereas other accounts (e.g., Klaczynski & Cottrell, 2004; Reyna & Ellis, 1994) claim that there will be no relationship between the two at all. As a third possibility, because heuristic responding is based on the contextualization of problems, we can predict that it will be dependent on general cognitive resources. Although the relationship between the different processing modes and cognitive capacity has been examined widely in adult populations, and to some extent in adolescent samples as well, so far it has not been investigated in the case of children. This is possibly because most theorists consider heuristic processing to be basic and hardwired, and the apparent emergence of new heuristics with development is attributed to knowledge acquisition, rather than to the development of new ways of combining information (see e.g., Stanovich et al., 2008; an exception from this trend is the fuzzy trace theory where the emergence of heuristics is attributed to the development of gist memory systems).

In this study, we examined primary school children's performance on reasoning and decision-making problems on which there are clearly classifiable heuristic and analytic responses. In addition, a series of measures of cognitive capacity were included to determine which of the approaches described above accounts for the development of children's thinking the most adequately.

## Method

### *Participants*

Participants were 84 school children recruited from two local primary schools. Our sample consisted of 42 girls and 42 boys, between 5 years 2 months and 11 years 7 months of age, with a mean age of 8 years 6 months. We had roughly equal numbers of children from all year groups of primary school (i.e., 10–15 children from Year 1 to Year 6). The schools received goodwill payment for participating in the study.

## *Materials*

*Cognitive capacity measure:* Participants completed a short form of the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1991) consisting of the Vocabulary and Block Design subtests. This particular short form is reported to have the highest validity and reliability compared with any other two subtest short forms of the WISC (Sattler, 1992). Because we were interested in the children's absolute computational capacity rather than their IQs, we used the raw scores on the tasks for our analyses. We also used the counting span task, a working memory measure with a processing and storage component (see also Handley et al., 2004). The task consists of a processing component that requires counting the number of coloured dots presented on a computer screen and a storage component that involves the storage of the products of a series of these counting operations. Children were presented with a series of cards on a computer screen, each card showing a number of blue and red coloured dots in a randomly determined irregular pattern. The children's task was to count out loud the number of red dots from each card. Participants received a practice set of two cards followed by three trials consisting of two cards, three trials of three cards, three trials of four cards, and three trials of five cards. That is, there were four levels of difficulty, and three series of trials at each level. After each series of trials the children had to recall the number of dots on each card in the order they were presented. All children worked through all of these trials, regardless of performance. A working memory global score was calculated by adding up the number of instances when the child recalled the number of dots on the preceding cards in the correct order. In the Handley et al. (2004) study the span score was used as a measure of performance (where children have to complete at least two series of trials correctly at each level, so that they get a span score for that level, and to move on to the next level). The disadvantage of this method is that it is conservative and it can possibly underestimate children's actual working memory capacity. Thus, we decided to use the global score in the present study. That is, children were given one point for each series of

trial where they were able to recall the results of the preceding operations in the correct order, regardless of their performance on previous series of trials. We used the combined z scores of the WISC and the counting span task as an indicator of cognitive capacity.

*Reasoning and decision-making tasks:* The four types of reasoning and decision-making tasks we used are known to elicit a high proportion of heuristic responses in the case of adults. We used two conjunction fallacy tasks, two if-only fallacy tasks, two sunk cost fallacy tasks, and four syllogistic reasoning tasks. The tasks were adapted from Epstein, Lipson, Holstein, and Huh (1992), Klaczynski and Cottrell (2004), Kokis et al. (2002), and Tversky and Kahneman (1983), and they were developed to ensure that they were appropriate for a developmental sample. We presented the stories as a PowerPoint slide show that included photographic illustrations, and the text was read by the experimenter. Participants were given a booklet with the response options for each task, and they responded by circling what they believed to be the most appropriate response. Before asking the children to circle the most appropriate response, we gave them a brief summary of the story to lessen the memory requirements of the task. In the case of three problems (the if-only fallacy, sunk cost fallacy, and syllogistic reasoning tasks), there were three response options: normative, heuristic, and “other.” On the conjunction fallacy task, there were only two response options: heuristic and normatively correct. All of the tasks we used were so-called “conflict” problems; that is, heuristic and normative responses were mutually exclusive. We formed a heuristic, normative, and “other” composite score by converting the raw scores on the reasoning and decision-making tasks to z scores and adding these up. In what follows, we give a short description of each type of task (the rest of the tasks can be found in Appendix A).

The *conjunction fallacy* (adapted from Tversky & Kahneman, 1983) violates the simplest and most fundamental rule of probability, that the likelihood of two independent

events occurring the same time (in “conjunction”) should always be less than or equal to the probability of either one occurring alone. However, reasoners generally fail to take this into consideration when the conjointly presented information meets their expectations more than just one piece of information alone. Here is an example of one of the two tasks that we used to measure the conjunction fallacy:

*Sarah is 12. She is very talkative and sociable. She goes to drama classes and she learns to play the guitar. She wants to be a pop singer or an actress.*

(The following part was included in order to give children a chance to practice allocating liking ratings to the statements. As they had to circle a single answer in the case of the other tasks, it seemed necessary to highlight that they had to respond to this task in a different way.)

Now, let’s see an example first. Mark the sentence which you think is more likely to be true with number 1 and the one which is less likely to be true with number 2.

*(a) Sarah doesn’t have many friends.*

*(b) Sarah likes music.*

Now read the following statements. Your task is to mark the statement which is the most likely to be true with number 1, the next one with number 2, and so on. Mark the statement which is the least likely to be true with number 4.

*(a) Sarah has lots of CDs and DVDs.*

*(b) Sarah likes to cook.*

*(c) Sarah has many friends at school.*



*(d) Sarah likes to cook and she collects pop magazines.*

Children who considered sentence *(d)* more likely to be true than sentence *(b)* committed the conjunction fallacy and were given a heuristic point. Those who considered the second sentence more likely to be true than the last one were given a normative point. In the case of this task, there was no third response option.

The *if-only fallacy* (Epstein, Lipson, Holstein, & Huh, 1992) is based on counterfactual thinking. Counterfactual thoughts are thoughts about “what might have been,” and they usually emerge when something bad happens (Byrne, 2002). They help people to learn from their experiences, but they also lead to cognitive biases such as the *if-only fallacy*, which gives the person the illusion that a certain negative outcome was inevitable under given circumstances and so could have been foreseen and prevented from happening. A task that cues this kind of reasoning that we used in the current study is the following:

*Tom went camping with his family and he put his bike inside the caravan. His mother told him to put his bike on the roof rack like his sister did, but he didn't listen to her. As luck would have it, Tom's bike got broken. Now listen to the next story. Robert went camping with his family and he put his bike inside the caravan. He had to put it in there because there was a kayak on the roof rack. As luck would have it, Robert's bike got broken. What do you think about the two stories?*

*(a) Tom made a worse decision than Robert.*

*(b) Robert made a worse decision than Tom.*

*(c) It wasn't their fault, it was just bad luck.*

The third response is considered to be the normatively correct one (Epstein et al.,

1992). A heuristic response here is if one thinks that the first boy made a worse decision than the second one. It is easier to understand why this response is not normatively correct if we put the question this way: “Who made a better decision?” In this case, the majority of participants would agree that neither of the boys made a better decision, and they would be inclined to choose the normative response (Klaczynski & Cottrell, 2004). An “other” response is to say that the second boy made a worse decision than the first one.

The *sunk cost fallacy* (adapted from Klaczynski & Cottrell, 2004) occurs when people base their current decisions on inconsequential past decisions. A goal in which people invested much time, money, and/or effort turns out to be worthless and not desirable anymore. In this situation, they can choose between investing even more personal resources in the goal to reach it anyway (“throwing good money after bad”) or they can abandon it, letting the previously invested resources go to waste (but at the same time saving further investments in a worthless goal). A task that cues the sunk cost fallacy is the following:

*You bought a cinema ticket from your pocket money. You paid £5. You start to watch the movie, but after 5 minutes you are bored and the film seems pretty bad. How much longer would you continue to watch it?*

*(a) 10 more minutes*

*(b) 30 more minutes*

*(c) watch until the end*

*Now listen to the next story. You've got a free cinema ticket voucher and you go to the cinema and get a ticket for it. You start to watch the movie, but after 5 minutes you are bored and the film seems pretty bad. How much longer would you continue to watch it?*

(a) 10 more minutes

(b) 30 more minutes

(c) watch until the end

As sunk costs are not retrievable anymore, they should be ignored. Therefore, decisions in the two situations should be the same whether one paid for the ticket or not. This is considered to be the normative answer. The “waste not” heuristic (Arkes & Ayton, 1999) dictates that people continue to watch the movie longer if they paid for the ticket. An “other” response is if one decides to watch it longer when the ticket was for free.

The *sylogistic reasoning tasks* (adapted from Kokis et al., 2002) required children to judge whether different types of inferences—modus ponens (MP), modus tollens (MT), denial of the antecedent (DA), and affirmation of the consequent (AC)—led to valid or invalid conclusions (which were either believable or unbelievable). The tasks measure the belief bias (Evans, Barston, & Pollard, 1983), which is the tendency to accept believable conclusions and reject unbelievable ones irrespective of logical validity. Participants solved an MP (if p, then q; p, therefore q), an MT (if p, then q; not q, therefore not p), an AC (if p, then q; q, therefore p), and a DA (if p, then q; not p, therefore not q) syllogism. MP and MT are valid syllogisms, so the normative answer to them is to accept the conclusion. AC and DA, however, are invalid syllogisms, and the normatively correct answer to them is “not certain.” Apart from the normative point, children were also given a belief point whenever they accepted a believable conclusion or rejected an unbelievable one. All of these syllogisms were conflict problems; that is, there was a conflict between the believability and the validity of the conclusion. Thus, higher amount of belief-based responses on these tasks indicated lower adherence to logical necessity.

Before starting the task, children were given instructions emphasizing that they needed to accept the premises as true even if they sounded strange or funny, and then they needed to think about whether the conclusion followed from the premises or not. They had three response options from which to choose: “yes,” “no,” and “not certain.” The instructions stated that children needed to circle “yes” if they thought that the conclusion definitely followed from the premises, they needed to circle “no” if they thought that the conclusion did not follow from the premises, and they needed to circle “not certain” if they thought that it was possible that the conclusion followed from the premises but there was still a possibility that it did not. These options were illustrated on an example problem. Then the following script was read out for the children:

*Scientists have discovered a planet in our galaxy which has animals and plants living on it. This planet is similar to Earth in many respects, but there are some differences as well. I give you some questions. Please choose the answer which describes the animals and plants living on this planet the best.*

Children were given two practice problems first, after which they solved the actual tasks. Here is an example for an unbelievable valid syllogism:

*On this planet flowers have thorns on them.*

*Daffodils are flowers.*

*Does it follow that on this planet daffodils have thorns on them?*

yes                      no                      not certain

Children were given a normative point if they circled “yes,” they were given a belief point if they circled “no,” and they were given an “other” point if they circled “not certain.”

### *Procedure*

Participants took part in two testing sessions, each lasting approximately 25 minutes. Session 1 consisted of two conjunction fallacy tasks, two counterfactual reasoning tasks, two sunk cost fallacy tasks, and four syllogistic reasoning tasks. Participants were tested in groups of five to eight. Session 2 consisted of the counting span task and the Vocabulary and Block Design subtests of the WISC-III. On the second occasion, participants were tested individually.

### **Results**

In order to analyze age trends in giving heuristic, normative, and “other” responses to the reasoning and decision-making tasks, we divided our sample into three age groups. The age groups correspond to years 1-2, 3-4, and 5-6 in the British primary school system. The mean age of the children was 6 years 7 months in age group 1, 8 years 6 months in age group 2, and 10 years 5 months in age group 3. Table 3.1. presents the mean number of heuristic, normative and “other” responses on the tasks across the three age groups. There appears to be a tendency for heuristic responding to increase, and normative and “other” responding to decrease with age across the tasks.

ANOVAs conducted on the number of heuristic, normative and “other” responses on the syllogistic reasoning task across age groups indicated a main effect of age in the case of the heuristic responses  $F(2,83)=5.24$ ,  $p<.01$ ,  $\eta_p^2=.12$ , and the “other” responses  $F(2,83)=4.61$ ,  $p<.05$ ,  $\eta_p^2=.10$ , but the main effect of age for the normative responses was not significant,  $F(2,83)=3.07$ ,  $p=.052$ . In the case of the sunk cost fallacy, the main effect of age on “other” responses was significant  $F(2,83)=3.68$ ,  $p<.05$ ,  $\eta_p^2=.08$ , but it was not significant for the heuristic,  $F(2,83)=1.7$ ,  $p=.19$ , or the normative responses,  $F(2,83)=.25$ ,  $p=.78$ . On the if-only fallacy task there was a significant main effect of age on the heuristic  $F(2,83)=6.72$ ,  $p<.01$ ,  $\eta_p^2=.14$ , and the normative  $F(2,83)=4.8$ ,  $p<.05$ ,  $\eta_p^2=.11$ , but not on the “other” responses  $F(2,83)=.39$ ,  $p=.68$ . The main effect of age on the heuristic and

normative responses in the case of the conjunction fallacy task was significant  $F(2,83)=8.73, p<.01, \eta_p^2=.18$ , for both response types.

Table 3.1. *The mean proportion of heuristic, normative and “other” responses on the different tasks across the three age groups.*

		Age group 1 (n=30)	Age group 2 (n=26)	Age group 3 (n=28)
Syllogistic reasoning	Heuristic	.29 (.26)	.48 (.25)	.48 (.25)
	Normative	.36 (.23)	.22 (.20)	.35 (.26)
	“Other”	.35 (.21)	.30 (.26)	.17 (.19)
Sunk cost fallacy	Heuristic	.42 (.42)	.45 (.38)	.59 (.33)
	Normative	.27 (.36)	.32 (.34)	.30 (.32)
	“Other”	.31 (.38)	.23 (.26)	.11 (.21)
If-only fallacy	Heuristic	.18 (.25)	.40 (.38)	.50 (.38)
	Normative	.65 (.27)	.50 (.40)	.40 (.35)
	“Other”	.17 (.23)	.10 (.20)	.10 (.28)
Conjunction fallacy	Heuristic	.68 (.31)	.86 (.30)	.96 (.13)
	Normative	.32 (.31)	.14 (.30)	.04 (.13)
	“Other”	.28 (.27)	.21 (.24)	.13 (.23)
Overall	Heuristic	.39 (.31)	.55 (.33)	.63 (.27)
	Normative	.40 (.29)	.30 (.31)	.27 (.27)
	“Other”	.28 (.27)	.21 (.24)	.13 (.23)

We also conducted an ANOVA on the heuristic, normative and “other” composite scores (i.e., the sum of the z scores of heuristic, normative and “other” responses on the different tasks) across age groups, and this indicated a main effect of age on each type of response  $F(2,83) = 13.82, p<.01, \eta_p^2=.25$ ;  $F(2,83)=5.29, p<.01, \eta_p^2=.12$ , and  $F(2,83) = 5.24, p <.01, \eta_p^2=.12$ , respectively. Tukey post-hoc tests revealed that the children in age group 1

gave significantly fewer heuristic responses and significantly more normative responses than the children in age group 2. In addition, children in age group 1 gave significantly more “other” responses than children in age group 3. No other specific post-hoc contrasts were significant.

Thus, the discrete ANOVA analyses indicated significant age-related changes on the different types of responses for most of the reasoning and decision-making tasks, and the non-significant trends were also in the same direction. The analyses on the composite scores showed a tendency for heuristic responding to increase, and normative and “other” responding to decrease with age. The post-hoc comparisons indicated that the changes in heuristic and normative responding occurred mostly between the ages of 6 years 7 months and 8 years 6 months, whereas the changes in “other” responding took place gradually between the ages of 6 years 7 months and 10 years 5 months.

*Relationships between the individual differences measures*

The next set of analyses examined the relationship between age (in months) and cognitive ability (as indicated by the counting-span task, and the short form of the WISC-III). First we calculated the reliability of the counting-span tasks by performing a split-half correlation between odd and even trials between each set level, and then we used the Spearman-Brown formula to give an estimate of the reliability of the task. The reliability of this measure was 0.77, which is at an acceptable level.

Table 3.2. *Correlations between the individual differences measures.*

	1.	2.	3.
1. age in months	--		
2. working memory	.66**	--	
3. WISC score	.78**	.78**	--

Next we combined the WISC score and the working memory score into a single variable, by computing the combined z score for these measures. The vocabulary subtest of the WISC is a good indicator of crystallized intelligence, whereas the block design and the working-memory tasks are measures of fluid intelligence (see e.g., Blair, 2006). The shared variance between the two scores (i.e., the WISC score, and the counting span score) was also indicated by a considerable correlation between the two measures (see Table 3.2.). We also found a significant correlation between age and the combined cognitive ability measure ( $r=.78, p<.001$ ).

#### *Relationships between age, cognitive ability and the type of responses given*

As a next step we examined the correlations between the individual differences measures and the three types of responses (heuristic, normative, and “other”) on the reasoning and decision-making tasks (the results are displayed in Table 3.3.). Heuristic responding on three tasks (the exception being the sunk cost fallacy) increased significantly with age. Moreover, heuristic responding on all tasks was positively correlated with cognitive abilities. Normative responding was either unrelated to age and cognitive abilities (as in the case of the syllogistic reasoning and the sunk cost fallacy task), or it was negatively correlated with both age and cognitive abilities. In the case of the conjunction fallacy, there were only two response options, thus, the decrease in normative responses was probably a by-product of children’s increasing preference for the heuristic response. However, in the case of the if-only fallacy the decrease in normative responses had to be at least partly independent of the increase in heuristic responses, as there were three response options. The number of “other” (i.e., atypical) responses decreased with development on both the syllogistic reasoning and the sunk cost fallacy task, but it was unrelated to age and cognitive abilities in the case of the if-only fallacy task.



In sum, we found evidence for a general increase in heuristic responding with age and cognitive abilities, accompanied in some cases by a reduction in normative responding (if-only fallacy and conjunction fallacy) or the absence of any increase (sunk cost and syllogistic reasoning tasks).

Table 3.3. *Correlations between age, cognitive ability and response types across tasks.*

		Age in months	Cognitive ability
Syllogistic reasoning	heuristic responses	.27*	.31**
	normative responses	-.06	-.03
	“other” responses	-.24*	-.32**
Sunk cost fallacy	heuristic responses	.18	.32**
	normative responses	.03	-.04
	“other” responses	-.27*	-.36**
If-only fallacy	heuristic responses	.34**	.38**
	normative responses	-.30**	-.32**
	“other” responses	-.06	-.10
Conjunction fallacy	heuristic responses	.37**	.42**
	normative responses	-.37**	-.42**

The positive correlation between heuristic responding and cognitive ability is inconsistent with the predictions of dual-process theories which either assume a reduction in heuristic responses with increasing cognitive capacity (e.g., Kokis et al., 2002), or no relationship (Klaczynski & Cottrell, 2004; Reyna & Ellis, 1994). However, because cognitive capacity is strongly correlated with age, at this point it is not clear whether cognitive capacity and heuristic responding are genuinely correlated, or whether age is the mediating variable.

### *Correlations between heuristic responding across the reasoning and decision-making tasks*

As a next step we focused on the relationship between heuristic responding across the different tasks. We found a significant correlation between belief bias and committing the conjunction fallacy. In addition, heuristic responding on the if-only task was significantly correlated with heuristic responding on the sunk cost, and the conjunction fallacy tasks (see Table 3.4). Overall we found some evidence that giving heuristic responses on one task is related to giving heuristic responses on other tasks. This was also confirmed by the fact that the Cronbach alpha computed on an overall index of heuristic responding across all tasks was .52, which can be considered a satisfactory level of reliability (Rust & Golombok, 1999).

Table 3.4. *Intercorrelations between heuristic responses on the reasoning and decision-making tasks.*

	1.	2.	3.	4.
1. Belief bias	--			
2.Sunk cost fallacy	.17	--		
3. If-only fallacy	.06	.29**	--	
4.Conjunction fallacy	.32**	.17	.27*	--

### *Predictors of heuristic responding*

Finally, in order to decide whether the positive correlation between heuristic responding and the cognitive ability measures we observed was genuine or was mediated by age we ran a regression analysis with the heuristic response composite score as the dependent variable and the individual differences measures (age and cognitive capacity ) as predictors. Before doing this we combined the WISC score and the working memory score into a single variable, by computing the combined z score for these measures.

A linear regression analysis revealed that when age and cognitive ability were both entered simultaneously, cognitive capacity ( $\beta=.52$   $p<.001$ ) but not age ( $\beta=.05$   $p=.32$ ), was a significant predictor of heuristic responding. These two variables accounted for one third of the variance in heuristic responding ( $R^2=.29$ ), which was significant ( $F(2,81) = 18.29$ ,  $p<.001$ ). When age was entered first, it explained 19% of the variance, whereas cognitive ability predicted a significant additional 10%. In contrast, after cognitive ability was entered first ( $R^2=.30$ ), entering age did not increase further the proportion of variance explained (see Appendix S1 for a supplementary table for these analyses). Thus, the regression analysis confirmed that the emergence of these heuristics with age was genuinely related to increases in cognitive capacity.

## Discussion

In this study we looked at developmental changes in giving normative, heuristic, and “other” responses to four different reasoning and decision-making problems. We found that in general heuristic responding increased with age, and the number of normative and “other” responses decreased between the ages of 5 and 11.

It is not surprising that the number of atypical (“other”) responses decreased with development. However, the fact that normative responding also decreased, and that the number of heuristic responses increased with age is more striking. Although certain dual-process theories (e.g., Klaczynski & Cottrell, 2004; Reyna & Ellis, 1994) are able to account for these findings, the evidence that the increase in heuristic responding was positively associated with increases in cognitive capacity goes against the predictions of each version of dual-process theories presented in the introduction. These approaches propose that heuristic responses are the products of automatic processing, therefore they are activated and applied independently of a general processing mechanism.

However, our results are in line with some findings from the area of conditional reasoning that indicate a relationship between reasoning about contextualized problems

(specifically, the retrieval and integration of relevant knowledge) and working memory capacity (e.g., De Neys et al., 2005; Janveau-Brennan & Markovits, 1999). The effortful nature of the retrieval of relevant knowledge from long-term memory might explain why the belief bias and heuristic responding on the conjunction fallacy problems were related to cognitive ability.

To sum up our results, there is evidence that cognitive capacity (and System 2 processes) not only play a role in abstract, hypothetical, logical thinking, but are also needed for the contextualization of problems. The reason that this has been overlooked for so long is probably that this relationship between reasoning about familiar topics and cognitive capacity may be demonstrated only under special conditions (i.e., dual-tasking; Cho, Holyoak & Cannon, 2007) or in the case of children (Janveau-Brennan & Markovits, 1999), or elderly people (McKinnon & Moscovitch, 2007); that is, in cases where participants' cognitive capacity is relatively limited. Since normal adults generally find these problems easy, differences in their cognitive capacity will be unrelated to making these "heuristic" inferences (although it does not necessarily imply that these processes are completely undemanding of cognitive capacity in the case of adults). On the other hand, adults giving normative responses have to consciously inhibit the easily available heuristic responses, and engage in further processing in order to produce a normative response. This is supported by the fact that normative responding in the case of adults and adolescents has been found to be related to cognitive capacity and thinking styles (Stanovich, 1999).

If we accept the above argument, a novel and interesting question emerges. Namely, at some point in development the contextualisation processes involved in heuristic reasoning should become virtually effortless. Given the relationship between contextualisation and general cognitive resources, we can expect that the automation of contextual processing should be driven by the maturation of cognitive and executive systems. As we described in Chapter 2, some executive components continue to develop into adulthood whereas others reach a mature state during adolescence (see e.g., Huizinga

et al., 2006). Building on these assumptions, in Experiment 2 we administered the same tasks as in Experiment 1 together with measures of executive functioning to a sample of adolescents between the age of 12 and 16.

### *3.2. Experiment 2: The development of heuristic reasoning in adolescence and its relationship to executive functioning*

The main aim of Experiment 2 was to extend the investigation of Experiment 1 to an older age group, namely a group of early and mid-adolescent children between the ages of 12 and 16. A number of reasoning theorists predict that reasoning performance should change during adolescence (see also Section 2.1.2.). According to Klaczynski (e.g., Klaczynski & Cottrell, 2004) metacognitive abilities develop by the age of 14. As a result, adolescents' thinking is more flexible than children's thinking, and they are able to compare and integrate multiple options, instead of just responding on the basis of their intuitions and feelings, or following rules rigidly (e.g., Klaczynski, 2007). On the other hand, Klaczynski and colleagues also note that heuristic responding remains predominant even in the case of adults, and normative responding does not necessarily increase with development (e.g., Klaczynski, 2007).

Fuzzy-trace theorists (e.g., Brainerd et al., 2004) claim that gist (i.e., contextualised) processing increases most profoundly between the ages of 7 and 11 (in line with the results of Experiment 1), whereas verbatim processing (i.e., the decontextualised, fact-based representation of problems) shows a marked increase between the ages of 11 and 14. Verbatim processing can suppress gist processing, leading to a reliance on facts and precise details instead of fuzzy gists. This results in inverted U-shaped patterns of performance on tasks where contextualisation leads to inappropriate responses. Similarly, studies that have investigated age-related changes in everyday conditional reasoning showed that in the case of inferences, where performance was based on an interplay between the retrieval and inhibition of contextual information (e.g., Markovits & Barrouillet, 2002) performance

initially increased then decreased. More specifically, the ability to retrieve counterexamples increases during childhood, but after the onset of adolescence (i.e., 12–13 years of age) children start to inhibit logically inappropriate counterexample activation in the case of the MP inference (DeNeys & Everaerts, 2008). Even in the case of late adolescents, MP acceptance depends on the efficiency of the inhibition process (see Section 2.2.2.).

Other theorists, who expect reasoning abilities to steadily improve with development also propose that reasoning performance should change considerably during adolescence. Piaget (e.g., Piaget, 1976) claimed that children reach the formal operational stage around the age of 12, which is the time when they start to understand and apply the rules of formal logic. Stanovich and colleagues (e.g., Stanovich et al., 2008) also predict that children's performance should get closer to the normative standards, but in this view the basis of these changes is children's increasing cognitive capacity. Although people with higher cognitive capacity do not always perform better on reasoning tasks, this usually happens when there is no detectable conflict within tasks, for example when participants are unaware of the relevant normative rule (Stanovich & West, 2008). According to Stanovich et al. (2008) this often happens in the case of tasks that require probability knowledge. However, children's probability knowledge should considerably increase during adolescence. Moreover, the tasks used in the present study employ a within-subjects design which makes the conflict between intuitions and logic more transparent.

Apart from examining age trends in reasoning performance, another aim of this study was to explore the links between reasoning performance and the development of cognitive abilities during adolescence. Stanovich et al. (2008) propose that children's increasing working memory capacity is the most important driving force behind the increases in logical competence. In fact, many studies have found a relationship between working memory capacity and logical reasoning. However, it is less clear whether working memory alone can explain logical performance on tasks where there is a conflict between logic and

intuitions. For example, De Neys (2006), using a dual-task manipulation, demonstrated that cognitive load selectively impairs performance on conflict problems, whereas performance on non-conflict problems was unaffected by cognitive load. By contrast, developmental studies found that performance on both conflict and non-conflict tasks was related to working memory capacity. However, performance on conflict tasks was additionally predicted by measures of inhibition (see e.g., Handley et al., 2004; Simoneau & Markovits, 2003).

Working memory and inhibition are two out of the three components of executive functions (EFs) which Miyake et al. (2000) identified. The third component is shifting (or set-shifting) which is the ability to flexibly switch between mental sets, representations or rules. Although moderately correlated, working memory, set-shifting, and inhibition are separable constructs, and they predict performance differentially on different tasks that are designed to measure EFs (Miyake et al., 2000). Developmental research revealed that EFs have a protracted course of development, beginning in early childhood and continuing into adolescence. Moreover, different EF-components show distinct developmental trajectories. A recent study (Huizinga, Dolan, & van der Molen, 2006) found that set-shifting and performance on the stop-signal task have reached mature levels by adolescence, while working memory and basic processing speed followed a more protracted course of development into young adulthood.

Based on the predictions of reasoning theorists, children's reasoning abilities can be expected to change when they reach adolescence. More specifically, some theorists predict that children become more logical (e.g., Stanovich et al., 2008), or that their thinking becomes more flexible and more subject to metacognitive control (e.g., Klaczynski, 2007), or that they become more able to inhibit irrelevant associations (e.g., Brainerd et al., 2004; De Neys & Everaets, 2008) around the age of 12 or 14. These changes have been variously attributed to the development of working memory capacity, inhibition, or metacognitive skills. To test these predictions in Experiment 2 we administered the same heuristics and

biases tasks as in Experiment 1 together with measures of working memory, general intelligence, inhibition and set-shifting ability to a group of adolescents between the ages of 12 and 16. Experiment 1 indicated that giving “heuristic” responses (i.e., the responses that are considered to be heuristic in the case of adults) increased with age between the ages of 5 and 11. Heuristic responding in this age group was positively correlated with cognitive ability, and heuristic responses became dominant around the age of 9. Based on the predictions described above, it can be expected that in the adolescent group heuristic responding will gradually become effortless (independent of cognitive ability), whereas normative responding will become positively correlated with cognitive capacity (especially working memory) or measures of executive functioning (especially inhibition), as children will start to inhibit inappropriate responses at this age. Consequently, the number of normative responses should increase. Finally, it was expected that heuristic responding will either decrease or remain stable during this period, and the number of atypical (i.e., “other”) responses will also decrease (see e.g., Klaczynski, 2001).

## Method

### *Participants*

Participants were 49 children (31 boys) from a Plymouth secondary school. The mean age of the participants was 13 years 9 months (age range: 11 years 11 months to 15 years 8 months).

### *Materials and procedure*

The same reasoning tasks were used as in Experiment 1, together with the same short form of the WISC-III (Wechsler, 1991) consisting of the Vocabulary and Block Design subtests. As before, the raw scores on the tasks (indicating raw computational capacity, rather than IQs) were used for the analyses. In addition, as in Experiment 1 the counting span task (a working memory measure with a processing and storage component) was also administered. In this study the stop-signal task, a measure of inhibition (based on Handley



et al., 2004), and a set-shifting task (based on Miyake, Emerson, Padilla & Ahn, 2004) were also administered.

The stop signal task is a computerised measure of inhibitory control. The task was selected because Handley et al. (2004) found that it was a good predictor of reasoning performance in the case of children. The task consists of two types of trial; primary task trials and stop signal trials. On primary task trials a fixation point was presented on the screen for 500 ms followed by the presentation of an X or an O. The child responded by pressing an X button or an O button on one of two button boxes. On stop signal trials the X or the O is presented along with a tone (the stop signal). On these trials the child is instructed to withhold their response. The tones were presented randomly at 200, 300, 400 or 500 ms before the child's mean reaction time (MRT) to primary trials. Because the tone delay is set relative to the MRT for each child, the inhibitory demands are independent of the response time to primary trials and approximately equivalent for each participant. Children completed a practice block of trials first. These trials familiarised the child with the task while establishing the MRT that was used for setting the tone delays for subsequent blocks of trials. The main measure of performance on the task was accuracy on the stop signal trials.

Set-shifting requires shifting back and forth between multiple tasks, operations, or mental sets (Monsell, 1996). Recent work suggests that when switching to a new operation, it may be necessary to overcome proactive interference due to having previously performed a different operation on the same type of stimuli (see Miyake et al., 2000). Thus, individual differences in set-shifting ability may not be a simple reflection of the ability to engage and disengage appropriate task sets *per se*, but may instead involve the ability to perform a new operation in the face of proactive interference. The *colour-shape task* (based on Miyake et al., 2004) used in the present study required participants to switch between classifying shapes and classifying colours, and the dependent variable was the MRT on switch trials divided by the MRT on non-switch trials. On each trial, a random

cue indicating which subtask to perform was presented just before the stimulus and remained on the screen until the participant responded. Children responded by pressing one of two button boxes. Each button box corresponded to one of the colours, as well as one of the shape categories.

Participants took part in two testing sessions. Session 1 took about 15 minutes, and participants solved the reasoning tasks. Participants were tested in small groups. After listening to the instructions given by the experimenter they worked through the tasks individually. Session 2 consisted of the counting span task, the Vocabulary and Block Design subtests of the WISC-III, the stop-signal task, and the set-shifting task. This session took about 40 minutes, and children did the tasks on a computer (apart from the WISC subtests), supported by the experimenter.

## Results

First, in order to analyze age trends in giving heuristic, normative, and “other” responses to the reasoning and decision-making tasks, the sample was divided into two age groups. The age groups correspond to Years 7–8, and 9–10 in the British school system. The mean age of the children was 12 years 9 months in the younger age group, and 14 years 9 months in the older age group.

The number of heuristic, normative, and “other” responses given by the two age groups on the different tasks were compared using independent samples *t* tests (see Table 3.5.). In the case of the syllogistic reasoning tasks the *t* tests indicated no change in terms of heuristic and normative responding. However, there was a marginal increase in “other” responses across age groups. In the case of the sunk cost tasks the number of heuristic and other responses remained stable with development, but there was a marginal increase in normative responding. On the if-only tasks there was no difference between age groups in any type of response. There was also no change in responding on the conjunction fallacy tasks. Finally, the overall proportion of heuristic, normative, and other responses given to

the four different types of task was analysed. Again, there was no evidence of age-related change.

Table 3.5. *Proportion of heuristic, normative and "other" responses on the different types of reasoning tasks across age groups.*

		Younger group	Older group		
		(n=23)	(n=26)	t (47)	p
Syllogistic reasoning	Heuristic responses	.46 (.23)	.40 (.27)	.73	.47
	Normative responses	.39 (.21)	.36 (.20)	.60	.55
	"Other" responses	.16 (.03)	.24 (.18)	1.78	.08
Sunk cost fallacy	Heuristic responses	.24 (.30)	.21 (.29)	.76	.74
	Normative responses	.54 (.30)	.71 (.35)	1.81	.08
	"Other" responses	.11 (.21)	.06 (.16)	.95	.35
If-only fallacy	Heuristic responses	.52 (.35)	.56 (.41)	.33	.75
	Normative responses	.43 (.31)	.42 (.39)	.11	.91
	"Other" responses	.04 (.14)	.02 (.10)	.70	.49
Conjunction fallacy	Heuristic responses	.93 (.17)	.98 (.10)	1.13	.27
	Normative responses	.07 (.17)	.02 (.10)	1.16	.27
Total	Heuristic responses	.54 (.14)	.54 (.14)	.21	.84
	Normative responses	.36 (.12)	.38 (.15)	.06	.95
	"Other" responses	.10 (.13)	.11 (.09)	.04	.97

#### *Relationships between the individual differences measures*

The next set of analyses examined the relationship between age (in months), cognitive ability (as indicated by the counting-span task, and the short form of the WISC-III), and executive functioning (as measured by the stop-signal, and the set-shifting tasks).

Before computing correlations between these measures, we calculated the reliability of the stop-signal task, and the set-shifting task. We computed Cronbach's alpha for both tasks (including the correct trials only). In the case of the stop-signal task performance was indexed by the mean number of correct responses to the stop signal trials collapsed across the different delays. Consequently, when computing Cronbach's alpha for this task, we included the stop-signal items only (regardless of signal delay) which resulted in  $r = .90$ . In the case of the set-shifting task the score was computed by dividing the mean RT on switch trials by the mean RT on non-switch trials. Thus, we computed Cronbach's alpha for both the switch ( $r = .97$ ) and non-switch ( $r = .96$ ) trials separately.

Table 3.6. displays the correlations between age and the cognitive ability and executive functioning measures (note that higher set-shifting cost indicates worse performance). Working memory significantly increased with age, but there was no age-related change on the other measures. Similarly to Experiment 1, WISC scores correlated significantly with working memory scores. There was some indication of a relationship between the WISC scores and the measures of executive functioning, but it did not reach significance (possibly because of the relatively small sample size). This was in line with the idea that the executive function measures tapped into different aspects of cognitive functioning than the measures of cognitive capacity. Similarly, the three measures of executive functioning were positively related, but the correlations did not reach significance. This corresponds to the idea of the relative independence of different aspects of executive functioning (see e.g., Huizinga et al., 2006; Miyake et al., 2000).

Table 3.6. *Correlations between age and the cognitive ability measures.*

	1.	2.	3.	4.	5.
1. age in months	--				
2. WISC	.23	--			
3. working memory	.37*	.49**	--		
4. inhibition	.24	.25	.15	--	
5. set-shifting cost	-.19	-.20	-.25	-.16	--

*Relationships between the individual differences measures and heuristic responding.*

As a next step the correlations between the individual differences measures and heuristic responding on the reasoning and decision-making tasks were analysed (the results are displayed in Table 3.7.; correlations between normative and “other” responses and the individual differences measures are displayed in Appendix S2 Table 1). As in Experiment 1 we used the combined z scores of the WISC subtests and the working memory task as an indicator of cognitive capacity. Due to the relatively small sample size we also discuss the possible implications of marginally significant correlations.

In the case of the syllogistic reasoning task neither heuristic nor normative responding changed with age. However, the number of “other” responses increased with age. Moreover, there was a tendency for children with higher inhibition scores to give less belief-based and more “other” responses on this task.

In the case of the sunk cost fallacy normative responding significantly increased with age, whereas the number of “other” responses (which is the opposite of the heuristic response) significantly decreased. In addition, giving heuristic responses was negatively correlated with cognitive capacity, whereas children displaying higher cognitive flexibility (as indicated by a lower cost of set-shifting) tended to give more normatively correct responses.

Response patterns on the if-only fallacy task did not change with age, but atypical “other” responses (which were the opposite of heuristic responses) were negatively correlated with cognitive ability, and positively correlated with set-shifting cost. Moreover, children with higher inhibition scores tended to give more heuristic and less normative responses on this tasks.

Table 3.7. *Correlations between the individual differences measures and heuristic responding on the different tasks.*

	Age	Cognitive ability	Inhibition	Set-shifting cost
Belief bias	-.19	-.18	-.25 <sup>†</sup>	.07
Sunk cost fallacy	-.06	-.14	-.32 <sup>*</sup>	.08
If-only fallacy	.10	.18	.25 <sup>†</sup>	.08
Conjunction fallacy	.29 <sup>*</sup>	.30 <sup>*</sup>	.15	.20
Total heuristic responses	.07	.09	-.08	.21

<sup>†</sup> correlation significant at the  $p < .10$  level (Note that higher set-shifting cost indicates worse performance.)

The tendency to commit the conjunction fallacy significantly increased with age, and it was also positively correlated with cognitive ability. To check if this correlation between cognitive ability and committing the fallacy still held after controlling for age, partial correlations were computed. After controlling for the effect of age the correlation between cognitive ability and conjunction fallacy responses dropped below significance ( $r(43) = .23$ ,  $p = .14$ ).

Finally, we also analysed the relationship between overall heuristic, normative and “other” responding and the measures of age, cognitive ability and executive functioning. Normative responding was negatively correlated with set-shifting cost. This relationship

was also confirmed by a regression analysis (see Appendix S2 Table 4). Set-shifting cost explained 11% of the variance in normative responding ( $F(1, 47) = 5.44, p < .05$ ). However, it is worth noting that there was no significant correlation between either heuristic or normative responding on the different tasks (see Appendix S2, Tables 2 and 3), so although in the dual-processes literature it is common practice to combine different measures of heuristic/normative responding based solely on theoretical considerations (see e.g., Kokis et al., 2002) in the case of the present study this does not seem to be justified, at least empirically.

## Discussion

Although comparisons between the young adolescent and mid-adolescent group using *t* tests indicated virtually no change in reasoning performance, the analysis of response patterns with a more sensitive measure (i.e., correlations) did show some evidence of age-related change. Using very similar tasks to the present study, and with the participation of children from a similar age group, Klaczynski (2001) found that normative responding on the sunk cost and if-only tasks increased with age, whereas heuristic responding remained stable, and the number of “other” responses decreased. In line with this, in our sample the number of heuristic responses on the sunk cost fallacy task remained stable, whereas normative responding increased and “other” responding decreased. Heuristic responding on this task was unrelated to cognitive ability, and negatively correlated with inhibition scores, which could be interpreted as a sign that the automatic and effortless tendency to give heuristic responses needs to be suppressed in order that children can give the normatively correct response on this task. However, it is worth noting that a lack of correlation between heuristic responding and cognitive ability does not necessarily imply that producing the response is completely effortless. Rather, it indicates that children, regardless of their cognitive ability, are able to produce the response (that is, they all have the necessary capacity to do this). Additionally, there was a marginally

significant correlation between cognitive flexibility (i.e., set-shifting ability) and normative responding which indicated that children who are able to flexibly switch between concepts are more likely to give the normatively correct response, possibly because they are less likely to rigidly overuse the “waste not” rule (see Arkes & Ayton, 1999).

Results on the if-only tasks also correspond to the findings of Klaczynski (2001) in some respects. Heuristic responding on this task remained stable with development, and although the number of “other” responses remained stable, giving these atypical responses correlated positively with the cost of set-shifting, and negatively with the ability to inhibit impulsive response tendencies. That is, children who show a reversed if-only effect performed lower on the measures of executive functioning. Although other correlations only approached significance in the case of this task, it is interesting to note that the correlations between heuristic and normative responding and inhibition seem to be reversed compared to what most theorists would predict (i.e., children higher in inhibition tended to give more heuristic responses, whereas children lower in inhibition tended to give more normative responses). A possible interpretation of these trends is that judging past actions on the basis of the outcome that they lead to may require some sort of cognitive effort even in adolescence. Although hindsight bias seems to develop early (e.g., Bernstein, Atance, Loftus & Meltzoff, 2004), counterfactual thinking develops relatively late in childhood (e.g., Beck, Robinson, Carroll & Apperly, 2006), and even in adulthood people find it hard to consider more than one possibility at a time (e.g., Evans, 2006; Evans et al., 2003).

Similarly to Experiment 1, in the early to mid-adolescent age group the tendency to commit the conjunction fallacy was positively correlated with both age and cognitive capacity (although the correlation with cognitive capacity fell below significance after controlling for age). Moreover, virtually every child in the older age group (98%) committed the conjunction fallacy, indicating that participants, regardless of cognitive ability, were susceptible to the fallacy. Stanovich and West (1998) found that people with



higher cognitive capacity were less likely to commit the conjunction fallacy than lower ability people. De Neys (2006) also demonstrated that giving normatively correct responses on the conjunction fallacy task required more time than giving heuristic responses (i.e., response times were significantly lower for incorrect responses). De Neys also used a dual-task manipulation to demonstrate that experimentally restricting participants' working memory capacity leads to higher levels of heuristic responding on the conjunction fallacy task. However, Stanovich and West (2008) using between-subjects versions of the task with university students found that the conjunction fallacy was *more* likely to occur in the case of high, as opposed to lower ability students. They argued that this was because in a between-subjects design the conflict between the normative rule and participants' intuitions is virtually impossible to detect (although this does not explain why the fallacy should be more common across high ability participants). Stanovich and West (2008) also cited Kahneman and Tversky (1982) who reported that statistically sophisticated psychology graduate students were aware of the conjunction rule they had violated (thus, in Stanovich & West's terms they possessed the relevant mindware but did not detect the necessity for override). However, statistically naïve undergraduate students failed to endorse the conjunction rule.

It is important to note that in the studies that found a relationship between cognitive ability and resisting the conjunction fallacy, participants were university students, and the proportion of correct responding was only around 10%. Apart from the possibility that even university students do not endorse the conjunction rule, based on the present developmental data (the results of this experiment together with Experiment 1) it can be proposed that the reason for the high failure rates might result from the fact that committing the fallacy is not fully automatic even in the case of adolescents. Given that producing the normatively *incorrect* response seems to require a certain amount of effortful thinking, it is no surprise that virtually none of the participants in the present study invested further effort into (or considered the possibility of) suppressing this response.

In the case of the syllogistic reasoning task there was no relationship between age or the measures of cognitive capacity and executive functioning, and giving heuristic or normative responses, apart from a marginally significant negative correlation between children's inhibition scores and their giving heuristic responses. However, the number of "other" (i.e., neither belief-based, nor logical) responses significantly increased with age. In this task children were instructed to reason on the basis of logical validity, ignoring the believability of the premises. It is possible that between the age of 12 and 16 adolescents become increasingly capable of inhibiting their beliefs, without necessarily being able to work out the logically valid solution. This is somewhat supported by the fact that "other" responses were given more often by children with higher inhibition scores.

Finally, although the overall number of heuristic, normative or "other" responses did not change with development (which is no surprise, looking at the mixed age trends on the different tasks), there was a significant correlation between overall normative scores and set-shifting ability. Due to the lack of significant relationships between normative responding on the different tasks this finding could be considered as an artefact. However, if this relationship is meaningful, this could indicate that giving normative/analytic responses on these tasks requires mental flexibility and the consideration of multiple possibilities. In fact, set-shifting ability explained 11% of the variance in normative responding. This finding seems somewhat related to the idea that adolescents' thinking becomes increasingly flexible around the age of 14 (e.g., Klaczynski, 2009), although neither normative responding, nor set-shifting ability increased with age in the present sample. Thus, instead of claiming that flexibility increases with age, it could be argued that once heuristic responding becomes automatic (which happens with an increasing range of tasks as children get older) the role of mental flexibility increases, as this makes children more able to consider alternative responses besides the easily available heuristic response. Mental flexibility might also provide a way of giving a non-heuristic response without necessarily having to fully inhibit the heuristic response. Interestingly, although most

reasoning theorists predict that cognitive ability and inhibition play a crucial role in normative responding, there was little evidence for this in the present investigation. It is possible that inhibition plays a more important role once heuristic responses become automatic (which might be related to the increasing ease of memory retrieval, which, in turn, could be related to the age-related increase in processing speed, or processing fluency), and well-practised, and come with a greater level of fluency and confidence.

Similarly to Huizinga et al. (2006) we found that children's performance on the working memory task significantly increased during adolescence, whereas their inhibition and set-shifting performance remained reasonably stable between the age of 12 and 16 (possibly because children already reached their final level of performance on these tasks). Moreover, like Huizinga et al. (2006) we also found that our measures of executive functioning were not very strongly correlated with each other, which lends some support to the notion of the relative independence of working memory, inhibition and set-shifting.

To sum up the most important conclusions of this study, instead of finding evidence that children over 12 reach a mature level of cognition (see e.g., Piaget, 1976), their performance on different tasks showed mixed patterns. This is also supported by the fact that, unlike in Experiment 1, where the same tasks were administered to a group of younger children, in the present study there was no relationship between giving heuristic responses on the different tasks (see Appendix S2 Table 2). Importantly, heuristic responding either remained stable or increased with age in the case of the set of tasks that were employed. Moreover, children with higher cognitive ability did not perform any better on these tasks than children with lower ability, although other aspects of executive functioning (i.e., inhibition and set-shifting) seemed to enhance normative responding. These findings are not very easy to reconcile with some of the claims that dual-process theorists make about the development of reasoning abilities. On the other hand, the present findings are in line with research on executive functioning that shows a protracted development of certain EF skills that continues into young adulthood.

*3. 3. Experiment 3. The development of heuristic reasoning in adolescence, and its relationship with analogical reasoning ability.*

Although some developmental reasoning theorists predict an increase in normative responding during adolescence, Experiment 2 showed mixed findings with little change in normative responding between the ages of 12 and 16, whereas heuristic responding increased with age in the case of the conjunction fallacy task. In Experiment 3 we followed up on some of the results of Experiment 2 by using syllogistic reasoning and conjunction fallacy problems.

The tasks used in this experiment were different from the tasks used previously, and we also have not included sunk cost and if-only fallacy problems (the reason for this is that these tasks are related to emotional processing which makes the results harder to interpret, and also makes the use of normative standards in judging the responses harder to justify).

In addition to the syllogisms and the conjunction fallacy problems, we used some tasks from the heuristics and biases literature which are related to probabilistic reasoning: the positive/negative recency effects, and the two versions of the engineers and lawyers problem (see Section 1.1.1.). As in the previous experiments, we used modified versions of the classic problems in order to make them more appropriate for children and adolescents. In this experiment we also introduced a methodological innovation, which consisted of manipulating the presence of conflict between intuitions and logical structure (see Figure 3.1.).

<b>Engineers and lawyers problem – representative description</b>	
<p><b>Conflict</b> A group of tourists visit the Eiffel tower in Paris. There are 15 old people in the group and 3 young ones. The tourists can choose between taking the lift or climbing up the tower (which takes half an hour). Only one person wants to climb up the tower. Do you think it is more likely that it's an old person, or that it's a young one?</p> <ul style="list-style-type: none"> <li>• It's more likely that it's a young person. (representative)</li> <li>• It's more likely that it's an old person. (base rate)</li> <li>• Both are equally likely.</li> </ul>	<p><b>Non-conflict</b> In a bird-watching club there are 6 women and 30 men. After a nice bird-watching trip a member of the club decides on going to a pub to watch a cricket-match. Who do you think is going to the pub?</p> <ul style="list-style-type: none"> <li>• A man. (representative/base rate)</li> <li>• A woman. (other)</li> <li>• Both are equally likely</li> </ul>
<b>Engineers and lawyers problem – non-representative description</b>	
<p><b>Conflict</b> Laura is a member of a sailing club and she also sings in a choir. She has 20 friends in the sailing club and 4 friends in the choir. Now she's going to Spain with a friend. Do you think it is more likely that she goes to Spain with somebody from the sailing club or with somebody from the choir?</p> <ul style="list-style-type: none"> <li>• It's more likely that it's somebody from the choir. (other)</li> <li>• It's more likely that it's somebody from the sailing club. (base rate)</li> <li>• Both are equally likely.</li> </ul>	<p><b>Non-conflict</b> In a chocolate factory 10 people's offices are in Building A, and 2 people are working in Building B. They are working on a new chocolate drink, and they want to decide who should try the drink first, so they organize a raffle. Do you think it is a person who works in building A or is it a person who works in Building B who's going to try out the drink first?</p> <ul style="list-style-type: none"> <li>• It's more likely that it's somebody who works in Building A. (base rate)</li> <li>• It's more likely that it's somebody who works in Building B. (other)</li> <li>• Both are equally likely.</li> </ul>
<b>Positive/negative recency</b>	
<p><b>Conflict</b> Sarah and Jack have four sons already. Now they are going to have another baby. Do you think it is more likely that the new baby is a boy or that it's a girl?</p> <ul style="list-style-type: none"> <li>• It's more likely that it's a boy. (positive recency)</li> <li>• It's more likely that it's a girl. (negative recency)</li> <li>• Both are equally likely.</li> </ul>	<p><b>Non-conflict</b> The Black Eagles basketball team are playing a match against the Deer Devils. Both teams won all their games so far in this season. Basketball experts say that both teams are in very good shape, and that it's impossible to tell in advance who will win the match tonight. The last eight times when Deer Devils played against Black Eagles, Deer Devils won 4 times, and Black Eagles also won 4 times. Who do you think will win the match tonight?</p> <ul style="list-style-type: none"> <li>• It's more likely that the Black Eagles will win. (response 1)</li> <li>• It's more likely that the Deer Devils will win. (response 2)</li> <li>• Both are equally likely.</li> </ul>
<b>Conjunction fallacy</b>	
<p><b>Conflict</b> Sue is a very intelligent woman, who works in a hospital. She wears glasses and a green uniform. Her bookshelves in her office are full of medical books. Mark the following statements with number 1 to 4 according to how likely they are. (1: most likely, 4: least likely)</p> <p>_____ Sue is a plumber. (non-representative 1)</p> <p>_____ Sue is a doctor. (representative)</p> <p>_____ Sue is a doctor and a mechanic. (representative + non-representative 2)</p> <p>_____ Sue is a mechanic. (non-representative 2)</p>	<p><b>Non-conflict</b> Brian has a studio, where he works alone. He is a very creative man, and he likes to experiment with colours. He takes his work to exhibitions, and sells some of them too. Mark the following statements with number 1 to 4 according to how likely they are. (1: most likely, 4: least likely)</p> <p>----- Brian is an aerobics instructor. (non-representative)</p> <p>----- Brian is a painter. (representative)</p> <p>----- Brian is an aerobics instructor and an accountant. (representative + non-representative 2)</p> <p>----- Brian is an accountant. (non-representative 2)</p>

Figure 3.1. Examples of the conflict and non-conflict versions of the conjunction fallacy task, the two types of engineers and lawyers task, and the positive/negative recency problem.

The problems used in the heuristics and biases tradition typically involve a conflict between people's intuitions and a statistical rule (e.g., the conjunction rule), or statistical information (e.g., base rates). As a result, an increase in correct responses necessarily implies a decrease in incorrect responding. According to Ferreira, Garcia-Marques, Sherman and Garrido (2006) this is problematic, because no response is "process pure". For example, giving a normatively correct response does not imply that the participant relied solely on Type 2 thinking when they gave that response. Using conflict tasks alone does not make it possible to separate out the contribution of automatic and rule-based processes to each response. For example, in the case of the *engineers and lawyers problem with a representative description* (see Figure 3.1.) there is evidence that participants' responses are influenced by the conflicting base rate information, even though their responses are more strongly influenced by the description of the person (Kahneman & Tversky, 1972). Ferreira et al. (2006) applied the process dissociation framework to investigate judgments under uncertainty in order to separate the contributions of automatic and controlled processes. The procedure makes use of a facilitation paradigm by including a condition in which automatic and controlled processes act in concert (we called it the *non-conflict* condition), as opposed to an interference condition (which we called the *conflict* condition) in which the two processes act in opposition. Assuming that both processes contribute to performance and operate independently, estimates of each can be obtained by comparing performance across the two conditions. Ferreira et al. (2006) used a formula to estimate the contribution of automatic and rule based processes to performance on the tasks. Instead of adopting the formula used by Ferreira and colleagues, in the present study we looked at performance on the conflict and non-conflict problems separately. In contrast with the previous studies we investigated normative performance (instead of heuristic responding) on both versions of the task, and we also looked at the difference between normative performance on conflict and non-conflict problems (i.e., we subtracted normative responding on conflict

problems from normative responding on non-conflict problems). In this way we gained an estimate of how susceptible participants are to salient information that interferes with normative performance.

In sum, we used five different types of task, each measuring a different reasoning heuristic (the two different versions of the engineers and lawyers problem, the positive/negative recency effect, the conjunction fallacy, and the belief bias). We used 16 syllogistic reasoning problems, and we had four different tasks measuring each of the other heuristics. That is, there were 16 judgment tasks altogether. In the case of the judgment tasks there were two sets of problem. Problems with the same content were presented in one set in a conflict version, and in the other set in a no-conflict version. We used a between-subjects design where children solved the non-conflict version of two tasks out of the four tasks of the same type, and the conflict version of the other two. Thus, they never solved both the conflict and non-conflict version of the same problem, and each heuristic was measured by two conflict and two non-conflict problems. We computed the effect of heuristic pull by comparing the proportion of children who chose the normative response in the non-conflict tasks and in the conflict versions of the same tasks. The full list of judgment tasks can be found in Appendix B, and the syllogistic reasoning problems are included in Appendix C. (Instructions can also be found in the appendices.)

A final question addressed by the present experiment is the relationship between analogical reasoning and heuristic responses. Analogy enables us to reason about novel phenomena, to identify relevant information on the basis of relational similarity, and to transfer knowledge from an initial learning context to future environments (e.g., Hummel & Holyoak, 2005). Analogy is a conceptual strategy in which a source object is represented as similar to a target object and correspondences are mapped between the two, based on the similarity of relations in each. During the mapping process both analogues have to be represented and manipulated in working memory (e.g., Holyoak & Thagard, 1989). These

operations require considerable cognitive effort. Although even 3-year-olds can solve simple analogical problems, this ability develops considerably during childhood (Halford, 1993).

A recent developmental study (Richland, Morrison & Holyoak, 2006) with participants from several age groups (3-4, 6-7, 9-11, and 13-14 years of age) systematically manipulated both relational complexity (i.e., the number of distinct units of information that must be processed in parallel while being maintained in working memory in order for a reasoner to complete a task - Andrews & Halford, 2002), and featural distraction in an analogy task based on finding correspondences between simple visual scenes (line drawings). Pre-tests confirmed that the critical relations (e.g., "chasing") were recognized by even the youngest children. Richland et al.'s results revealed that the development of the ability to reason analogically interacts with both relational complexity and featural distraction, with both of these sources of difficulty having a greater impact on younger than older age groups. In addition, younger children were especially likely to make the error of choosing a perceptual/semantic distractor (e.g., matching a cat to another cat when these did not play parallel roles). These findings suggest that changes in analogical reasoning with age depend on the interplay among increases in relevant knowledge, the capacity to integrate multiple relations, and inhibitory control over featural distraction (cf. Diamond, 2006). Similar conclusions have been reached based on studies with adults (Cho, Holyoak & Cannon, 2007). For example, a dual-task study demonstrated that either a verbal working memory load or an executive load disrupts analogical reasoning (Waltz, Lau, Grewal & Holyoak, 2000), increasing featural relative to relational mapping. It has also been found that people become more susceptible to interference effects at older ages (Viskontas et al., 2004). Patients with damage to the prefrontal cortex (a neural substrate of executive functions) are poor at analogical reasoning, especially in the presence of distraction (Krawczyk et al., 2008; Morrison et al., 2004).

Kahneman and Frederick (2002) propose that heuristic responses are the result of attribute substitution, and others (e.g., Stanovich, 1999; Klaczynski, 2001) suggest that the



reason that heuristics can lead to biases is that sometimes they are inappropriately overgeneralized. Based on these claims, it could be hypothesized that heuristic reasoning might be reliant on processes similar to that underlying analogical inference. On the other hand, analogical reasoning is demanding of executive resources, and it requires the ability to control interference, and to decouple representations (i.e., to maintain representations in the presence of distraction). According to Stanovich (1999, 2006) decoupling is one of the main functions of Type 2 processes. Based on this conjecture, we could expect analogical reasoning to be closely related with Type 2 reasoning and the ability to inhibit heuristic responses (i.e., to resist the interference stemming from salient, but irrelevant information).

Experiment 1 demonstrated that a number of (and possibly all) reasoning heuristics increase with age, and producing them requires conscious effort in the case of children between the age of 5 and 11. In Experiment 2 we found that during adolescence the use of some heuristics shows different developmental patterns. Children committed the conjunction fallacy more often as they got older, whereas other heuristics remained stable with development. As noted by Klaczynski (e.g., 2009), and as predicted by the fuzzy-trace theory (e.g., Reyna & Ellis, 1994) normative/heuristic responding on these tasks and cognitive capacity were unrelated. Finally, executive function skills (i.e., inhibition and set-shifting ability) seemed to be related with reasoning performance on these tasks, although the direction of this relationship was inconsistent across tasks. Building on these findings we can expect that susceptibility to heuristics in the present study will either remain stable, or it will increase with age. We can also expect that susceptibility to heuristics will be related to analogical reasoning performance. Stanovich (1999, 2006) proposed that decoupling is the most important Type 2 process, and he related this process with the ability to resist heuristic processing. That is, in this view, analogical reasoning ability should be negatively correlated with susceptibility to heuristics. On the other hand, analogical reasoning is closely related with executive functioning. Based on the findings of Experiment 2, we could expect a mixed relationship between executive

functions/analogical reasoning and heuristic reasoning (i.e., both positive and negative relationships are possible, as well as a lack of relationship).

## Method

### *Participants*

76 children (43 girls) between the age of 11 and 16 years took part in the study. Participants were recruited from two Plymouth secondary schools. In order to study age effects we divided our sample into two age groups (43 younger children with a mean age of 12 years, and 34 older children with a mean age of 14.9 years).

### *Materials*

*Judgment tasks:* Each type of heuristic was measured by two conflict and two non-conflict tasks. Examples of the conflict and non-conflict versions of each task can be found in Figure 3.1. Children were given a normative point for choosing a response that corresponded to the statistical information given in the task, or the logical structure of the task. In the case of conflict tasks normative responses were in conflict with a salient cue, whereas in the case of non-conflict tasks normative responses were supported by salient cues.

The first task was the *engineers and lawyers problem with a representative description*. In the non-conflict version of this task base-rates and representativeness pointed to the same response, whereas in the other version they were in conflict. In the case of a conflict people tend to favour the representative option instead of the response that corresponds to the base-rates, which is considered to be the normative response (Kahneman & Tversky, 1973).

The second task was the *engineers and lawyers problem with a non-representative description*. This task looked at the tendency to disregard base-rates when participants are provided with irrelevant information. The non-conflict version involved scenarios that emphasized a random selection of a person/object from two groups (with no individuating

information). In the conflict version there was no indication of randomness, and the two groups had different characteristics (although these differences were not informative.) Kahneman and Tversky (1973) found that in this situation people are inclined to disregard base rates (which would be the normative response) and judge the probability of both outcomes as equally likely.

The third task investigated the *positive/negative recency effect*. In the non-conflict version of the task the event to be predicted followed a sequence where both possible outcomes occurred with equal probability. In the conflict version the event to be predicted followed a sequence where only one outcome occurred several times. The *positive recency* (or *hot hand*) effect consists of expecting the same outcome that repeatedly occurred to happen again. The *negative recency effect* (or *gambler's fallacy*) is to predict the end of the sequence, and to expect the occurrence of a different outcome. The normative response is to say that both possible outcomes are equally likely.

The *conjunction fallacy* task required children to rate the probability of a representative event, two non-representative events, and the conjunction of either one of the non-representative events and the representative event, or the two non-representative events happening the same time. In the non-conflict version of the task children committed the fallacy if they judged the probability of one of the non-representative events happening alone to be smaller than the probability of the conjunction of the two non-representative events happening the same time. In the case of conflict tasks children committed the fallacy if they rated the conjunction of the representative and non-representative event as more likely to be true than the non-representative event alone (the non-representative event included in the analyses was the same for both the conflict and non-conflict version of the same task, although these were administered to a different group of participants). Children who did *not* commit the fallacy were given a normative point.

*Syllogistic reasoning tasks:* There were 16 problems, and the validity and believability of the tasks were systematically manipulated. There were 8 non-conflict problems (4 valid believable, and 4 invalid unbelievable syllogisms), and 8 conflict problems (4 valid unbelievable and 4 invalid believable syllogisms). Participants were given a normative point if they accepted a valid or rejected an invalid conclusion (regardless of believability). Examples for each type of task can be seen in Figure 3.2.).

	<b>Valid</b>	<b>Invalid</b>
<b>Believable</b>	<p><i>All nurses are sabs. All sabs are caring. Therefore, in this village all nurses are caring.</i></p> <p>Yes                      No, not necessarily</p>	<p><i>All funny people are bocs. All bocs are comedians. Therefore, in this village all comedians are funny.</i></p> <p>Yes                      No, not necessarily</p>
<b>Unbelievable</b>	<p><i>All politicians are mids. All mids are shy. Therefore, in this village all politicians are shy.</i></p> <p>Yes                      No, not necessarily</p>	<p><i>All quiet people are curs. All curs are salesmen. Therefore, in this village all salesmen are quiet.</i></p> <p>Yes                      No, not necessarily</p>

Figure 3.2. Examples for the different types of syllogistic reasoning problems.

*Measures of cognitive capacity:* as in the previous experiments we used the short form of the WISC as a measure of general intelligence, and the counting span task as a measure of working memory. In addition, we used Set I of the Advanced Progressive Matrices test (Raven, Raven & Court, 1998). Set I of the APM consists of 12 items, and it is usually used as a practice and screening set for the full test, and draws upon all the intellectual processes sampled on the full test (although it does not extend to the highest complexity levels). The WISC provides an estimate of both general knowledge and computational capacity, whereas the Raven test is an indicator of abstract reasoning ability. Moreover, the APM can be considered as a non-thematic test of analogical reasoning (measuring the ability to create and maintain temporary relationships in working memory, although there is no distraction present).

*Measures of analogical reasoning ability:* the *picture analogy task* was adapted from Krawczyk et al. (2008). The instructions can be found in Appendix D. These problems

were presented as pictures in the format A:B::C:?.(e.g., sandwich: lunchbox :: hammer: ?). Participants had to complete the analogy with one of four answer choices presented as separate pictures beneath the problem. In *distractor* problems, the four response options contained the correct solution (e.g., toolbox), together with a perceptual distractor (e.g., gavel), a semantic distractor (e.g., nail), and a picture unrelated to the C term (e.g., ribbon). In *non-distractor* problems, the two distractors were replaced by two additional unrelated items. The first two problems (a distractor and a non-distractor one) were for practice and were not included in the analyses. These were followed by another 8 distractor and 8 non-distractor problems (in a fixed random order). See Figure 3.3. for examples of a distractor and a non-distractor problem. The particular items used in the distractor and non-distractor conditions were counterbalanced across participants within each group.

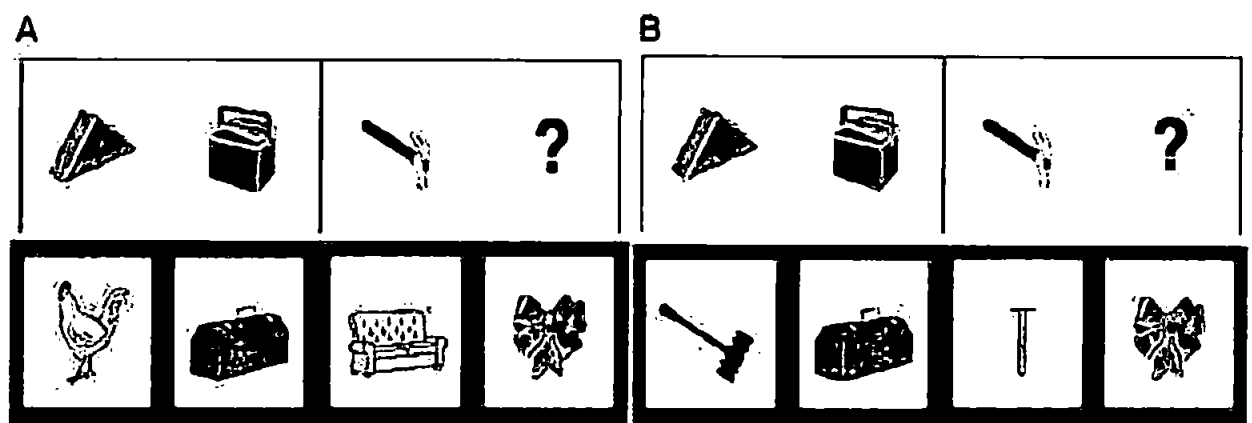


Figure 3.3. *Non-distractor and distractor versions of a picture analogy problem (A: non-distractor; B: distractor).*

The *scene analogy task* was adapted from Richland et al. (2006). The problems consisted of a pair of pictures (source and target analogues, arranged vertically). In each picture four or five objects appeared in simple relationships (see Figure 3.4.). Two or three of these objects were involved in a critical relationship, such as “hanging” (e.g., a baby monkey hanging on an adult monkey, which was hanging on the trunk of an elephant). An arrow pointed to one of the objects in the source scene (e.g., the adult monkey), and participants had to find the object corresponding

to this object in the target scene (e.g., a girl, given that a doll is hanging on a girl who in turn is hanging on a tree). There were 20 problems, five in each of four conditions that systematically varied the number of relation tokens that needed to be mapped, one or two (e.g., one or two tokens of the “hanging” relation in each picture), and the presence or absence of an object in the target scene that was either featurally similar to (distractor condition) or dissimilar to (non-distractor condition) the object to be mapped in the source scene (e.g., a monkey in the target picture that visually and semantically resembled the adult monkey in the source, versus a cat that did not resemble the object in the source). The particular pairs of scenes used in each of the four conditions were counterbalanced across participants within each group. The instructions used can be found in Appendix D.

### *Procedure*

Children were administered the judgment and the syllogistic reasoning tasks, and the analogy problems (including the Raven test) first. Due to the limited availability of the children from one school (n=33), the syllogistic reasoning problems were not administered to them. The working memory and general intelligence measures were administered at a later session. Both sessions took approximately 40 minutes.

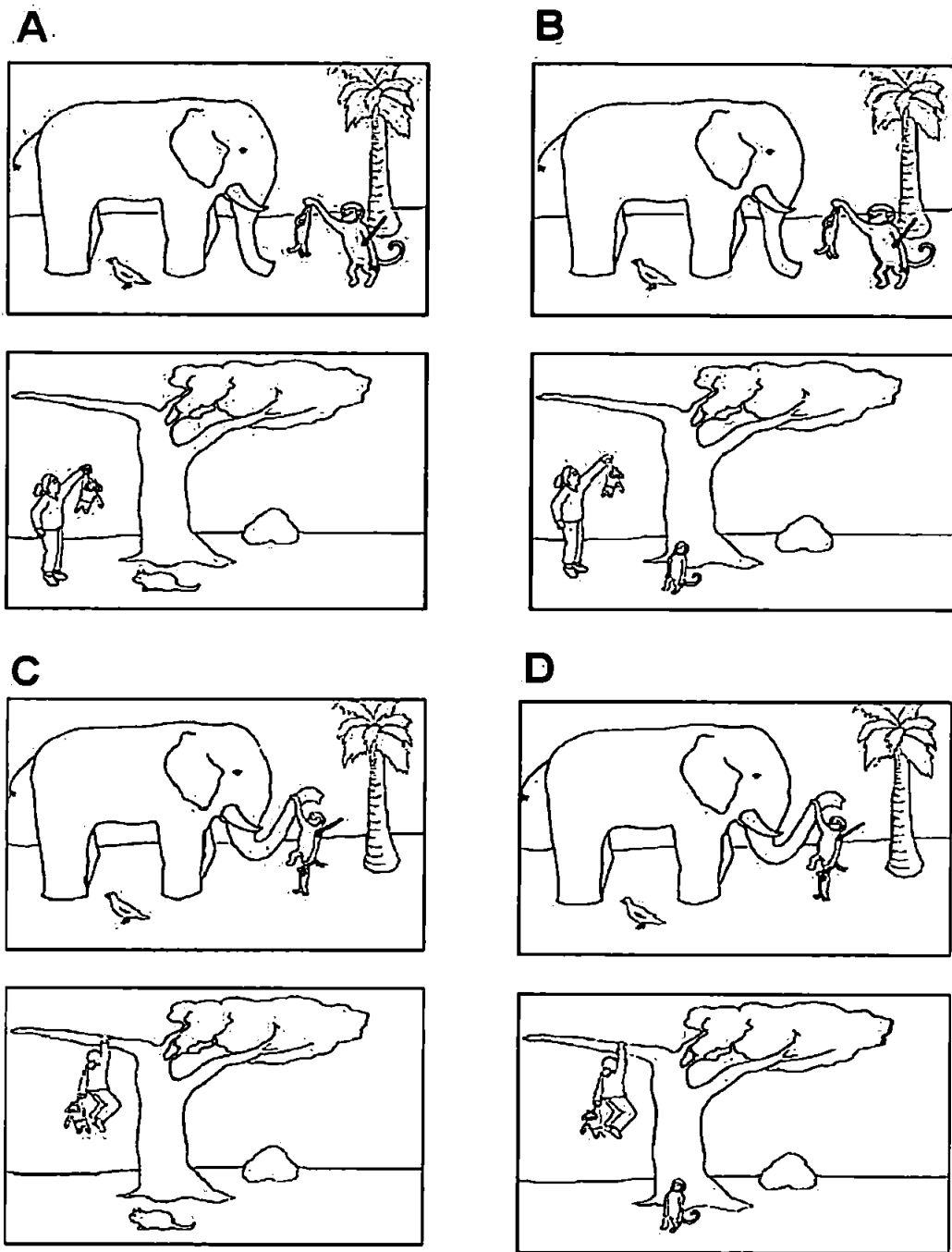


Figure 3.4. *Different versions of a scene analogy problem: (A) one relation/non-distractor; (B) one relation/distractor; (C) two relations/non-distractor; (D) two relations/distractor.*

## Results

### *Age-, conflict- and content-effects on the judgment tasks*

2x2 mixed ANOVA analyses were run for each type of task with age group as a between-subjects variable and problem type (conflict/non-conflict) as a within-subjects

variable. The difference in normative responding across conflict and non-conflict tasks was analysed. In the case of non-conflict tasks both normative considerations and the use of salient cues result in the same (normative) response. By contrast, in the case of conflict tasks participants giving a normative response have to be able to suppress the interference resulting from salient, conflicting information. That is, analysing the difference between normative responding on the conflict vs. non-conflict tasks makes it possible to estimate the amount of interference (i.e., “heuristic pull”) that participants experience in the case of a heuristic-normative conflict. A supplementary table for the following analyses is Appendix S3 Table 1.

The first task was the *engineers and lawyers problem with a representative description*. Table 3.8. displays the proportion of children who chose each type of response. The ANOVA indicated that children chose the normative response (corresponding to the base rates) more often in the case of non-conflict tasks than in the case of conflict tasks ( $F(1,73)=149.10, p<.001, \eta_p^2=.67$ ). That is, conflicting representative information had a large effect on children’s responses. There was no effect of age or a conflict by age group interaction ( $F_s<2.07, p_s>.15$ ).

Table 3.8. *The proportion of participants choosing each type of response across age groups and conflict conditions on the engineers and lawyers problem with a representative description.*

	Conflict			Non-conflict		
	repr.	base rate (non-repr.)	equally likely	repr. / base rate	non-repr.	equally likely
Younger (n=42)	.74 (.34)	.15 (.26)	.11 (.21)	.78 (.35)	.09 (.22)	.13 (.27)
Older (n=34)	.68 (.32)	.16 (.24)	.16 (.24)	.81 (.25)	.01 (.09)	.18 (.04)



The second task was the *engineers and lawyers problem with a non-representative description* (see Table 3.9. for descriptives). Data from 15 children had to be eliminated because they were given a different version of this task by mistake. The ANOVA indicated a significant effect of conflict ( $F(1,58)=30.51, p<.001, \eta_p^2=.35$ ). That is, children took base rate information less into consideration in the case of conflict tasks, when they were provided with some irrelevant individuating information. Additionally, there was a significant effect of age ( $F(1,58)=6.51, p<.05, \eta_p^2=.10$ ). This showed that older children took base rate information more into account than younger children.

Table 3.9. *The proportion of children choosing each type of response across age groups and conflict/non-conflict tasks on the engineers and lawyers problem with a non-representative description.*

	Conflict			Non-conflict		
	other	base rate	equally likely	other	base rate	equally likely
younger	.38 (.37)	<b>.24 (.33)</b>	.38 (.37)	.15 (.23)	<b>.56 (.44)</b>	.29 (.35)
older	.14 (.22)	<b>.40 (.29)</b>	.46 (.37)	.04 (.14)	<b>.77 (.33)</b>	.19 (.24)

The third task investigated the *positive/negative recency effects* (see Table 3.10. for descriptive statistics). A between-subjects *t* test indicated that children displayed no preference for giving either positive or negative recency responses ( $t(75)=1.07, p=.29$ ). Consequently, we merged the two types of response into a single heuristic score. As a result, normative and heuristic responses were mutually exclusive in the case of this task.

Table 3.10. *The proportion of participants giving normative/non-normative responses on the positive/negative recency problems across age groups and conflict/non-conflict tasks.*

	Conflict		Non-conflict	
	Positive/negative recency	equally likely	other	equally likely
Younger	.44 (.43)	<b>.56 (.43)</b>	.22 (.37)	<b>.78 (.37)</b>
Older	.34 (.40)	<b>.66 (.40)</b>	.03 (.12)	<b>.97 (.12)</b>

The ANOVA on the normative (as well as the heuristic/“other”) responses indicated a main effect of conflict ( $F(1,73)=32.87, p<.001, \eta_p^2=.31$ ) and a main effect of age ( $F(1,73)=4.01, p<.05, \eta_p^2=.05$ ). There was no age by problem type interaction. Older children were more likely to give the correct equally likely response, and all the children gave less correct responses in the case of conflict tasks.

Table 3.11. *The proportion of participants choosing each type of response on the conjunction fallacy problems across age groups and conflict/non-conflict tasks.*

	Conflict		Non-conflict	
	conjunction fallacy	normative	conjunction fallacy	normative
younger	.89 (.21)	<b>.11 (.21)</b>	.32 (.32)	<b>.68 (.32)</b>
older	.82 (.33)	<b>.18 (.33)</b>	.45 (.37)	<b>.55 (.37)</b>

In the case of *the conjunction fallacy* task (see Table 3.11.) data from 18 children had to be eliminated because they either responded by circling a single response or they rated multiple responses as equally likely. As in the case of the positive/negative recency task, heuristic and normative responses were mutually exclusive. The ANOVA indicated that children were more likely to commit the conjunction fallacy if the conjunction contained a representative event than when it consisted of two non-representative events

( $F(1,55)=80.49, p<.001, \eta_p^2=.59$ ). There was no age effect or age by problem type interaction.

*Performance on the syllogistic reasoning problems*

This task was completed by only 44 out of the 77 students who solved the other tasks (this was because participants from one school were only available for a limited time). 28 children were from the younger group (mean age 12.1 years) and 16 children were from the older group (mean age 14.6). Due to the fact that there were only two response options available, belief-based and logical responses in the case of the conflict tasks, and belief-based/normative and other responses in the case of non-conflict tasks were mutually exclusive (see Table 3.12). The ANOVA on the normative responses showed a significant effect of conflict ( $F(1,43)=72.25, p<.001, \eta_p^2=.63$ ) which indicated that participants gave less normative responses when there was a conflict between logical structure and believability. The effect of age, and the conflict by age interaction were not significant ( $F_s>1.85, p_s>.18$ ).

Table 3.12. *The proportion of belief-based and logical responses across problem types and age groups on the syllogistic reasoning problems.*

	Conflict		Non-conflict	
	Belief-based	Normative	Belief-based/normative	Other
Younger ( n=28)	.78 (.22)	.22 (.22)	.79 (.20)	.21 (.20)
Older (n=16)	.64 (.25)	.36 (.25)	.77 (.17)	.23 (.17)

*Analogical reasoning performance*

A supplementary table for the following analyses is Appendix S3 Table 2. The first analysis was a comparison between younger and older children’s performance on the Raven test. This test was solved by 71 out of the 75 children in the sample. The

independent samples *t* test indicated that older children ( $M=7.97$ ,  $SD=2.33$ ) performed significantly better than younger children ( $M=6.39$ ,  $SD=2.92$ ;  $t(69)=2.44$ ,  $p<.05$ ).

The next analysis investigated children’s performance on the picture analogy tasks (see Table 3.13). This task was solved by 75 out of the 76 children in the sample. The 2x2 mixed ANOVA with distraction (present/absent) as a within-subjects and age group as a between-subjects variable indicated a significant effect of distraction ( $F(1,73)=100.95$ ,  $p<.001$ ,  $\eta_p^2=.58$ ). There was no effect of age.

Table 3.13. *The proportion of correct responses on the distractor and non-distractor picture analogy problems across age groups.*

	non-distractor	distractor
younger	.94 (.10)	.71 (.19)
older	.94 (.09)	.73 (.23)

A comparison between the proportion of perceptual, semantic and “other” (neither perceptual, nor semantic) errors across age groups (see Table 3.14) indicated no difference between younger and older children. Paired samples *t* tests indicated that participants made significantly more semantic than perceptual ( $t(74)=9.71$ ,  $p<.001$ ) or “other” ( $t(74)=7.65$ ,  $p<.001$ ) errors.

Table 3.14. *The proportion of different types of errors on the picture analogy problems.*

	perceptual	semantic	other
younger	.02 (.03)	.12 (.09)	.04 (.06)
older	.02 (.03)	.12 (.10)	.03 (.06)

The following analyses concern students’ performance on the scene analogy tasks (see Table 3.15.). A 2x2x2 mixed ANOVA with relational complexity (one/two relations) and distraction (non-distractor/distractor) as within-subjects variables, and age groups as between-subjects variables indicated main effects of relational complexity ( $F(1,73)=22.88$ ,

$p < .001$ ,  $\eta_p^2 = .24$ ), distraction ( $F(1,73) = 48.92$ ,  $p < .001$ ,  $\eta_p^2 = .40$ ), and age ( $F(1,73) = 5.68$ ,  $p < .05$ ,  $\eta_p^2 = .07$ ), and a relational complexity by distraction interaction ( $F(1,73) = 8.96$ ,  $p < .01$ ,  $\eta_p^2 = .11$ ). That is, children performed worse on the scene analogy problems when the tasks required them to map two relations simultaneously, and they also performed worse in the presence of distraction. Older children generally performed better on these tasks. Finally, when they had to do a single mapping, children's performance decreased in the presence of distraction. By contrast, when they had to map two relations, the presence or absence of distraction did not have a further effect on their performance.

Table 3.15. *The proportion of correct responses on the scene analogy problems across age groups, distraction conditions and relational complexity.*

	One relation		Two relations	
	Non-distractor	Distractor	Non-distractor	Distractor
younger	.83 (.19)	.63 (.26)	.70 (.24)	.57 (.26)
older	.92 (.16)	.69 (.23)	.75 (.17)	.70 (.22)

A comparison between the number of perceptual, relational and "other" (neither perceptual nor relational) errors made by the two age groups on the scene analogy problems (see Table 3.16) indicated that older children made less "other" errors than younger children ( $t(73) = 2.03$ ,  $p < .05$ ). Additionally, the number of "other" errors was significantly less than the number of perceptual ( $t(74) = 4.88$ ,  $p < .001$ ) and relational errors ( $t(74) = 3.77$ ,  $p < .001$ ), whereas there was no difference between the number of relational and perceptual errors.

Table 3.16. *The proportion of perceptual, relational and other errors on the scene analogy problems across age groups.*

	perceptual	relational	other
younger	.12 (.11)	.13 (.07)	.07 (.11)
older	.10 (.09)	.10 (.06)	.03 (.04)

Finally, Table 3.17. displays the correlations between the different measures of analogical reasoning. For this analysis the distractor and non-distractor items across relational complexity levels on the scene analogy problems were combined. The reasonable sized correlations between tasks indicated that analogical reasoning on the different problems was underlain by similar processes, although the lower correlations across tasks as compared to within tasks indicated that there were some task-specific processing requirements as well which were not shared by all analogical reasoning tasks. Although there was no significant relationship between the non-distractor versions of the picture and scene analogy problems, performance on the distractor versions of these problems were related.

Table 3.17. *Correlations between the different measures of analogical reasoning.*

	1.	2.	3.	4.	5.
1. picture analogy non-distractor	--				
2. picture analogy distractor	.41**	--			
3. scene analogy non-distractor	.18	.25*	--		
4. scene analogy distractor	.32**	.31**	.54**	--	
5. Raven	.40**	.40**	.41**	.41**	--

*Relationships between the measures of individual differences in age, cognitive capacity, and analogical reasoning performance*

Table 3.18. displays the correlations between age, working memory capacity, general intelligence, and analogical reasoning ability (the combined z scores of students' performance on the three analogy tasks). *N*s for these measures vary between 50 and 70 due to software problems, experimenter errors or the fact that some participants were not available for a second testing session (see details below).

There was a positive correlation between age and the measures of intelligence and analogical reasoning, but unlike in Experiment 2 there was no relationship between age and working memory. Working memory scores were positively related with the measures of intelligence, and analogical reasoning. The positive correlation between analogical reasoning and working memory is in line with results reported in the analogical reasoning literature (e.g., Waltz et al., 2000). General intelligence and analogical reasoning ability were also correlated with each other.

Table 3.18. *Correlations between age, working memory, general intelligence, and analogical reasoning ability.*

	1.	2.	3.	4.
1. age	--			
2. counting span	.18 (n=59)	--		
3. WISC	.42** (n=66)	.42** (n=59)	--	
4. analogical reasoning	.28* (n=70)	.37** (n=54)	.59** (n=59)	--

*Relationships between performance on the judgment and reasoning tasks*

Table 3.19. displays the correlations between susceptibility to the different types of biases (i.e., we computed the difference between normative responding on the conflict vs.

non-conflict version of each task by subtracting the conflict normative score from the non-conflict normative score). That is, higher scores indicate greater susceptibility to the biases.

Ns for these correlations vary between 32 and 75 due to some children failing to respond to some tasks, and due to the fact that children from one school were not administered the syllogistic reasoning tasks. In the case of the positive/negative recency tasks we considered positive and negative recency responses together, because children did not seem to display a preference for either bias on these tasks (see above).

Table 3.19. *Correlations between susceptibility to different biases.*

	1.	2.	3.	4.
1. Engineers and lawyers problem, representative	--			
2. Engineers and lawyers problem, non-representative	.34** (n=60)	--		
3. Positive/negative recency	.08 (n=75)	.08 (n=60)	--	
4. Conjunction fallacy	.07 (n=57)	-.20 (n=45)	-.35** (n=57)	--
5. Syllogistic reasoning	.38** (n=45)	.46** (n=45)	-.09 (n=45)	.10 (n=32)

There was a positive relationship between susceptibility to giving heuristic responses to the different versions of the engineers and lawyers problem, and these were also positively correlated with susceptibility to the belief bias on the syllogistic reasoning problems. Additionally, there was a negative relationship between susceptibility to the conjunction fallacy and to the positive/negative recency effect. In contrast with the results of Experiment 2 where there was no evidence for a relationship between giving heuristic responses to different tasks, the present method seems to be more appropriate for detecting similarities across tasks. However, the fact that there are both positive and negative (as well as non-significant) relationships across tasks is in line with the results of Experiment 2 inasmuch as they suggest that the underlying processes can be different across tasks.



Finally, we also looked at the correlations between age, cognitive capacity and analogical reasoning and susceptibility to heuristics (that is, the difference between normative responding between non-conflict and conflict tasks – see Table 3.20.). Susceptibility scores were unrelated to age and cognitive ability. However, there was a positive relationship between susceptibility to the representativeness heuristic on the engineers and lawyers problem.

Table 3.20. *Correlations between susceptibility to different biases and age and the measures of cognitive capacity.*

	Age	Cognitive ability	Analogical reasoning
Engineers and lawyers problem, representative	.04 (n=75)	.12 (n=58)	.36** (n=69)
Engineers and lawyers problem, non-representative	.15 (n=60)	.04 (n=53)	.17 (n=57)
Positive/negative recency	.16 (n=75)	.13 (n=58)	-.06 (n=69)
Conjunction fallacy	-.22 (n=57)	-.15 (n=43)	-.14 (n=51)
Syllogistic reasoning	-.20 (n=45)	-.08 (n=44)	-.17 (n=44)

## Discussion

The main purpose of Experiment 3 was to further investigate the changes in heuristic responding during adolescence using a new set of tasks, and a novel method that makes it possible to estimate more precisely the effect of salient, distracting information on reasoning performance. Additionally, the relationships between analogical reasoning and susceptibility to heuristics were also analyzed. The fact that children tended to give fewer normative responses to conflict than to non-conflict tasks indicates that we were successful in measuring the effect of heuristics on using normative information.

An important finding of this study is that although susceptibility to heuristics remained stable with age, children's general intelligence, and their performance on the scene analogy problems, and on the Raven test improved with age. This is in line with the idea that normative competence increases during adolescence (e.g., Stanovich et al., 2008). However, this increase in competence is not obvious from children's performance on traditional heuristics and biases tasks (see e.g., Klaczynski, 2009). In fact, there was no relationship between susceptibility to heuristics and cognitive capacity. On the other hand, children were more likely to rely on base rates/probability information in the case of the non-representative version of the engineers and lawyers task, and in the positive/negative recency problems as they got older. This could be the result of children's increasing knowledge about how to make use of probabilistic information.

In contrast with Experiment 2, in the present experiment we found some evidence for relationships between susceptibility to the different biases (a positive relationship between susceptibility to heuristics on the engineers and lawyers problems, and on the syllogistic reasoning task, and a negative relationship between susceptibility to the conjunction fallacy and the positive/negative recency effect). This probably indicates that the measure of heuristic pull that we used in the present study was a more sensitive and purer measure of heuristic reasoning than conflict tasks alone. However, as in Experiment 2, the relationships were mixed, and they could not easily be explained by the individual differences variables that we investigated (i.e., age, cognitive ability, and analogical reasoning).

A final question that this study aimed to answer was whether there is a relationship between analogical reasoning and heuristic responding. Based on dual-process theories (e.g., Stanovich, 1999, 2006) we could expect that suppressing heuristic reasoning (i.e., the decoupling of representations) should be similar to analogical reasoning, which also requires the reasoner to extract the structure of the tasks, and maintain temporary links between representations in the presence of distraction. On the other hand, the description

of attribute substitution (e.g., Kahneman & Frederick, 2002) which defines heuristic reasoning as making links between similar situations, and using existing knowledge to answer novel problems sounds very similar to a description of analogical reasoning (see e.g., Hummel & Holyoak, 2005). These two possibilities might be reconciled if we suppose that initially heuristics are based on effortful processing, and a conscious attempt to integrate information from different sources (as indicated by the results of Experiment 1), and thus are positively related to participants' analogical reasoning ability. However, with practice heuristics become automatic and autonomous, and at this stage conscious effort is needed to inhibit the inappropriate use of heuristics, and thus we can expect a negative relationship between heuristic use and analogical reasoning. Finally, there could be a transition phase between these two stages, when participants are able to produce heuristic responses without much effort, but they are unable to suppress these responses. It is also possible that participants do not suppress an effortlessly produced heuristic response because they are unaware of the relevant normative rule (or they are unaware that the rule applies in the given situation). The present study provided evidence for a positive relationship between susceptibility to a certain heuristic (i.e., the representativeness heuristic in the case of the engineers and lawyers problem) and analogical reasoning ability, and there were also some tasks where there was no relationship between susceptibility to a heuristic and analogical reasoning. However, there was no evidence for a negative relationship between heuristic reasoning (or susceptibility to heuristics) and analogical reasoning (i.e., decoupling). Thus, although it is possible that decoupling plays a role in reasoning performance on heuristics and biases tasks in the case of adult participants, there was no evidence for such a link in our adolescent sample.

In sum, Experiment 3 demonstrated an increase in logical/statistical competence in the case of some tasks, and we also found an increase in general intelligence and performance on the Raven test and the scene analogy problems. At the same time and within the same population, the number of heuristic responses on conflict tasks remained

stable, and there was little evidence for an improvement in detecting the conflict between intuitions and normative rules, regardless of children's age or cognitive capacity. Finally, analogical reasoning was positively related to susceptibility to a certain heuristic which is the opposite of what we could expect based on an influential dual-process account of reasoning development (Stanovich et al., 2008).

#### *3.4. General discussion (Experiments 1-3.)*

The three experiments described in this chapter demonstrated that dual-process theories in their original form (e.g., Evans & Over, 1996; Stanovich 1999) which identify heuristics and normative rules with two separate reasoning systems (i.e, heuristic = Type 1 – automatic, effortless and simple; and normative = Type 2 - conscious, effortful and complex) cannot be generalised to children. In fact, Experiment 1 showed that in the case of children (between the age of 5 and 11) Type 2 reasoning leads to responses that are traditionally considered to be heuristic (i.e., Type 1). Our analyses also warn against the interpretation of the lack of correlation between cognitive capacity and reasoning performance as an indication that the reasoning process is effortless. Although there was no relationship between cognitive capacity (as measured by a composite score of the WISC and working memory) and heuristic reasoning in Experiment 2, there was some evidence for relationships between reasoning performance and executive functioning (as measured by the inhibition and the set-shifting tasks). Similarly, in Experiment 3 we developed a more sensitive measure of heuristic pull/susceptibility to heuristics, and whilst this was unrelated to cognitive capacity, it was related to analogical reasoning performance (which indexes the ability to make links between decoupled representations) – at least in the case of one task.

One aim of this series of studies was to explore the developmental trends in heuristic reasoning between the age of 5 and 16. The results of Experiment 1 showed an increase in heuristic responding between the age of 5 and 11 on 4 out of 5 tasks (with the exception of

the sunk cost fallacy). At the same time both normative and “other” responding tended to decrease on these tasks. This increase in heuristic responding was positively correlated with children’s cognitive capacity, and we proposed that this was because of the effortful nature of the contextualisation process (or memory retrieval process) that is involved in producing these responses.

Experiments 2 and 3 showed that after the initial increase in heuristic responding during mid-, and late childhood, there is little change in heuristic responding between the ages of 11 and 16. However, if there was a change, heuristic responding tended to increase rather than decrease (which would be predicted by an influential developmental dual-process account – see e.g., Stanovich et al., 2008). In Experiment 2 there was an increase in the conjunction fallacy. All other heuristics remained stable across the young and mid-adolescent age groups. Although there was evidence for an increase in working memory capacity (in Experiment 2), and in WISC scores, and analogical reasoning performance (in Experiment 3), the changes in cognitive abilities were mostly unrelated to (or inconsistently related to) changes in reasoning performance.

Many theorists predict a change in children’s reasoning abilities when they reach adolescence. Some theorists propose that children become more logical (e.g., Stanovich et al., 2008), or that their thinking becomes more flexible and more subject to metacognitive control (e.g., Klaczynski, 2007), or that they become more able to inhibit irrelevant associations (e.g., Brainerd et al., 2004; De Neys & Everaets, 2008) around the age of 12 or 14. These changes have been variously attributed to the development of working memory capacity, inhibition, or metacognitive skills. We found little direct supporting evidence for these claims, apart from the increase in analogical reasoning performance which requires the inhibition of distracting stimuli. However, it is still worth noting that as the processing demands of contextualisation become relatively lower after mid-childhood, there should be more cognitive resources available for other operations. Nevertheless, this

is clearly not reflected in children's performance on the heuristics and biases tasks that we administered in our studies.

These results are similar to the patterns described by Klaczynski (2009) in a recent review paper. Adopting his argument, these mixed results can be considered as the consequences of differences in individual children's experiences, their emotional responses, and beliefs. This interpretation is also consistent with the fuzzy-trace theorists' claim that reasoning depends on the gist of experiences that individuals extract from the events that they encounter (which is based on their past experiences). These results can also be reconciled with the latest model of Stanovich and West (2008) which places a great emphasis on individual's knowledge of relevant normative rules, as well as the necessity of recognizing the applicability of these rules to individual cases. (Thus, this model differs from the other two approaches in its emphasis on the role of Type 2 rather than Type 1 processes). However, the problem with all of these claims is that they provide little help with predicting individuals' future behaviour. Ironically, the latest developments of dual-process theories are very similar to the ideas of the fast and frugal heuristics approach. For example, Gigerenzer, et al. (2008) emphasize that the heuristics that people use in a given situation will largely depend on their individual learning history. Thus, it is hard to predict what sort of heuristic a given individual will use in a particular situation. In addition, according to this approach, even if they use a particular heuristic, this can lead to different responses depending on the person's individual learning history.

Overall, the results of Experiments 1-3 highlight the dynamic nature of the development of reasoning heuristics, and they also question the validity of approaches that identify normative responding with intelligence and cognitive effort, and heuristic responses with sloppy, low-effort reasoning. The present findings contradict the view which considers heuristics to be based on hardwired cognitive processes that are part of the basic human cognitive architecture. Instead they point to the role of cognitive development, experiences, and learning in the emergence of reasoning heuristics. Starting

from these assumptions we can hypothesize that the development of reasoning heuristics might be delayed or could show atypical patterns in children with developmental disorders, as their cognitive development (or at least some aspects of it) are usually delayed, and they might lack some experiences that typically developing children have access to.

A particularly interesting group in this respect is autistic children. On the one hand, individuals with autism are claimed to suffer from executive functioning problems (see e.g., Hill, 2004 for a review), that is, in dual-process terms they have a Type 2 deficit. On the other hand, mainstream cognitive theories of autism, such as the weak central coherence theory (e.g., Happé & Frith, 2006), and the “extreme male brain” theory of autism (e.g., Baron-Cohen, 2002) propose that individuals with autism have a decontextualised thinking style, and they are very good at detecting regularities, and extracting rules which is more in line with a Type 2 advantage. Chapter 4 gives a review of the literature on the cognitive theories of autism, and the existing research on reasoning in autism. Chapter 5 presents two experiments investigating the development of reasoning heuristics in autism.

## Chapter 4. Cognitive theories of autism, and research on reasoning in autism.

### 4.0. Introduction to Chapter 4.

The main aim of this chapter is to give a brief introduction into autism (and the cognitive theories of autism), to review the literature on reasoning and autism, and finally to explain how research with children with autism can possibly lead to a better understanding of the development of reasoning heuristics. In relation to the cognitive theories of autism, I will also discuss the issues of *specificity*, *uniqueness*, and *universality* (i.e., whether autism arises from a domain-specific factor or if there are multiple factors involved; whether these factors are unique to autism or if they are also involved in other developmental disorders; and whether these factors are present in every individual with autism or just in the majority).

### 4.1. Autism and the cognitive theories of autism

Autism is a behaviourally defined disorder, characterised by impairments in social communication, social interaction, and social imagination, together with a restricted range of interests, stereotyped and repetitive behaviours, and mannerisms (see e.g., DSM-IV). These behavioural symptoms are present in very early childhood, before the age of 3. Many children with autism (50 – 70%) have additional learning disabilities (i.e. an IQ lower than 70), while ‘high-functioning’ autistic children have IQs in the normal range. For some children with autism, language is limited or absent altogether. The term Asperger syndrome is usually used in the case of children with no apparent language delay. The prevalence of the broad spectrum of autistic disorders is around 5-6 per 1000 (Le Couteur & Baird, 2003).

Autism is defined by behavioural symptoms, and it can be caused by different factors. In a minority of individuals with autism (between 6-10%) there is a specific



medical cause. This is more common in autistic individuals with a more serious learning disability (Fombonne, 1999). There is also evidence for a genetic origin, where multiple genes are likely to be involved. In monozygotic twins the chance of concordance for autism is 60%, with a higher chance of a social impairment (Rutter, Silberg, O'Connor & Simonoff, 1999). Autism is also characterised by pervasive neurobiological abnormalities, which are not confined to particular regions of the brain (Belmonte, Cook, Anderson et al., 2004).

The core impairments and behaviours of autism vary greatly from person to person, as well as within a single individual. As autism is a developmental disorder, symptoms can be expected to change over time (Wing, 1988). In general, social and communicative symptoms improve significantly with time, whereas restricted and repetitive behaviours are more stable. This suggests that there might be different developmental trajectories for each of the symptom domains (Charman, Taylor, Drew et al., 2005). For these reasons, it is hard to precisely diagnose individuals within the autism spectrum, which is reflected in the (often inconsistent) use of various terms (i.e., autism, Asperger's syndrome, pervasive developmental disorder, etc.).

Individuals with autism sometimes possess some special skills (usually in the areas of music; drawing; mathematics; rote memory; constructional skills or manipulation of mechanical objects). These skills might be above the individual's general level of functioning, however, they usually do not exceed the level of what could be expected based on the individual's chronological age. Only in the case of high functioning individuals with no intellectual impairment can we expect some abilities well above chronological age (Baird, Cass & Slonims, 2003).

#### *4.1.1. The theory of mind deficit hypothesis of autism*

The most widely used test of theory of mind (ToM) is the unexpected transfer test, which was devised by Wimmer and Perner (1983). Participants watch a sequence of

events, where (as a result of an unexpected transfer of an object) one of the characters will have a false belief about the location of the object. Participants then have to make a judgement as to where the character will look for the object, and in order to give the correct answer they have to infer the mental state of the character (i.e., “the character believes that...”). The majority of children pass the false belief task when they are about 4 years old (Wellman, Cross & Watson, 2001). The first experiment to show that children with autism had a problem with solving the task was Baron-Cohen, Lesile and Frith (1985) who found that 80 percent of children with autism failed, and they concluded that these children had a ToM deficit. However, 20% of the sample did pass the task, which indicated that the deficit was not universal in autism (Happé, 1994). A meta-analysis by Happé (1995) also showed that ToM was strongly associated with verbal mental age and false belief in children with autism, and participants with a verbal mental age of 9 and above were almost certain to be able to pass.

In fact, individuals with high functioning autism with no general intellectual/language impairment can pass even second order false belief tasks (e.g., Baron-Cohen, 1989). Interestingly, as a response to this and similar findings, many researchers concluded that individuals with autism do have a ToM deficit; however, the tests most commonly used are not sensitive enough, or inappropriate to show this. For example, Frith, Happé and Siddons (1994) suggested that autistic passers of ToM tasks might have used non-mentalistic methods, and arrived at the correct solutions without actually inferring mental states. In order to overcome these problems, a range of “advanced” tests of ToM were developed specifically for the high functioning autistic population. These include the Eyes Test (Baron-Cohen, Jolliffe, Mortimore & Robertson, 1997), the Recognition of Faux Pas Test (Baron-Cohen, O’Riordan, Stone, Jones & Plaisted, 1999), and the Strange Stories test (Happé, 1994). Although even high functioning individuals tend to perform less well on these tasks relative to typically developing controls, they do not fail entirely (e.g., Baron-Cohen et al., 1997).

In sum, it seems that most individuals with autism have some problems with inferring the mental states of others in certain situations. This is no surprise, as the diagnostic criteria of autism involve problems with social interactions, social communication, and social imagination, thus, autistic people, by definition, have social impairments. However, it is less clear where these problems stem from. For example, performance on the false belief task does not only correlate with children's verbal abilities, but it is also related to their executive functioning (e.g., Carlson, Moses & Hix, 1998; Perner & Lang, 1999), and their ability to deal with relational complexity at the tertiary level (that is, to integrate three variables in a single cognitive representation - see e.g., Andrews, Halford, Bunch, Bowden, & Jones, 2003). The false belief task was also criticised on the basis that it is not only a confounded measure of ToM, but that the ability to reason about false beliefs is not even necessary for somebody to have a theory of mind (Bloom & German, 2000). Similar questions can be raised about "advanced" measures of theory of mind. Bull, Phillips, and Conway (2008) used dual-task manipulations of executive functions (inhibition, updating and switching) to investigate the role of these control functions in mental state and non-mental state tasks. They used the Eyes test (Baron-Cohen et al., 1997) and the Strange Stories test (Happé, 1994), together with control tasks that were very similar, but did not require the ability to reason about false beliefs. The Eyes test showed specific dual-task costs when concurrently performed with an inhibitory secondary task. By contrast, interference effects on the Stories task were general, occurring on the control tasks as well, and across all types of executive function, indicating that the processing demands were not related to the ToM component of the task.

The original conceptualisation of ToM deficit in autism (Baron-Cohen, 1995) involved the idea that ToM was based on modular processing, and the ToM module was selectively impaired in autism. Based on the above evidence, it seems clear that performance on almost all ToM tasks requires executive skills as well. However, the question remains whether there is also a modular/automatic and inborn component of

ToM. There is evidence from brain imaging research (e.g., Saxe, Schulz & Jiang, 2006) that certain parts of the brain are selectively implicated in belief attribution in adults. In addition, the behavioural evidence for the ability to quickly and effortlessly infer other's mental states in complex social situations, and the fact that even babies and animals are able to attribute belief-like states to others (see Apperly & Butterfill, 2009 for a review) seems to suggest that there should be some aspects of ToM reasoning which are independent of language and executive functions. Although we could expect that individuals with autism will be impaired in this aspect of ToM, there is no experimental evidence so far to support this.

In a number of influential publications, Karmiloff-Smith and colleagues (e.g., Karmiloff-Smith, Scerif & Ansari, 2003; Thomas & Karmiloff-Smith, 2002) criticized the idea that certain "cognitive modules" can be selectively impaired, and double dissociations might occur in developmental disorders. They based their argument on the following claims. Although there are specialised areas or cognitive modules in the adult brain, these emerge from a developmental process of modularization. This process is based on some innately specified starting points, but at the early stages of development, these are only domain-relevant, rather than domain-specific. The modularization process unfolds during development, and also as a result of specific environmental interactions. In the case of individuals with developmental disorders the effects of genetic mutation during early development are likely to be widespread across the developing system. Instead of neatly segregated patterns of intact and impaired cognitive modules, it is more likely that all domains develop atypically, although some domains may be more affected than others due to their particular problem space. For example, although individuals with Williams syndrome (WS) are better at verbal than non-verbal tasks, their language is not intact. Moreover, their language development follows an atypical developmental trajectory. As a result of the dynamic nature of development (for example, the changes in the regulation of gene expression over time), and the effects of environmental interactions, it is also likely

that the patterns of strengths and weaknesses within the same developmental disorder change over time. For example, children with both WS and Down syndrome (DS) show equal delay in language development, despite DS adult language being significantly worse than that of the WS adults'.

In sum, although it seems clear that individuals with autism are generally impaired in ToM reasoning, the nature and severity of their actual problems can vary widely, depending on the individual's cognitive abilities, and verbal skills. As autism is a developmental disorder, and its diagnosis is based on behavioural symptoms rather than its aetiology, it is possible that these impairments in different individuals are the results of a number of different factors, and it might not be possible to give a single explanation for these. At the very least it seems highly unlikely that these symptoms can be explained by the impairment of a single ToM module.

#### *4.1.2 The weak central coherence theory*

Another influential cognitive account of autism is the Weak Central Coherence (WCC) theory (e.g., Frith & Happé, 1994; Happé, 1999). One of the key strengths of this theory is that it explains some of the non-social, as well as the social features of autism. The WCC theory proposes that typically developing individuals tend to engage in global processing, building up a gist-based representation, whereas autistic individuals engage in more detailed, local or piecemeal processing. The WCC theory is supported by evidence that autistic children show very good performance on tasks that require attention to local features, such as the embedded figures (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983) and block design tests (Shah & Frith, 1993). The key feature of both the block design and the embedded figures tests is that a global pattern includes smaller constituent components. For individuals in the neurotypical population, the global figure is more salient than its constituent parts. Frith (1989, 2003) argues that individuals with autism show better performance on these tasks because they lack the cognitive drive to attend to the "big

picture” (i.e., they exhibit WCC) and as a result they can process local parts without global interference. In fact, there is evidence for a local to global interference in autism, which might result from a deficiency in inhibiting further processing of irrelevant details, that is executive function deficits (Rinehart, Bradshaw, Moss, et al., 2000 – see also Section 4.1.3.).

Happé (1996) found that participants with autism were less susceptible to visual illusions (e.g., Ponzo, Muller-Lyer, Titchener) than other groups. Happé claimed that this was because individuals with autism processed the stimuli in a piecemeal manner without integrating individual features with the context. However, later research showed that participants with autism were susceptible to the illusions when they were asked, for example, “which line *looks* longer”, but not when asked “which line *is* longer” (Brosnan, Scott, Fox & Pye, 2004).

In the domain of verbal processing, there is evidence that individuals with autism are less able than typically-developing children to benefit from sentence context in disambiguating the meaning of homographs (e.g., Frith & Snowling, 1983; Jolliffe & Baron-Cohen, 1999, although see Brock, Norbury, Einav & Nation, 2008). Other evidence suggests that autistic people are impaired in generativity (i.e., using their knowledge spontaneously and flexibly in novel situations; Peterson & Bowler, 2000). Autistic people appear to exhibit reduced processing of similarities between stimuli; instead, they are better at processing differences (Plaisted, 2001), although this characteristic may be restricted to perceptual processing. Autistic people also seem to have difficulty understanding metaphors and other types of non-literal language, such as irony (e.g., Pexman, 2008). Importantly, it appears that these effects cannot be attributed to deficits in the automatic inferences involved in text comprehension or to the lack of activation of relevant knowledge (Saldana & Frith, 2007).

The WCC theory has, however, been criticized based on evidence that in the general population various tasks that supposedly measure central coherence are not

correlated with each other (Pellicano, Maybery & Durkin, 2005). Moreover, some studies suggest that people with autism are able to process information globally when they are instructed to do so, although they process information locally when no such instructions are offered (e.g., Mottron, Burack, Stauder & Robaey, 1999). A recent study using different versions of the Block Design task demonstrated that locally oriented processing in autistic participants did not imply a deficit in forming global representations (Caron, Mottron, Berthiaume & Dawson, 2006). After reviewing a large number of empirical studies of coherence, Happé and Frith (2006) concluded that the finding of a local bias was robust. On the other hand, the findings regarding weak global processing are mixed. Happé and Frith also concluded that the local bias was not a side effect of executive dysfunction or theory of mind deficits, and that it could be overcome through conscious effort.

At the biological level, it has been suggested that WCC in autism might be the result of abnormal neuronal connectivity (due to either structural or functional differences). The suggested mechanisms include a lack of synchronisation in activation between relevant brain areas (Brock, Brown, Boucher & Rippon, 2002), reduced long-range and increased short-range physical connectivity (Just, Cherkassky, Keller & Minshew, 2004), or numerous and inefficient feedback connections resulting in a lack of top-down modulation of early sensory processing and a lack of integration of sensory processing with cognitive monitoring (Frith, 2003). These ideas are supported by a growing number of functional imaging studies showing reduced connectivity and a lack of top-down modulation particularly between the frontal cortex and other brain regions (e.g. Bird, Catmur, Silani, et al., 2006; Just, Cherkassky, Keller, et al., 2007). All of these theorists predict that the abnormal connectivity of different parts of the brain would result in a preserved or enhanced ability for detail-focused, piecemeal processing. What is unclear, however, is whether abnormal neuronal connectivity is universal in autism or if it is only present in a subgroup of autistic individuals. For example, White, O'Reilly and Frith (2009) have suggested that this might be only characteristic of those individuals with

autism who have macrocephaly (that is, increased head/brain size which is present in approximately 20% of the autistic population).

In sum, the strengths of the WCC theory include that it offers explanations for both some social and some non-social aspects of autism (for example, the restricted range of interests and repetitive behaviours, and autistic people's problems with understanding non-literal language). Another strength of the theory is that there are some very robust findings from the area of visual-spatial processing supporting it. However, it seems like the preference for local processing does not imply an inability to process information globally. The fact that explicit instruction increases global processing suggests that WCC (at least in the case of some autistic individuals) might result from executive function deficits (see Section 4.1.3.). Together with the evidence that WCC possibly arises due to abnormal neural connectivity in a subgroup of individuals, this raises the possibility that WCC in autism might emerge for different reasons in different individuals.

#### *4.1.3. The executive dysfunction theory of autism*

Similarly to the WCC theory, and in contrast with the ToM deficit hypothesis, the executive dysfunction account of autism (e.g., Ozonoff, Pennington & Rogers, 1991) is intrinsically domain-general. This theory was based on the observation that some symptoms of autism (i.e., the need for sameness, the tendency to perseverate, the need to be prompted in order to initiate certain behaviours or to switch set, and autistic individuals' problems with impulse control) are very similar to those shown by individuals with Dysexecutive Syndrome (Baddeley & Wilson, 1988). Such individuals have problems with executive functioning (EF) usually due to the damage of the prefrontal cortex (PFC), although PFC damage does not necessarily imply EF deficits (Shallice & Burgess, 1991), while some people with damage outside the PFC do show impairments (e.g., Levisohn, Cronin-Golomb & Schmahmann, 2000). Interestingly, the PFC is also thought to be related to ToM reasoning, and, thus, PFC might be a shared brain site for these functions (see



Apperly, Samson & Humphreys, 2005, for an overview), although this might stem from the intrinsic executive demands of most ToM tasks.

Executive function is defined as the ability to maintain an appropriate problem-solving set for attainment of a future goal which involves planning, the inhibition of prepotent but irrelevant responses, and flexibility of thought and action (Ozonoff, Pennington & Rogers, 1991). According to Miyake et al. (2000) the three main components of EFs are working memory, inhibition, and set-shifting (the ability to flexibly switch between mental sets, representations or rules). The executive dysfunction account of autism was first suggested by Ozonoff et al. (1991) who found that 96% of the autistic participants in their study performed below the mean scores of the control group, which they interpreted as an evidence for a universal EF deficit in autism. However, later studies did not replicate this finding. For example, Pellicano, Maybery, Durkin and Maley (2006) reported that executive problems were only found in 50 % of their sample (note that they defined a deficit as performance at least one standard deviation below the mean of the control group).

A large number of studies reported EF deficits in autism (see e.g., Hill, 2004 for a review). However, according to Hill (2004) the interpretation of these findings is problematic. One issue is that several different tasks (and even different versions of the same tasks) have been used to assess EFs, and the comparison groups also varied widely across studies. The way tasks are presented can also affect the performance of autistic participants (for example, in the case of the Wisconsin Card Sorting Task participants with autism showed greater deficit when the task was presented on a computer - Ozonoff, 1995). The differences between autistic and control participants often disappear when full-scale IQ or verbal IQ is controlled for. Individuals with Attention Deficit Hyperactivity Disorder, Schizophrenia, Obsessive Compulsive Disorder and Tourette syndrome perform similarly to autistic individuals on some tests of executive function. Therefore, studies

trying to differentiate disorders on the basis of performance on EF tasks have yielded mixed and inconsistent results (see Rajendran & Mitchell, 2007 for a review).

In addition to these problems, there is little evidence for a widespread EF deficit in preschool children with autism (Hill, 2004), although this might be because young children in general do not perform very well on EF tasks. However, Frith (2003) speculated that EF deficits in autism might emerge at a later point in development, as a result of poor pruning (i.e., the elimination of faulty connections between neurons). On the other hand, Ozonoff and Jensen, (1999), as well as Happé, Booth, Charlton and Hughes (2006) reported age-related improvements in EF in autism relative to typically developing control groups.

An interesting issue is the relationship between EF and ToM. Some researchers have argued that ToM is required for executive control (e.g., Perner, Lang & Kloo, 2002), whereas others suggested that theory of mind tasks could be reduced to executive processes (e.g., Russell, Mauthner, Sharpe, & Tidswell, 1991). According to Apperly, Samson and Humphreys (2009) both EF and language play an important role in the emergence of ToM, and EF continues to play an important role in ToM even in the case of adults. Pellicano (2007) in a group of young children found evidence that some children with autism failed ToM tasks, but performed well on EF tasks, while no children had the opposite profile. This supports the idea that EF is necessary (although not sufficient) for the emergence of ToM. On the other hand, it is still possible that ToM might play a role in the performance of certain EF tasks.

In sum, it seems that the EF hypothesis can explain many of the characteristics of autism. However, it has several limitations. Most importantly, not all individuals with autism show executive problems and those who do may have differing profiles of EF. Moreover, executive difficulties are not unique to autism. It is also hard to disentangle the complex relationships between EF, verbal abilities, and general intelligence.

#### *4.1.4. The extreme male brain theory of autism*

Baron-Cohen's (2002) extreme male brain (EMB) theory of autism proposes that WCC, and the deficits in ToM can be explained as the features of an extreme male brain with a strong drive for systemising, and an impaired ability for empathising. Empathising is the drive to understand the emotions and thoughts of others, and to respond to these with an appropriate emotion. This makes it possible to predict a person's behaviour, and to care about how they feel. Systemising is the drive to analyse and understand the variables and rules that govern the behaviour of a system, and the drive to construct systems (which can be mechanical, natural, abstract or other). Systemising makes it possible to predict the behaviour of systems, and to control them. The main claims of the EMB theory are the following. There are reliable sex differences in ToM and WCC, as well as between empathising and systemizing in the general population. Autistic individuals' performance on these tasks/characteristics is at the extreme male end of a continuum. Finally, both systemizing and empathising skills are dependent on a single biological factor (pre-natal testosterone secretion), which also implies a negative correlation between these characteristics.

Baron-Cohen and colleagues developed two questionnaires that measure empathising (EQ) and systemising (SQ) skills in the general population, and a third questionnaire, which measures both (Autism Spectrum Quotient, AQ - Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001). The idea is that "autistic traits" are present in the general population, and adults with high-functioning autism or Asperger Syndrome are just an extreme on that dimension. Baron-Cohen and colleagues claim that males tend to score higher on AQ than females, and high functioning autistic individuals tend to score higher than typical males. They also claim that scientists tend to score higher than non-scientists. On the other hand, AQ scores are not diagnostic in themselves.

The bulk of evidence for the EMB theory comes from the Cambridge fT Project. Baron-Cohen and colleagues followed the development of a sample of around 500 children

whose foetal testosterone (fT) levels were measured through amniocentesis. They found that fT levels were negatively correlated with the frequency of eye-contact in males at 12 months (Lutchmaya, Baron-Cohen & Raggatt, 2002a), fT was also negatively correlated with vocabulary development between 18 and 24 months (Lutchmaya, Baron-Cohen & Raggatt, 2002b). At 4 years fT was associated with poorer quality of social relationships and narrow interests, and at 8 years it was positively correlated with the child version of the Systemising Quotient. Between 6 and 10 years of age children's scores on the child version of AQ and on the Childhood Autism Spectrum Test (CAST, which is a diagnostic tool) were significantly correlated, although the correlation was not very strong ( $r=.25$ ). Boys scored higher on both tests than girls, although the effect size was very small for the CAST. fT levels were positively associated with scores on both tests when groups were combined and also within the sample of boys, but fT levels were not associated with CAST scores in the case of girls. There was no difference between genders on the block design test, and fT was unrelated to block design performance. Baron-Cohen, Wheelwright, Hill, Raste and Plumb (2001) also reported that in an adult sample women scored higher on the Eyes test, whereas men performed better on the embedded figures test.

On the other hand, the findings of researchers outside Baron-Cohen's lab are more mixed. Jarrold, Butler, Cottington, and Jimenez (2000) found a negative relationship between ToM and embedded figures performance in the case of university students. However, they did not find evidence for sex differences on these tasks. By contrast, Morgan, Maybery and Durkin (2003) reported that social and non-social autistic characteristics were independent in preschool children with autism. Carroll and Kin Yung (2006) conducted a big study with university students, and compared the performance of male and female science and non-science students. Some of their findings were in line with the EMB theory. For example, men had significantly higher AQ scores, and they scored lower on the Eyes test, and the Social Skills Inventory, whereas they scored higher on the block design test, and the Mechanical Reasoning test. Female humanities students

performed worse on the embedded figures test than the other groups, and science students scored higher on the block design test than humanities students. On the other hand, Mechanical reasoning and performance on the block design and embedded figures tests was unrelated to Social Skills. Mechanical reasoning was also unrelated to SQ. There was a significant relationship between sex and “brain type” (i.e., “male brain” with high systemizing and low empathising skills or “female brain”, which is the opposite), but there was no relationship between discipline or AQ and “brain type”.

In sum, there seem to be reliable sex differences on Baron-Cohen et al.’s measures of AQ, SQ and EQ, although the relationships between these measures are not quite clear (no evidence for strong negative correlations between AQ and EQ, and also not very strong correlations between AQ and the other two measures). In fact, there is recent genetic evidence that social skills/impairments, communication abilities, and restricted and repetitive behaviours are inherited through largely independent genes (Happé & Ronald, 2008). Additionally, the AQ measure is not correlated very strongly with diagnostic tests of autism (according to Baron-Cohen, Auyeung, Ashwin & Knickmeyer, 2009 it only accounts for approximately 20% of autistic traits).

Probably the most problematic aspect of the theory is that it mixes the question of natural sex differences with the issue of how the autistic cognitive profile differs from that of the general population. One reason that autism is associated with a male brain type is the fact that autism is more common amongst males than females. Children with autism have a sex ratio of 4:1 (males:females) across the full IQ range (Chakrabarti & Fombonne, 2005), and a ratio of 9:1 in the case of Asperger syndrome (Wing, 1981). That is, in the population that the AQ is aimed at approximately 90% of the individuals with autism are males. However, there are a number of other developmental disorders which predominantly affect males (just to give two examples, the male:female ratio is 5:1, and 10:1 in dyslexia and ADHD, respectively). The idea that sex differences emerge because of the same reason as differences between autistic and typically developing groups is also

problematic. For example, men generally perform better than women on the mental rotation task, and individuals with autism perform better than typically developing males (Falter, Plaisted & Davis, 2008). However, men are better than women on the rotational aspect of the task (rotation speed), whereas autistic individuals are superior on the non-rotational aspects (speed of mental comparison of objects). Similarly, poor performance on certain tasks can arise due to different reasons in different groups. For example, there are obvious and important differences between typically developing 3-year-olds who do not pass the false belief task, and older autistic children, who also fail. As Bloom and German (2000, p. B29) note “in all interesting regards, normal 3-year-olds are nothing like older children with autism”.

The measure of systemizing also seems to be seriously gender biased with representative items including: “If I were buying a car I would want to obtain specific information about its engine capacity.” “I find it difficult to read and understand maps.” “If there is a problem with the electrical wiring in my home I would be able to fix it myself.” The SQ questionnaire has also little to do with input-output operations, and it is hard to see why systemising should only be relevant to understanding non-social systems, rather than understanding any kind of system. As a recent psychometric analysis by Ling, Burton, Salt and Muncer (2009) shows, the scale is also better understood as consisting of four moderately correlated subscales (technicity, topography, DIY, and structure) than a measure of a unitary construct. Finally, the ecological validity of the questionnaire is dubious, as it does not correlate well with measures that tap into understanding systems, such as the Mechanical reasoning test (although it does correlate with performance on some tests where there is a known male advantage, such as the mental rotation test – when the effects of gender are not controlled for).

In conclusion, the EMB theory seems to be aimed at collecting evidence to support the claim that “autism can be described as a manifestation of extreme male characteristics” rather than at a better understanding of the cognitive profile of actual autistic people. This

is done through making circular claims, where cause and effect relationships (as well as the influence of both societal and biological factors) are hard to disentangle. Sadly, besides being devoid of practical significance, the theory even turned out to be dangerous in that recently Geier and Geier (2007) started to use testosterone-blockers for the “treatment” of children with autism, which is a form of chemical castration, usually used in the case of adult sex offenders.

#### *4.2. Reasoning and autism*

In the following section I will review the (sparse) existing research on reasoning in autism in the areas of analogical reasoning, memory retrieval and conditional reasoning, and some studies related to heuristics and biases in autism (two studies on counterfactual reasoning, and a study on the framing effect).

##### *4.2.1. Analogical reasoning in autism*

Analogy enables us to reason about novel phenomena, to identify relevant information on the basis of relational similarity, and to transfer knowledge from an initial learning context to future environments (e.g., Hummel & Holyoak, 2005; see Experiment 3 for a brief review on analogical reasoning). A small number of studies investigated analogical reasoning in autism. Based on some theoretical interpretations, it could be expected that individuals with autism would show impairments in this type of reasoning. For example, analogical reasoning is closely associated with the executive functions (EF) of the prefrontal cortex (Waltz et al., 1999), which are often impaired in autism (see Section 4.1.3.). However, such impairment is not universal, and autistic people exhibit varying performance profiles across different executive tasks. Some studies reported deficits on EF tasks in low- but not high-functioning autistic groups (e.g., Hughes, Russell & Robbins, 1994; Ozonoff et al., 2004). There is also some indication that EF deficits in autism as compared to typical controls decrease with age (Happé, Booth, Charlton &

Hughes, 2006; Ozonoff et al., 2004), although these findings could reflect sampling artefacts (as these studies were not longitudinal).

A deficit in analogical reasoning would also appear to be predicted by the WCC theory (see Section 4.1.2.). The WCC theory proposes that typically-developing individuals tend to engage in global processing, building up a gist-based representation, whereas autistic individuals engage in more detailed, local or piecemeal processing. Other evidence suggests that autistic people are impaired in generativity (i.e., using their knowledge spontaneously and flexibly in novel situations; Peterson & Bowler, 2000), and they appear to exhibit reduced processing of similarities between stimuli; instead, they are better at processing differences (Plaisted, 2001), although this characteristic may be restricted to perceptual processing. Autistic people also seem to have difficulty understanding metaphors and other types of non-literal language, such as irony (e.g., Pexman, 2008). Recently, Happé and Frith (2006) modified the WCC theory, and claimed that the finding of a local bias was robust. On the other hand, the evidence regarding weak global processing is mixed. Happé and Frith also concluded that the local bias was not a side effect of EF or ToM deficits, and that it could be overcome through conscious effort.

Given that relational integration involves a form of coherence, it would appear that WCC theory would predict that autism should lead to a general impairment of analogical reasoning. In fact, O'Loughlin and Thagard (2000), using a computational model in which coherence is established by constraint satisfaction, predicted (extrapolating from a computer simulation of performance on a homograph task) that the ability to reason analogically is very likely to be impaired in autism. However, the empirical evidence to date has been mixed. Scott and Baron-Cohen (1996) reported that autistic children with a learning disability were able to perform analogical reasoning tasks as well as both typically-developing children matched in mental age, and a group of children with learning disability who were matched in chronological and mental age. In contrast, Reed (1996) reported that autistic children showed poorer performance than a typically-developing group on two out of four analogical reasoning tasks. More recently, Dawson, Soulières, Gernsbacher, and Mottron (2007) examined autistic people's performance on



the Raven Progressive Matrices test (Raven, 1938), which is designed to measure the ability to form perceptual relations and to reason analogically independent of language and formal schooling. These investigators found that autistic people with a learning disability according to their Wechsler intelligence scores performed in the normal range on the Raven test.

The performance of autistic people on the Raven test suggests that their ability to reason relationally with complex, abstract materials is unimpaired. However, although the Raven test requires integration and processing of complex information (Carpenter, Just & Shell, 1990), it does not require activation and integration of relevant knowledge (as the problems are purely formal). In addition, the presence of distraction is not systematically varied in the Raven test. It thus remains possible that autistic people may have difficulty with analogical reasoning when they have to retrieve and integrate relevant knowledge, or when they have to resist interference (Diamond, 2006).

Another hypothesis is that autistic people solve Raven problems using a different strategy than controls. For example, increased ability to discriminate between and remember highly similar visual patterns is characteristic of autistic people, and has been claimed to underlie their superior performance on the Block Design task (Caron et al., 2006). This heightened discrimination ability could contribute to their success on the Raven test, which requires choosing the correct response from an array of visually similar patterns. Superior autistic performance in processing differences, coupled with reduced processing of similarities (e.g., Plaisted, 2001), might also contribute to the observed discrepancy between autistic people's WISC and Raven scores (e.g., Dawson et al., 2007).

#### *4.2.2. Memory retrieval and everyday conditional reasoning in autism*

Although the results of the analogical reasoning studies indicate that the ability to make links between explicitly presented, complex non-verbal information is intact in autism, it is less clear whether autistic people are also able to reason about things that have

to be inferred. One area of research that could be informative in this respect is the research on memory functioning in autism.

Traditionally children with autism are regarded as having good rote memory (e.g., Wing, 1981), but they seem to be impaired in the free recall of pictures and words relative to children matched on verbal and non-verbal ability (Boucher & Warrington, 1976), and even adults with Asperger syndrome show deficits in this regard (Bowler, Matthews & Gardiner, 1997). However, these studies also reported intact recall when participants were provided with acoustic, graphemic or semantic cues. In a series of studies Bowler and colleagues showed evidence of normal priming and recognition processes (that is, unimpaired implicit memory – e.g., Bowler, Gardiner & Grice, 2000), but an impairment in episodic memory requiring intentional recall (Bowler et al., 1997) and poor source memory (Bowler, Gardiner & Berthollier, 2004) in Asperger syndrome. Bowler et al. (2004) have proposed the “task support hypothesis” to account for these findings.

According to this hypothesis, autistic individuals show greater difficulty in retrieving and integrating their background knowledge with a problem context when the task provides little support for retrieval. Similarly, Minshew and Goldstein (2001) proposed that the memory impairment in autism originates from a failure to use organising strategies or meaning to support memory, although the underlying memory representation is intact.

Based on these considerations Happé and Frith (2006) proposed (using the terms of the fuzzy-trace theory – see Section 2.1.3.) that verbatim memory is intact in autism, whereas gist memory (which is likely to play a role in structuring the free recall of events) might be selectively impaired. In a study investigating eye-witness memory in autism (McCroy, Henry & Happé, 2007) children with Asperger syndrome mentioned less details of an event during free recall, and they were less likely to recall the most salient (“gist”) elements of the event than children in the control group. However, general and specific questioning elicited the same amount of new information in the autistic and in the control group, and in this case both groups recalled the most salient elements of the scenario. An

additional important finding of this study was that memory recall was correlated with executive functioning in the autistic, but not in the control group. A possible explanation for this is that children with autism did not benefit from a spontaneous organization of the material, and, consequently, they had to rely on their general executive resources during memory retrieval. This finding is in line with Bennetto, Pennington and Rogers (1996) who found correlations between memory and EF in autism, and also with Toichi and Kamio (2003) who reported an association between memory functioning and non-verbal reasoning in autism.

Everyday conditional reasoning (see Section 2.2.) relies heavily on memory retrieval. Namely, the logically valid MP inference (e.g., “if it’s raining, Tom gets wet”, “It’s raining, therefore...”) can be blocked by retrieving counterexamples, such as “Tom has an umbrella with him”. In this case, relying on real-life knowledge leads to worse logical performance. By contrast, in the case of the fallacious AC inference (e.g., “Tom gets wet, therefore...”) retrieving real-life knowledge about possible alternative antecedents (for example, “Tom has just had a shower”) helps blocking the invalid inference. Two recent studies looked at conditional reasoning in autism.

One of them (Pijnacker, Geurts, van Lambalgen, Kan et al., 2009) was conducted with high-functioning adults with autism, and of individually matched control participants (matched on age, gender, handedness, and verbal and non-verbal IQ). Two different tasks were used: traditional conditional reasoning problems (MP, MT, DA and AC), and “suppression task” versions of the same inferences. The suppression task consists of adding a further premise to a traditional conditional reasoning task, for example a MP inference with an additional premise would be: “If Mary has an exam, she will study in the library. If the library is open, Mary will study in the library. Mary has an exam. Will Mary study in the library?” Without the extra premise people usually endorse the conclusion. However, including the extra premise reduces endorsement rates to around 50% (Byrne, Espino & Santamaria, 1999). The results of this study showed that although there was no

difference between the autistic and control group on the traditional conditional inference problems, autistic participants were less influenced by the explicitly presented counterexamples in the suppression task. More specifically, they showed less suppression of MT and MP, although they showed equal suppression of DA and AC as the control group. Moreover, participants in both groups showed significantly lower endorsement rates in the suppression task than in the traditional conditional inference task. Overall these results give clear evidence that autistic participants were influenced by explicitly presented counterexamples, and their performance on the traditional tasks also suggest that they were equally influenced by retrieved counterexamples as the control group. According to Pijnacker et al. (2009) the reason that autistic participants showed less suppression of MP and MT (that is, they showed less suppression of valid inferences) is that they are less able than typically developing individuals to deal with exceptions, and it is a sign of their reduced mental flexibility. This argument is similar to Klaczynski's (2007) conclusion who claimed that children sometimes seem more rational than adults, because they follow rules more rigidly, whereas adolescents and adults integrate rule understanding with contextual considerations. Consequently, the decisions they make are more likely to deviate from the relevant normative rules.

The other recent study on conditional reasoning in autism (McKenzie, Evans & Handley, *in press*) was conducted with high-functioning adolescents with autism matched to a group of typically developing adolescents on age, verbal ability, verbal working memory and inhibition. In this study the availability of counterexamples was experimentally manipulated in order to examine the role of memory retrieval in participants' conditional reasoning performance (in this study only MP and AC inferences were included). In the case of typical samples the suppression of inferences is more likely when a high number of counterexamples are available, as opposed to when the availability of counterexamples is low (e.g., Quinn & Markovits, 1998). The results of this study showed that in the case of the MP inference typically developing adolescents were

influenced by the availability of counterexamples, whereas participants with autism showed no effect of availability. By contrast, in the case of AC inferences both groups were affected by the availability of counterexamples, although this effect was greater in the case of the typically developing group. The results of this study were somewhat similar to the Pijnacker et al. (2009) study in that the difference between the autistic and control groups was more pronounced in the case of valid inferences. On the other hand, in the McKenzie et al. (*in press*) study there was a bigger difference between the performance of the autistic and the control group. This possibly indicates that some of the differences between groups can be better understood as signs of a developmental delay in autism, although the difference on the suppression task in the case of valid inferences persisted into adulthood. Two other aspects of the McKenzie et al. (*in press*) study are interesting. Besides solving the conditional inference tasks, participants were also asked to rate the believability of the relations described in the premises. Moreover, they were asked to generate counterexamples that would render the conclusion of the conditional reasoning problems invalid. Believability ratings are based on implicit, associative processes and are known to have an influence on conditional reasoning, which is independent of the availability of counterexamples (e.g., Verschueren, Schaeken & d'Ydewalle, 2005). Neither believability ratings, nor the amount and quality of counterexamples generated by the two groups differed. This is in line with the claims of the task support hypothesis (Bowler et al., 2004) that implicit memory processes are unimpaired in autism, whereas explicit memory retrieval is deteriorated. However, autistic participants are able to retrieve the contents of explicit memory when they are provided with appropriate cues.

#### 4.2.3. *Heuristics and biases in autism*

Stanovich (2003) described the *fundamental computational bias* as the automatic and ubiquitous tendency to contextualize problems with prior knowledge. In addition to contextualisation, people also tend to "socialize" problems, and rely on pragmatic cues and

inferences even in impersonal situations; they are looking for deliberative design and pattern in randomly generated events; and they try and create a coherent narrative of the events of their lives (see also Section 1.1.1). Stanovich (2003) also claimed that these tendencies form the basis of the susceptibility to heuristics and biases. Based on the ToM deficit and the WCC theory (see Sections 4.1.1. – 4.1.2.) we could expect that individuals with autism will not experience the drive to contextualise and socialize problems, and thus we could expect them to be less susceptible to reasoning biases. An interesting relevant finding is related to the Social Attribution Task (Heider & Simmel, 1944) which is a silent animation of geometric objects interacting with each other. Although the figures are abstract, people tend to describe their behaviour in mentalistic terms. Klin (2000) asked a group of higher-functioning adolescents and adults with autism, and a neurotypical control group to describe the behaviour of the objects. Participants with autism tended to describe the animation in mainly geometric and physical terms, whereas the control group attributed social meaning to the movements of the geometric shapes. Thus, as could be expected based on the ToM deficit and WCC accounts, autistic participants did not exhibit the fundamental computational bias in the case of this task.

Another relevant area of research is the investigations into counterfactual reasoning in autism. Counterfactual reasoning is reasoning about what might have been if some past events had turned out differently (e.g., Byrne, 2007). This type of reasoning underlies certain reasoning biases, such as the if-only fallacy (see Chapters 2-3). Studies of typically developing children showed that counterfactual reasoning is closely related to false belief understanding, even when verbal mental age is controlled for (Riggs, Peterson, Robinson & Mitchell, 1998). Peterson and Bowler (2000) conducted a study with children with autism, children with severe learning difficulties (SLD), and a group of typically developing children who were matched to the other two groups on verbal mental age. They used false belief scenarios which were similar to the unexpected transfer test of false belief (see Section 4.1.1), and they asked participants both about the belief of the character (false

belief question), and the location of the object, *if the relocation hadn't taken place* (counterfactual question). Children with autism, but not children with SLD showed significantly poorer performance on the false belief task than the control group (the difference between the two clinical groups was not significant). The results also showed that for the clinical groups the counterfactual reasoning task was significantly easier than the false belief task, whereas the two tasks were equally difficult for the (much younger) control group. Verbal ability was significantly related to both counterfactual reasoning and ToM performance in all groups. Finally, verbal ability and counterfactual reasoning were significant predictors of ToM performance in each group. Peterson and Bowler (2000) claimed that counterfactual reasoning ability is a necessary, but not sufficient prerequisite of passing the false belief task. This is because the ToM task requires the spontaneous generation of counterfactual suppositions, whereas the counterfactual alternative is cued in the case of the counterfactual reasoning question.

In Peterson and Bowler's (2000) study only one type of counterfactual reasoning (i.e., subtractive reasoning) was investigated. A recent study on counterfactual reasoning in children with high functioning autism (Begeer, Terwogt, Lunenburg & Stegge, 2009) investigated the development of both additive ("If only I had done...") and subtractive ("If only I had not done...") counterfactuals in children between the age of 6 and 12. The autistic and control groups were matched on age, gender, verbal ability, and full-scale IQ. Besides the counterfactual reasoning tasks, children also had to engage in an ideational fluency task where they had to generate as many ideas within a certain category as possible (e.g., "What can you drink?"). Overall both groups generated a similar number of both additive and subtractive counterfactual statements. However, the developmental patterns in the two groups differed. In the autistic group, older children tended to generate more subtractive counterfactuals than younger children, whereas there was no age-related increase in the number of additive counterfactuals. By contrast, in the control group the number of additive counterfactuals increased with age, but the number of subtractive

counterfactuals remained stable. In addition, 6-8 year old autistic children generated fewer subtractive counterfactuals than the control group, whereas 10-12 year old autistic children generated fewer additive counterfactuals than controls. Autistic children showed poorer performance on the ideational fluency task than the control group, and performance on the ideational fluency task was significantly related to generating additive counterfactuals in both groups. Begeer et al. (2009) interpreted these findings as a sign of a tendency in the autistic group for “think within the box”. That is, a tendency to narrow down responses to the explicitly given information which results in a preference for generating subtractive, rather than additive counterfactuals. However, these patterns could also be interpreted as indicating a slight delay in the development of counterfactual reasoning in autism.

A final study that can be related to the heuristics and biases research programme is an investigation into the framing effect in autism (De Martino, Harrison, Knafo, Bird & Dolan, 2008). Participants in this experiment were presented with different scenarios where they had to choose between a sure option and a gamble option. Half of the trials were framed as “gain trials”, the other half were framed as “loss trials”. For example, in one scenario participants had received £50 and then they had to decide whether they wanted to keep £20 (gain scenario)/lose £30 (loss scenario) or they choose to gamble, where they had a 60% chance of losing all money and a 40% chance of keeping it all. Given the equivalence of the gamble and sure choices in terms of the expected utility, the measure of participants’ “rationality” was whether they were able to ignore the frame, and give the same response on equivalent tasks regardless of presentation (i.e. gain or loss frame). Participants had to make a decision within 4 seconds. De Martino and colleagues found a strong tendency for the control participants to exhibit a framing effect, which was markedly reduced in the autistic group. In addition to the behavioural differences, autistic participants exhibited no differential autonomic responses to contextual cues as indexed by skin conductance responses. De Martino et al. (2008) related these findings to Type 1 and



Type 2 processing, and they claimed that autistic participants have an increased tendency towards analytic decision making due to their impaired intuitive reasoning mechanisms.

#### *4.3. Discussion*

Historically, and in the interest of parsimony, cognitive theories of autism were aimed at identifying a single cause that could explain all the symptoms of autism. Four major cognitive theories of autism have been developed with this aim: the theory of mind deficit hypothesis, the weak central coherence theory, the executive function deficit theory and, most recently, the extreme male brain theory. Despite much research in all these areas, none of these accounts provide a complete picture of autism. Moreover, indices of autistic symptomatology have been found to be unrelated to performance in any cognitive domain, suggesting that cognitive capabilities and deficits are not associated with the behavioural symptoms of autism (Pellicano et al., 2006).

Recent evidence from twin-studies also indicates that largely independent genes may operate on social skills, communication abilities, and repetitive behaviours (Happé & Ronald, 2008). Due to the independent inheritance of the different behavioural symptoms of autism, as well as due to the developmental interactions between different symptoms which can lead to distinct profiles of strengths and weaknesses in each autistic individual, we can expect much heterogeneity within the autistic population. Autism is diagnosed on the basis of deficits in social communication, social interaction, and social imagination, thus impairments in the social domain are present in all individuals with autism, by definition. Many autistic individuals also suffer from executive functioning deficits, and they also tend to exhibit a particular cognitive style which results in a reduced level of contextualisation, although this can be overcome by conscious effort.

Investigations into analogical reasoning suggest that autistic individuals are able to process the relationships between explicitly presented, complex, non-verbal items. On the other hand, the evidence concerning verbal reasoning reveals some differences between

autistic and typical populations. Research in the area of memory retrieval suggests that autistic individuals have problems with the uncued retrieval of explicit memories, whereas they show unimpaired implicit and associative processes. Experiments regarding conditional inference show a reduced tendency in autism to contextualise problems with relevant background knowledge, even when counterexamples are explicitly presented. Similarly, counterfactual reasoning in autism tends to be more restricted than in typical development, with a decreased tendency to “think outside the box”, and to go beyond explicitly presented facts. However, it is also worth mentioning that in the case of high-functioning individuals with autism, and especially in the case of adults, these differences are subtle and participants with autism display similar patterns to control participants. Thus, the differences are often quantitative rather than qualitative.

From the point of view of dual-process theories, the performance of autistic participants sometimes seems more logical or normative than that of the control group. For example, in the studies on conditional reasoning (McKenzie et al., *in press*; Pijnacker et al., 2009), as well as in a study on the framing effect (De Martino et al., 2008) autistic participants were less influenced by contextual information, and thus, in effect, they were more “logical” than the control group. In dual-process terms, these findings can be interpreted as the results of missing or insufficient Type 1 input. However, certain Type 1 processes, such as implicit associations seem to be intact. On the other hand, Type 2 processing can also be impaired in the case of individuals with autism who suffer from EF deficits. However, there is little evidence for a general deficit in working memory or inhibition skills in autism (see Hill, 2004), and EF deficits are usually absent or mild in the case of high-functioning individuals with autism. Consequently, we can expect that high-functioning autistic individuals will show increased logical performance on some heuristics and biases tasks.

Some dual-process theorists (e.g., Evans & Over, 1996) suggest that heuristic processing leads to correct responses most of the time. Consequently, if individuals with

autism rely less on heuristic processing, they will be disadvantaged in many real-life situations where typically developing individuals are able to make decisions quickly and efficiently. From a developmental perspective, it seems that, at least in the case of some reasoning biases, individuals with autism develop the same kind of biases as typically developing individuals, but they achieve the same level of development more slowly, and sometimes their development does not reach the level of functioning of the mature typical system. It is also possible that although some autistic participants show typical performance on certain tasks, the underlying cognitive processes are different. In Chapter 5 I am going to describe two experiments that investigated the development of reasoning heuristics in autism.

## Chapter 5. Heuristic reasoning in autism (Experiments 4 and 5).

### *5.0. Introduction to Chapter 5*

In this Chapter we are going to describe two studies that were aimed at investigating the development of reasoning heuristics in autism. In these studies we administered the same reasoning problems and cognitive ability/executive function measures as in Experiments 2 and 3, and a subset of the children who took part in those studies were involved in the present experiments as a control group. The experimental findings with typically developing children (see Chapter 3) suggested that although in some respects children's normative competence increases with age, this is not always evident in their actual performance on reasoning tasks, especially in the case of typical heuristics and biases tasks where heuristic and normative responses are always in conflict with each other. In fact, in the case of young children (between the age of 5 and 11) giving responses that are traditionally considered to be heuristic (and also non-normative) was positively related to children's cognitive ability. At this stage reasoning performance on the different tasks was also moderately related across tasks. In the case of adolescents, performance across tasks was mostly unrelated, and the relationship between reasoning performance and cognitive capacity also showed very mixed patterns. This can be interpreted as a sign that at this stage of development some heuristic responses were produced relatively quickly and automatically, whereas others were still dependent on general cognitive resources. Based on these findings, we concluded that using reasoning heuristics is a developmental achievement which is closely related to the development of general cognitive resources, and possibly to the development of executive functions as well.

In Chapter 4, we gave an overview of the reasons why it can be expected that the development of heuristics will be atypical in autism. Based on the prominent cognitive theories of autism, two conflicting predictions can be made. The weak central coherence theory (Frith & Happé, 1994; Happé, 1999) proposes that individuals with autism engage

in a piecemeal, detail-focussed processing, and they are less inclined to extract the gist of stimuli that they are presented with than typically developing individuals. As heuristic responding is considered to be based on the perceived contextual links between the representation of a task and a particular response option, we can expect that autistic children will be less sensitive to these contextual cues, and will consequently produce fewer heuristic responses. Alternatively, it is possible that autistic individuals achieve global coherence, but this does not happen spontaneously. Instead contextualisation will be optional, and when it happens, it will require conscious effort (Happé & Frith, 2006).

We can also expect autistic children to be less susceptible to reasoning heuristics and biases on the basis of the extreme male brain theory (Baron-Cohen, 2002). This theory proposes that autistic cognition is characterized by a strong drive for systemizing. Systemizing makes it possible to analyse and understand the variables and rules that govern the behaviour of a system, and it also involves the drive to construct systems (which can be mechanical, natural, abstract or other). On the basis of this description we could expect that individuals with autism will be good at extracting the logical structure of tasks, and as a result, will be more likely to produce normatively correct responses.

On the other hand, based on the executive function deficit theory (see e.g., Hill, 2004 for a review) it could be hypothesized that individuals with autism have impaired Type 2 processing abilities. That is, they will be more susceptible to biases than typically developing children. However, as we have seen in Chapter 3, in the case of children Type 2 processing can lead to responses that are traditionally considered to be heuristic. Thus, on the basis of these findings we can expect that a Type 2 deficit could result in less bias in the case of individuals with autism. Taken together the claims of all these accounts, the most likely result in the case of children is decreased heuristic processing in autism.

Looking at the results of studies that have investigated the reasoning abilities of individuals with autism (see Section 4.2.), the findings are mixed. Research into analogical reasoning suggests that autistic individuals are able to process the relationships between

explicitly presented, complex, non-verbal items (i.e., they perform well on the Raven test – see Dawson et al., 2007). On the other hand, the evidence concerning verbal reasoning reveals some differences between autistic and typical populations. Research in the area of memory retrieval suggests that autistic individuals have problems with the uncued retrieval of explicit memories, whereas they show unimpaired implicit and associative processes. Experiments on conditional inference show a reduced tendency in autism to contextualise problems with relevant background knowledge, even when counterexamples are explicitly presented. Similarly, counterfactual reasoning in autism tends to be more restricted than in typical development, with a decreased tendency to “think outside the box”, and to go beyond explicitly presented facts. However, it is also worth mentioning that in the case of high-functioning individuals with autism, and especially in the case of adults, these differences are subtle and participants with autism display similar patterns to control participants. Thus, the differences are often quantitative rather than qualitative. In any case, the above findings support the idea that there is a reduced tendency in autism to contextualize complex verbal stimuli (see also Lopez & Leekam, 2003). In line with this assumption, in the studies on conditional reasoning (McKenzie et al., *in press*; Pijnacker et al., 2009), as well as in a study on the framing effect (De Martino et al., 2008) autistic participants were less influenced by contextual information, and as a result they were more “logical” than the control group. In developmental studies (McKenzie et al., *in press*) autistic individuals also seemed to show a developmental delay in displaying some typical biases.

In summary, it seems that individuals with autism develop the same kind of biases as typically developing individuals, but they achieve the same level of development more slowly, and sometimes their development does not reach the level of functioning of the mature typical system. As a result, they sometimes seem more normative or logical than typically developing individuals, as they display less biases. However, these differences between groups can be very subtle or even absent. Nevertheless, it is also possible that

although autistic participants show similar performance to the control group on certain tasks, the two groups rely on different cognitive processes whilst they solve the tasks (for example, they contextualise verbal materials through conscious effort).

*5.1. Experiment 4: The development of heuristic reasoning in autism and its relationship to cognitive abilities and executive functioning.*

The main aim of this study is to compare the tendency to rely on heuristic reasoning in autism and in typical development. Based on the theoretical approaches that we reviewed in the introduction, the most probable outcome is that children with autism will be less susceptible to heuristics than typically developing children (see Table 5.1. for a summary). Another aim is to explore the links between potential differences between groups both in their reasoning performance and their cognitive abilities/executive function (EF) skills. Finally, given that a prominent cognitive theory of autism is based on the assumption that EF deficits are widespread in autism, it can be informative to compare the EF skills of the two groups. Although EF components are usually investigated through extracting latent variables from a number of different tasks, using a large group of participants, many studies that have investigated EF-deficits in autism used small groups and they also employed tasks that required the simultaneous use of multiple EF-skills. In the present study we used tasks that measure working memory capacity, inhibition and set-shifting ability, which were identified as the main aspects of EF (Miyake et al, 1999). The fact that we use tasks that tap into all main components of EF, and that each of these tasks measure a single EF-component (rather than a combination of different skills) makes this aspect of the study interesting in itself.

Investigating EF skills and reasoning performance within the same group of participants can also be informative from another point of view. Frith (2003) speculated that EF deficits in autism might be the result of poor pruning (i.e., the elimination of faulty connections between neurons), which she also connected with weak central coherence in

autism. Consequently, EF-deficits could be related to a decontextualised thinking style (that is, less use of heuristics in reasoning tasks). There is a wave of synapse formation in the frontal cortex just before puberty (age 11 in girls, 12 in boys) which is followed by a pruning back in adolescence (Giedd, 2004). The differences between autistic and typically developing children should emerge at the time when pruning takes place, that is, after the age of 12. The idea that EF deficits should be related to a decontextualised thinking style goes against the claims of dual-process theorists, but it fits well with the concept that general cognitive resources are necessary for the development of reasoning heuristics.

Table 5.1. *Expectations about heuristic use in autism based on some prominent theoretical approaches.*

<b>Theoretical approach</b>	<b>Expectations about heuristic use</b>	<b>Mechanism</b>
Weak central coherence theory	Less heuristic use	Piecemeal processing style, reduced sensitivity to context
Extreme male brain theory	Less heuristic use	Piecemeal processing style, increased sensitivity to structure and factual information
Executive dysfunction theory	More/less heuristic use	Executive function deficits (reduced ability to resist interference / piecemeal processing style)
Task support hypothesis	Less heuristic use	Problems with uncued retrieval of contextually related concepts

In summary, the present study is the first one to examine the performance of autistic children on a range of classic heuristics and biases problems. Based on some prominent cognitive theories of autism, we expect that if there are any differences between groups, it is likely that autistic children will produce fewer heuristic responses to typical heuristics and biases tasks than the control group. This prediction is however complicated by the fact that once children are able to produce heuristic responses automatically, they might suppress these responses, if they are aware of the conflict between these heuristics and a relevant normative rule (see Stanovich & West, 2008). As a result, in the case of these tasks there might be no difference in the responses given by the two groups, or EF-deficits



can lead to *more* heuristic responses. Additionally, we can expect either no difference in the EF skills of the two groups, or some impairment in the autistic group. Alternatively, there might be no difference between heuristic responding across groups, because children with autism engage in an effortful contextualisation process (cf. Happé & Frith, 2006).

Thus, although it is possible that children with autism will produce the same number of heuristic responses as the control group, in these cases we expect that the relationships between heuristic responses and the measures of cognitive capacity and executive functioning will differ between groups. Namely, it is likely to be the result of either autistic children giving heuristic responses based on an effortful contextualisation process, or the lack of difference between groups is the result of typically developing children's effortfully suppressing heuristic responses. In the first case, there should be a positive correlation between cognitive capacity and heuristic responding in the autistic group only. In the latter case, there should be a negative correlation between heuristic responding and measures of executive functioning (especially inhibition) in the typically developing group only.

## Method

### *Participants*

Twenty-five high functioning children with autism (1 girl) took part in the study (mean age 14 years 4 months). Diagnostic records of the children showed that every child had received a diagnosis of autism by experienced clinicians. No child had a diagnosis of Asperger's syndrome or Pervasive Developmental Disorder. Additionally, 41 typically developing children (18 girls; mean age 13 years 10 months) participated in the study as a control group. The typically developing participants involved in this study were the same children as in Experiment 2, and the individual differences measures used in both studies were also identical (consisting of the Vocabulary and Block Design subtests of the WISC, the counting span task, the stop-signal task, and the colour-shape set-shifting task).

Table 5.2. displays the descriptive statistics for the measures of cognitive capacity and executive functioning across groups. We failed to collect data from 2 participants on the working memory task, 1 participant on the block design task, 3 participants on the inhibition task, and 2 participants on the set-shifting task due to software errors or the experimenter's failure. Autistic participants scored significantly lower on the working memory and vocabulary tasks. There was no difference between groups in terms of age, block design scores, inhibition or set-shifting. We also compared the two groups on the combined WISC score (the combined z scores of the two subtests), and there was no difference between groups on this measure either.

Table 5.2. *Comparisons between the autistic and the control group on the measures of cognitive abilities and executive functions.*

		M (SD)	t	df	Sig. (2-tailed)
Age (in months)	autistic	171.84 (3.54)	1.50	72	.14
	control	165.78 (1.96)			
Working memory	autistic	5.32 (6.1)	2.12	70	.04
	control	9.45 (8.65)			
Vocabulary	autistic	31.96 (9.66)	2.12	72	.04
	control	36.51 (6.52)			
Block design	autistic	48.76 (10.42)	.40	71	.69
	control	49.75 (9.67)			
WISC	autistic	.00 (1.47)	.06	71	.95
	control	-.02 (1.63)			
Inhibition	autistic	42.39 (12.13)	.14	69	.89
	control	42.02 (10.09)			
Set-shifting	autistic	.80 (.14)	.09	70	.93
	control	.79 (.14)			

### *Materials and procedure*

The same reasoning tasks as in Experiments 1 and 2 were used (see Appendix A), and the procedure was also the same. Participants took part in two sessions. Session 1 took about 15 minutes, and participants solved the reasoning tasks. Participants were tested in small groups. After listening to the instructions given by the experimenter they worked

through the tasks individually. Session 2 consisted of the counting span task, the vocabulary and block design subtests of the WISC-III, the stop-signal task, and the set-shifting task. This session took about 40 minutes, and children did the tasks on a computer (apart from the WISC subtests), supported by the experimenter.

## Results

ANOVAs were conducted on the number of heuristic, analytic and “other” responses (see Table 5.3.) in order to analyze the effects of participant group on each problem (supplementary tables can be found in Appendix S4). As there was a significant difference between groups in their working memory and vocabulary scores, these were included as covariates in the analyses. However, this did not significantly change the results (apart from in the case of the conjunction fallacy task, see below). Consequently, we report the results of the simple ANOVA analyses. It is also worth noting that controlling for differences in verbal ability is considered to be dubious by some theorists, as impairments in verbal ability are a fundamental aspect of autism. Thus, controlling for differences in verbal ability possibly results in indirectly “controlling for autism” (see Bishop, 1997).

On the syllogistic reasoning task the ANOVAs indicated a main effect of participant group in the case of the heuristic responses ( $F(1,73)=6.58, p<.05, \eta_p^2=.09$ ). This was because autistic participants gave more belief-based responses than the control group. No other effects or interactions reached significance ( $F_s<4, p_s>.05$ ).

In the case of the sunk cost fallacy, the main effect of participant group on normative responses was significant ( $F(1,73)=4.89, p<.05, \eta_p^2=.07$ ). This indicated that autistic children gave fewer normative responses than children in the control group. There was also a main effect of participant group in the case of “other” responses ( $F(1,73)=18.13, p<.001, \eta_p^2=.21$ ), indicating that autistic children gave significantly more “other” responses to the tasks than typically developing children, which corresponded to a reversed sunk cost effect

(a tendency to abandon an unpleasant situation more readily if previous investment has been made).

There was no effect of group on any type of response in the if-only fallacy problems.

In the case of the conjunction fallacy, there was a main effect of participant group on heuristic/normative responses ( $F(1,73)=5.28, p<.05, \eta_p^2=.07$ ). This indicated that autistic children committed the fallacy significantly less often than the control group. However, when we included vocabulary and verbal working memory as covariates, this effect disappeared ( $F(1,73)=2.16, p=.15$ ) and there was only a significant effect of vocabulary scores ( $F(1,73)=4.65, p<.05, \eta_p^2=.07$ ).

Table 5.3. *Proportion of heuristic, normative and "other" responses on the different types of reasoning and judgment tasks across participant groups (SDs in parentheses).*

		autistic	control
Syllogistic reasoning	Heuristic responses	.59 (.25)	.43 (.25)
	Normative responses	.27 (.22)	.37 (.20)
	Other responses	.14 (.13)	.20 (.18)
Sunk cost fallacy	Heuristic responses	.20 (.29)	.22 (.29)
	Normative responses	.44 (.36)	.63 (.34)
	Other responses	.36 (.37)	.08 (.19)
If-only fallacy	Heuristic responses	.44 (.36)	.54 (.38)
	Normative responses	.50 (.35)	.43 (.35)
	Other responses	.06 (.17)	.03 (.12)
Conjunction fallacy	Heuristic responses	.86 (.23)	.96 (.14)
	Normative responses	.14 (.23)	.04 (.14)

#### *Relationships between the individual differences measures*

The next set of analyses examined the relationship between age (in months), cognitive ability (as indicated by the counting-span task, and the short form of the WISC-III), and executive functioning (as measured by the stop-signal task, and the set-shifting task). Results are displayed in Table 5.4.

Table 5.4. *Correlations between the cognitive ability measures for the autistic sample (the results for the typically developing sample in Experiment 2 are displayed in brackets).*

	1.	2.	3.	4.	5.
1. age in months	--				
2. WISC	.20 (.23)	--			
3. working memory	.00 (.37*)	.29 (.49**)	--		
4. inhibition	.25 (.24)	-.08 (.25)	.26 (.15)	--	
5. set-shifting cost	-.02 (-.19)	.12 (-.20)	-.41* (-.25)	-.34 (-.16)	--

In the typically developing group (see Experiment 2) there was a significant correlation between age and working memory ( $r(45)=.37, p<.05$ ), and WISC scores and working memory ( $r(44)=.49, p<.01$ ). In the case of the autistic group, although there were some medium-sized correlations, because of the small sample size ( $n=25$ ) these correlations were mostly statistically unreliable. The only significant relationship was a negative correlation between set-shifting cost and working memory capacity. That is, the higher children scored on the verbal working memory task, the better they were at set-shifting. Miyake, Emerson, Padilla and Ahn (2004) reported that articulatory suppression increased the switch cost on the colour-shape task in an adult sample. They suggested that inner speech might be required for retrieving and activating relevant task goals. Thus, children with autism might show impairments on the set-shifting task (and might have problems with flexibly switching between representations in general) because they have verbal working memory deficits. In fact, many previous studies found a relationship between memory functioning and EF in autism, but not in typical development (e.g., Bennetto et al., 1996; McCroy et al., 2007). However, although we found a relationship between set-shifting and verbal working memory in the autistic sample, and autistic

children had significantly lower working memory scores than the control group, there was no difference between groups in their set-shifting ability.

*Relationships between the individual differences measures and heuristic responding*

As a next step, the correlations between the individual differences measures and heuristic responding on the reasoning and decision-making tasks were analysed (the results are displayed in Table 5.5.). In the case of the autistic sample we did not combine the WISC and working memory scores as we did in the case of typically developing children, since the two scores were not significantly correlated (although we did find a medium-size correlation between the two measures). Heuristic responding was not reliably related to age. However, there was a significant positive correlation between WISC scores and committing the if-only and the conjunction fallacy. These correlations remained significant even after controlling for the effects of age:  $r(20)=.55, p<.01$ , in the case of the if-only fallacy, and  $r(20)=.45, p<.05$ , in the case of the conjunction fallacy.

Finally, we analysed the relationships between heuristic/analytic responding on the different tasks in the autistic group (see Appendix S4, Tables 2 and 3). As in the case of the typically developing children, there was no relationship between giving heuristic/analytic responses to the different tasks.

Table 5.5. *Correlations between the individual differences measures and heuristic responding on the different tasks in the autistic group (results for the typically developing sample from Experiment 2 are displayed in brackets).*

	age	WISC scores	Working memory	Inhibition	Set-shifting cost
Belief bias	.23 (-.19)	.17 (-.18)	.02 (-.24)	.15 (-.25 <sup>†</sup> )	.13 (.07)
Sunk cost fallacy	.10 (-.06)	.13 (-.11)	-.01 (-.18)	.05 (-.32*)	.27 (.08)
If-only bias	.29 (.10)	.57** (.22)	.27 (.12)	.01 (.25 <sup>†</sup> )	-.23 (.08)

---

Conjunction fallacy	.06 (.29*)	.45* (.25)	.23 (.22)	-.02 (.15)	.10 (.20)
---------------------	------------	------------	-----------	------------	-----------

---

## Discussion

This study was aimed at investigating the differences between the development of reasoning heuristics in autism and in typical development, as well as the differences between groups in their general cognitive resources and EF skills. Our participants with autism showed no general impairment in their cognitive capacity and executive functioning. The two groups performed at the same level on the short form of the WISC, on the inhibition measure and on the set-shifting task. However, participants with autism scored significantly lower on the working memory measure and on the vocabulary task. Certain theoretical accounts (e.g., Baddeley, Gathercole & Papagno, 1998) propose that there is a direct causal link from poor verbal working memory and delayed vocabulary acquisition. However, poor vocabulary can also contribute to poor working memory performance (Hulme & Roodenrys, 1995). Verbal working memory is also involved in flexible switching between mental representations (Miyake et al., 2004), and we found evidence for a relationship between set-shifting ability and working memory scores in our sample. Nevertheless, we found no evidence for impairments in mental flexibility (as indexed by set-shifting performance) in the autistic group.

To the best of our knowledge this is the first experiment which has investigated heuristic reasoning in autistic people. Based on some prominent cognitive accounts of autism (see Table 5.1. for a summary), we expected that autistic children will produce less heuristic responses to typical heuristics and biases tasks than the control group, at least on a certain subset of tasks. Namely, we expected this in the case of tasks where there was no evidence in the control group that children suppress heuristic responses (i.e., the if-only fallacy and the conjunction fallacy –see Experiment 2). Alternatively, we expected autistic children to produce the same number of heuristic responses on these tasks as the control

group, but to do this through conscious effort. In fact, we found that although there was no clear evidence for less bias in the autistic sample (the difference between groups in the case of the conjunction fallacy task was the result of differences in their verbal abilities). However, the correlational patterns indicated that autistic children produced heuristic responses on these tasks relying on effortful (Type 2) processing. By contrast, typically developing children seemed to produce these responses effortlessly, although the non-significant correlations between cognitive capacity/EFs and heuristic responding on these tasks were positive. There was also a significant positive correlation between cognitive capacity and heuristic responding on the conjunction fallacy task in the typically developing group, but it disappeared after controlling for age.

There are also some heuristic responses that children can produce with little effort, but they tend to suppress them, because they are aware of the conflict between these heuristics and a relevant normative rule (see Stanovich & West, 2008). In the case of these tasks it can be expected that EF-deficits should lead to *more* heuristic responses in the case of the autistic group. Based on the results of Experiment 2, this was to be expected in the case of the belief bias and the sunk cost fallacy. In fact, participants with autism were more susceptible to the belief bias than the control group. Although there was no difference between groups in committing the sunk cost fallacy, typically developing children gave more normative responses to these tasks. The fact that autistic children did not suppress these heuristics as typically developing children did, might be the result of them not being able to produce these heuristics automatically. Alternatively, in the case of the syllogistic reasoning task, typically developing children's advantage could be the result of their benefitting from the context the tasks were presented in. When investigating the performance of young children, children with autism and children with a learning disability on syllogistic reasoning problems, Leever and Harris (1999) found that in general children benefitted from presenting the problems in a fantasy context. Specifically, providing a fantasy context made it possible for typically developing children, and children with a



learning disability to accept false premises as the basis of their reasoning, and to avoid belief bias (i.e., simply saying that believable conclusions followed from the premises, and unbelievable conclusions did not). By contrast, autistic children did not benefit from the fantasy context, and they performed at chance level.

Going back to our results, just as in the typically developing group, there was no relationship between heuristic/normative responding across tasks. This is in conflict with the predictions of dual-process theories, but supports the idea that heuristic reasoning develops through different stages. There is an initial effortful stage, followed by a more automatic stage, which is (sometimes) followed by a stage where heuristics are consciously suppressed. The exact timing of these stages can differ across different tasks, depending on the complexity of the contextualisation process. In the case of the autistic sample it seemed that belief-based responses on the syllogistic reasoning task, and sunk cost responses were produced relatively effortlessly, whereas if-only responses and committing the conjunction fallacy required conscious effort.

### *5.2. Experiment 5: Susceptibility to heuristics, and analogical reasoning ability in autism.*

Experiment 5 was a partial replication of Experiment 3. This study was conducted with a different group of autistic children (except for 7 children who also took part in Experiment 4). A subset of children from Experiment 3 was included as a control group. As we described in Chapter 3, each task was presented in two versions: one in which there was a conflict between logic/a normative rule and intuitions (conflict problems), and in another version where logical/normative considerations were in line with intuitions (non-conflict problems). In the case of non-conflict tasks it was possible to give a normative response without relying on a contextualization process, and also without the need to suppress a tendency to contextualize.

Experiment 4 demonstrated that children with autism were less biased than the control group in the case of some problems, and more biased in the case of others. We

interpreted this variability across tasks as the sign of differences in the nature of the contextualization process that was required to solve each task (that is, we assumed that the complexity and capacity demand of the contextualization process was not uniform across tasks). This interpretation was supported by the findings of Experiment 1, where the contextualization process required conscious effort in the case of young children, and even in the case of adolescents (in Experiments 2 and 3), heuristic responding on some tasks was positively related to measures of analogical reasoning. Nevertheless, if the differences between groups only affect contextualized reasoning, it could be expected that there will be no difference between groups in the case of tasks where a normative conclusion can be reached without contextualized reasoning (at least we could expect autistic children to reason about these problems at the level of their general cognitive functioning). That is, in the present experiment we can expect that the performance of the two groups will only differ on conflict problems, if there is a group difference in contextualized reasoning only. However, if there is a difference between groups in their reasoning ability at a more general level, then we can expect group differences on both conflict and non-conflict problems.

A further manipulation that we used in this study was related to the content of the problems. Each type of task was presented in two forms: first as involving objects or animals (non-social content), and second as involving people (social content). See Figure 5.1. for an illustration of the content and conflict manipulations. Note that the conflict and non-conflict version of the same problem (e.g., the “Brian problem” or the “tiger problem”) was administered to different groups of participants. That is, there were 8 conjunction fallacy problems altogether, and each participant had to solve four problems, which corresponded to one of each type. (Although the tasks used in Experiment 3 were the same, the effect of the content manipulation was not analysed there). It is well documented that autism often results in profound difficulties with everyday social interaction. It has been claimed that such deficits arise because autistic people have very

specific problems with reasoning within the social domain (see, for example, Adolphs, 1999). Whilst recent evidence suggests that stereotypes are activated and used as readily for autistic as typical populations (Hirschfeld, Bartmess, White & Frith, 2007), it is possible that participants with autism show impairments in social reasoning in the case of certain problems. Thus, it could be argued that participants with autism reason differently about a particular problem not because of differences in the way they contextualise information, but because these problems require reasoning about the social world.

<b><i>Conjunction fallacy, social, conflict</i></b>	<b><i>Conjunction fallacy, non-social, conflict</i></b>
<p>Brian has a studio, where he works alone. He is a very creative man, and he likes to experiment with colours. He takes his work to exhibitions, and sells some of them too. Mark the following statements with number 1 to 4 according to how likely they are.</p> <ul style="list-style-type: none"> <li>○ Brian is an aerobics instructor. (non-representative 1)</li> <li>○ Brian is a painter. (representative)</li> <li>○ Brian is a painter and an accountant. (representative + non-representative 2)</li> <li>○ Brian is an accountant (non-representative 2)</li> </ul>	<p>This animal lives in the jungles in Asia. It is quite big and strong, and has black stripes on its body. It feeds on smaller animals, and sometimes it also attacks people. Mark the following statements with number 1 to 4 according to how likely they are.</p> <ul style="list-style-type: none"> <li>○ This animal lives in large groups. (non-representative 1)</li> <li>○ This animal is a big cat. (representative)</li> <li>○ This animal is a big cat and it likes to eat small birds. (representative + non-representative 2)</li> <li>○ This animal likes to eat small birds. (non-representative 2)</li> </ul>
<b><i>Conjunction fallacy, social, non-conflict</i></b>	<b><i>Conjunction fallacy, non-social, non-conflict</i></b>
<p>Brian has a studio, where he works alone. He is a very creative man, and he likes to experiment with colours. He takes his work to exhibitions, and sells some of them too. Mark the following statements with number 1 to 4 according to how likely they are.</p> <ul style="list-style-type: none"> <li>○ Brian is an aerobics instructor. (non-representative 1)</li> <li>○ Brian is a painter. (representative)</li> <li>○ Brian is an aerobics instructor and an accountant. (non-representative 1 + non-representative 2)</li> <li>○ Brian is an accountant (non-representative 2)</li> </ul>	<p>This animal lives in the jungles in Asia. It is quite big and strong, and has black stripes on its body. It feeds on smaller animals, and sometimes it also attacks people. Mark the following statements with number 1 to 4 according to how likely they are.</p> <ul style="list-style-type: none"> <li>○ This animal lives in large groups. (non-representative 1)</li> <li>○ This animal is a big cat. (representative)</li> <li>○ This animal lives in large groups and it likes to eat small birds. (non-representative 1 + non-representative 2)</li> <li>○ This animal likes to eat small birds. (non-representative 2)</li> </ul>

Figure 5.1. An illustration of the content and conflict manipulations used in Experiment 5.

Another reason why this manipulation might be interesting is that there is a known content effect in the case of a classic heuristics and biases task. As we described in Chapter 1, when thinking about sequences of random events, people tend to give

responses corresponding to the *gambler's fallacy* (or *negative recency*) when the problem is about objects, whereas they are more likely to commit the *hot hand fallacy* (i.e., *negative recency*) when the problems are about people (Ayton & Fischer, 2004; Gilovich et al., 1985). Most of the other classic heuristics and biases tasks, however, are generally presented with a social content. Thus, it is not clear whether content effects exist in the case of other tasks as well (i.e., whether participants give different responses to otherwise logically equivalent problems depending on the content of the problems).

An additional topic that was investigated in this study was analogical reasoning in autism. Analogical reasoning requires comparing a source to a target analogue to identify systematic relational correspondences (Gentner, 1983), enabling a transfer of knowledge to novel environments (see Experiment 3 for a review on analogical reasoning, and Section 4.2.1. for a review of analogical reasoning in autism). On a theoretical basis we could hypothesize that analogical reasoning is impaired in autism. For example, a deficit in analogical reasoning would be predicted by the executive function deficit account of autism, as analogical reasoning is closely associated with the executive functions of the prefrontal cortex (Waltz et al., 1999). Given that relational integration involves a form of coherence, a deficit in analogical reasoning would also appear to be predicted by the weak central coherence theory (Frith & Happé, 1994; Happé, 1999). However, the empirical evidence to date has been mixed. Scott and Baron-Cohen (1996) reported that autistic children with a learning disability were able to perform analogical reasoning tasks as well as both typically-developing children matched in mental age, and a group of children with learning disability who were matched in chronological and mental age. By contrast, Reed (1996) reported that autistic children showed poorer performance than a typically-developing group on two out of four analogical reasoning tasks. More recently, Dawson, Soulières, Gernsbacher, and Mottron (2007) examined autistic people's performance on the Raven Progressive Matrices test (Raven, 1938). They found that autistic people with a learning disability according to their Wechsler intelligence scores performed in the normal

range on the Raven test. This suggests that autistic people are able to reason relationally with complex, abstract materials. However, although the Raven test requires integration and processing of complex information (Carpenter, Just & Shell, 1990), it does not require the activation and integration of relevant knowledge. In addition, the presence of distraction is not systematically varied in the Raven test. It thus remains possible that autistic people may have difficulty with analogical reasoning when they have to retrieve and integrate relevant knowledge, or when they have to resist interference. It is also possible that autistic people solve Raven problems using a different strategy than controls. The Raven test requires choosing the correct response from an array of visually similar patterns. Autistic people might rely on their increased ability to discriminate between and remember highly similar visual patterns (e.g., Caron et al., 2006) when they solve Raven problems. In the present study we also investigated the relative difficulty of the items involved in the analogical reasoning tasks across groups. If the autistic group relies on different abilities to solve analogical reasoning problems, then the item analyses should reveal this.

In summary, this experiment investigated the following questions. In order to examine further why autistic children perform differently from control children on some classic heuristics and biases tasks, we included some control tasks where heuristic and normative responses were in line with each other. The effect of task content (social/non-social) on children's performance was also investigated in order to decide whether the differences in reasoning performance are generally present, or they are restricted to reasoning about the social domain. Children also solved different types of analogical reasoning problems (which were aimed at measuring their ability to reason analogically about both abstract and thematic materials). Finally, the relationships between reasoning performance on the heuristics and biases and analogical reasoning tasks were also investigated. According to Kahneman and Frederick's (2004) notion of attribute substitution, heuristic reasoning is based on making links between situations that are

similar in some respect, and transferring knowledge from one context to another.

Analogical reasoning has been described in very similar terms (see e.g. Gentner, 1983). On the other hand, Stanovich (e.g., 1999) emphasizes the need to decontextualise and decouple representations while solving heuristic reasoning problems. Analogical reasoning also requires the decoupling of representations, and based on this view, we could expect it to be related to normative (rather than heuristic) responding. As we have seen in Experiment 3, there was in fact some evidence for a relationship between heuristic and analogical reasoning in the case of typically developing children, and the relationship was positive (which is in line with the attribute substitution, rather than the decoupling view of heuristic reasoning).

## Method

### *Participants*

Twenty children (2 females) with autism (Autistic Disorder: AD) took part in the study (7 children from this group also took part in Experiment 4). The children were between the age of 11 and 16 (mean age 13 years 11 months). Diagnostic records of the children showed that every child had received a diagnosis of autism by experienced clinicians using the guidelines of DSM-IV (American Psychological Association, 1994). No child had a diagnosis of Asperger's syndrome or Pervasive Developmental Disorder. All participants in the AD group had a language delay, and at the time of the testing they took part in language and communication development classes in specialist units within their schools.

In addition, thirty-five typically-developing children (17 females), between the age of 11 and 16 (mean age 13 years 1 month) participated in the study as a control group. An independent samples *t* test indicated that children in the control group were significantly younger than the children with autism ( $t(53)=2.19, p<.05$ ). Children in the control group had no known clinically significant impairment or diagnosis. The criterion for including children in this experiment was that they had to solve all reasoning tasks and individual

differences measures that were administered to the autistic group. Due to the fact that we did not manage to collect data on all tasks from many children in Experiment 3, the control group included in this experiment was around half the size of the original sample. Whether a child was able to complete all the tasks did not reflect their abilities, rather it was determined by the availability of the class that they belonged to.

The same measures of cognitive ability and executive functioning were used as in Experiment 3: the short form of the WISC-III (Wechsler, 1991) consisting of the block design and the vocabulary subtests, Set 1 of the Raven Advanced Progressive Matrices (Raven, Raven & Court, 1998), and the counting span task (see Handley et al., 2004). In addition, children were also administered a computerized version of the Corsi blocks task (based on Vandierendonck, Kemps, Fastame & Szmalec, 2004), a visuospatial counterpart to verbal working-memory span tasks. We used both a forward and a backward recall version of the task. The forward recall version is a purer indicator of visual-spatial working memory, whereas the backwards version additionally loads on central executive resources (Vandierendonck et al., 2004).

Table 5.6. *Comparisons between the autistic and the control group on the measures of cognitive capacity and executive functioning.*

		Mean	<i>t</i> (53)	<i>p</i>
Raven	autistic	7.25 (2.20)	.63	.53
	control	6.83 (2.49)		
Vocabulary	autistic	27.85 (8.96)	.67	.50
	control	29.31 (7.01)		
Block design	autistic	46.55 (12.61)	.08	.94
	control	46.80 (10.47)		
Counting span	autistic	20.85 (10.11)	1.71	.09
	control	25.11 (8.11)		
Corsi forwards	autistic	10.25 (4.22)	.31	.76
	control	9.94 (2.96)		
Corsi backwards	autistic	7.80 (4.10)	1.16	.25
	control	8.84 (2.28)		

Table 5.6. presents the results of the comparisons between the autistic and the control groups on the measures of cognitive capacity and executive functioning. There was no significant difference between the groups on any of these measures apart from a marginally significant difference on the verbal working memory measure, indicating a trend for autistic participants to score lower on this task. In order to take account of differences between groups, we included age as a covariate in our main analyses. Where there was no effect of age, we report the simpler ANOVA models.

### *Materials and procedure*

There were sixteen tasks altogether, four tasks measuring each type of heuristic: the representativeness heuristic (on the engineers and lawyers problem with a representative description), the equiprobability bias (on the engineers and lawyers problem with a non-representative description), the hot hand / the gambler's fallacy (i.e., positive/negative recency effects), and the conjunction fallacy. We used a 2x2 design where tasks measuring each heuristic were presented both with a social and a non-social content, and two out of the four tasks were conflict problems, and the other two were non-conflict problems. In the present study we employed three analogical reasoning tasks. In addition to the APM, children were also administered a picture analogy task (Krawczyk et al., 2008), and a scene analogy task (Richland et al., 2006). Both the picture analogy and the scene analogy task require the activation and application of real-life knowledge. In the more difficult versions of both tasks, participants have to resist perceptual and semantic interference. For a description of each task see the method section of Experiment 3. The full list of tasks can be found in Appendix B.

All the problems were presented in a booklet. The different types of reasoning problems were mixed together in a random order. The analogical reasoning tasks were presented after the heuristics and biases problems in a fixed order. The experimenter read



out the instructions and participants worked through the problems at their own pace. The individual differences measures were administered individually in a separate testing session.

## Results

*The effects of conflict, content and group membership on performance on the reasoning and judgment tasks.*

2x2x2 mixed ANOVAs were run for each type of task with group (autistic/control) as a between-subjects variable and problem type (conflict/non-conflict), and content (social/non-social) as within-subjects variables. As in Experiment 3, the differences in normative responding across conflict and non-conflict tasks were analysed. The purpose of these analyses was to assess children's susceptibility to the different heuristics measured by the tasks. Due to the significant difference in age between groups, we included this measure as a covariate in our analyses. Where age had no effect on the results, we report the simple ANOVA analyses. A supplementary table for the analyses can be found in Appendix S5 (Table 1).

The first task was the *engineers and lawyers problem with a representative description*. Table 5.7. displays the proportion of children who chose each type of response. The ANOVA indicated a significant effect of conflict ( $F(1,51)=20.54, p<.001, \eta_p^2=.29$ ), and a significant conflict by age interaction ( $F(1,51)=7.01, p<.05, \eta_p^2=.12$ ) on normative responses.

In order to examine the nature of this interaction further, we divided our sample into two age groups. There were 29 children in the younger group (8 autistic children and 21 typically developing children; mean age: 12 years 2 months) and 26 children in the older group (12 autistic and 14 typically developing children; mean age: 14 years 8 months), and we compared the difference between normative responses on the conflict vs. non-conflict problems. That is, we subtracted the number of normative responses on the conflict

problems from normative responding on the non-conflict problems. As in Experiment 3, we used this measure as an indicator of susceptibility to heuristics (in the present case, the measure of susceptibility to representative information that is in conflict with base rates). An independent samples *t* test indicated that younger children ( $M=1.62, SD=.56$ ) were more susceptible to representativeness than older children ( $M=1.12, SD=.71; t(53)=2.94, p>.01$ ). That is, base rate information had a greater effect on children's responses as they got older.

Table 5.7. *The proportion of participants choosing each type of response across groups, content and conflict conditions on the engineers and lawyers problem with a representative description.*

		Conflict			Non-conflict		
		rep.	base rate (non-rep.)	equally likely	rep. / base rate	non-rep.	equally likely
Aut.	Social	.75 (.44)	<b>.20 (.41)</b>	.05 (.22)	<b>.80 (.41)</b>	.05 (.22)	.15 (.37)
	Non-social	.80 (.41)	<b>.15 (.37)</b>	.05 (.22)	<b>.80 (.41)</b>	.05 (.22)	.15 (.37)
Control	Social	.77 (.43)	<b>.09 (.28)</b>	.14 (.36)	<b>.80 (.41)</b>	.03 (.17)	.17 (.38)
	Non-social	.83 (.38)	<b>.14 (.36)</b>	.03 (.17)	<b>.89 (.32)</b>	.07 (.24)	.04 (.17)

The second task was the *engineers and lawyers problem with a non-representative description* (see Table 5.8. for descriptives). First, we analysed the effect of conflict, content and group membership on giving base rate (i.e., normative) responses. There was a significant effect of conflict, indicating that children took base rate information less into consideration in the case of conflict tasks, where they were provided with some irrelevant individuating information ( $F(1,53)=15.47, p<.001, \eta_p^2=.23$ ). Additionally, there was a

significant conflict by group interaction ( $F(1,53)=4.45, p<.05, \eta_p^2=.08$ ). This indicated that in the case of conflict tasks autistic children took base rate information more into account than children in the control group (that is, they were less sensitive to the conflict between base rate information and representativeness). This interaction remained significant even after including age as a covariate ( $F(1,52)=4.21, p<.05, \eta_p^2=.08$ ).

Table 5.8. *The proportion of children choosing each type of response across groups, content and conflict/non-conflict tasks on the engineers and lawyers problem with a non-representative description.*

		Conflict			Non-conflict		
		other	base rate	equally likely	other	base rate	equally likely
Autistic	Social	.15 (.37)	<b>.50 (.51)</b>	.35 (.49)	0 (0)	<b>.75 (.44)</b>	.25 (.44)
	Non-social	.25 (.44)	<b>.60 (.50)</b>	.15 (.37)	.25 (.44)	<b>.60 (.50)</b>	.15 (.37)
Control	Social	.34 (.48)	<b>.31 (.47)</b>	.34 (.48)	.06 (.23)	<b>.74 (.44)</b>	.20 (.41)
	Non-social	.26 (.44)	<b>.29 (.46)</b>	.43 (.50)	.09 (.28)	<b>.69 (.47)</b>	.23 (.43)

The third task investigated the *positive/negative recency effects* (see Table 5.9. for descriptive statistics). The analysis of the normative responses indicated a significant effect of conflict ( $F(1,53)=16.12, p<.001; \eta_p^2=.24$ ). As expected, participants gave more normative responses in the case of non-conflict problems.

In order to see whether we find the content effects that would be expected based on the literature, the responses corresponding to both the positive and the negative recency effect were analysed separately. The ANOVA on the positive recency responses indicated a main effect of content ( $F(1,53)=6.43, p<.05, \eta_p^2=.11$ ) and a main effect of conflict

( $F(1,53)=14.00, p<.001, \eta_p^2=.22$ ). That is, children were more likely to exhibit the positive recency effect when they had to predict an outcome which followed a seemingly non-random sequence (as opposed to a sequence that seemed random). In line with the literature on this bias, the positive recency effect was more pronounced when the task had a social (as opposed to a non-social) content. There was also a content by conflict interaction ( $F(1,53)=17.86, p<.001, \eta_p^2=.26$ ). This indicated that the content manipulation only affected performance on the conflict problems (i.e., where the sequence seemed non-random).

Table 5.9. *The proportion of participants choosing each type of response on the positive/negative recency problems across groups, content and conflict/non-conflict tasks.*

		Conflict			Non-conflict		
		positive recency	negative recency	equally likely	Response 1	Response 2	equally likely
Autistic	Social	.47 (.51)	0 (0)	<b>.53 (.51)</b>	0 (0)	.06 (.24)	<b>.94 (.24)</b>
	Non-social	.06 (.24)	.29 (.47)	<b>.65 (.49)</b>	.06 (.24)	.12 (.33)	<b>.82 (.39)</b>
Control	Social	.26 (.44)	.03 (.17)	<b>.69 (.47)</b>	.03 (.17)	.03 (.18)	<b>.94 (.24)</b>
	Non-social	.11 (.32)	.11 (.32)	<b>.77 (.43)</b>	.06 (.24)	0 (0)	<b>.94 (.24)</b>

A similar analysis regarding the negative recency effect indicated a significant effect of content ( $F(1,53)=8.18, p<.01; \eta_p^2=.14$ ), a content by group interaction ( $F(1,53)=4.25, p<.05; \eta_p^2=.08$ ), and a content by conflict interaction ( $F(1,53)=8.13, p<.05; \eta_p^2=.14$ ). This indicated that the negative recency effect was stronger in the case of tasks with a non-social (as opposed to a social) content. The effect was also stronger in the case of conflict tasks. The content by group interaction showed that autistic children were more likely to

commit the fallacy when the tasks were about objects, whereas in the control group there was no difference between tasks with social and non-social contents (in fact, children in the control group did not show a susceptibility to this fallacy). The content by group interaction remained significant even after controlling for age ( $F(1,52)=4.42, p<.05; \eta_p^2=.08$ ).

In the case of *the conjunction fallacy* task (see Table 5.10.) heuristic and normative responses were mutually exclusive. The ANOVA indicated a significant effect of conflict ( $F(1,53)=8.85, p<.01, \eta_p^2=.14$ ). That is, children were more likely to commit the conjunction fallacy if the conjunction contained a representative event than when it consisted of two non-representative events. There was also a significant group by conflict interaction ( $F(1,53)=4.36, p<.05, \eta_p^2=.08$ ). This indicated that autistic children were less sensitive to conflict than children in the control group. However, after controlling for the effects of age, only the conflict by group interaction remained significant ( $F(1,52)=4.00, p<.05, \eta_p^2=.07$ ).

Table 5.10. *The proportion of participants choosing each type of response on the conjunction fallacy problems across groups, content and conflict/non-conflict tasks.*

		Conflict		Non-conflict	
		conjunction fallacy	normative	conjunction fallacy	normative
Autistic	Social	.80 (.41)	<b>.20 (.41)</b>	.70 (.47)	<b>.30 (.47)</b>
	Non-social	.65 (.49)	<b>.35 (.49)</b>	.65 (.49)	<b>.35 (.49)</b>
Control	Social	.94 (.24)	<b>.06 (.24)</b>	.54 (.50)	<b>.46 (.50)</b>
	Non-social	.86 (.36)	<b>.14 (.36)</b>	.69 (.47)	<b>.31 (.47)</b>

In summary, on three out of four tasks (with the exception of the engineers and lawyers problem with a representative description) we have found evidence for differences between groups. Autistic children were less susceptible to the equiprobability bias in the

case of the engineers and lawyers problem with a non-representative description, and they were also less susceptible to the conjunction fallacy. On the other hand, the autistic group showed a stronger negative recency effect than the control group.

*Analogical reasoning performance*

The following analyses concerned children’s performance on the analogical reasoning tasks. As before, we also ran these analyses including age as a covariate. As including this measure as a covariate did not change the results, we report the results of the simple ANOVAs below. We have already reported the comparison between groups in the case of the APM in the participants section. We also analysed the relative difficulty of the Raven items in both groups. The purpose of this analysis was to check whether the test measured the same construct in the case of the two groups. Dawson et al. (2007) reported that item difficulty on the Raven test was highly correlated between autistic and control children. Similarly, in the present sample the correlation was  $r(10)=.89, p<.001$ .

Performance on the picture analogy tasks (see Table 5.11.) was analyzed using a 2x2 mixed ANOVA with distraction (present/absent) as a within-subjects factor, and group (autistic/control) as a between-subjects factor. A main effect of distraction was obtained,  $F(1,53)= 88.46, p<.001, \eta_p^2=.63$ ). There was no effect of group, and no distraction by group interaction. This indicated that autistic and control participants suffered equal impairment from the presence of distractors as response options.

Table 5.11. *Proportion of correct responses on the picture analogy tasks.*

	Non-distractor M (SD)	Distractor M (SD)
Autistic	.73 (.08)	.54 (.15)
Control	.75 (.08)	.53 (.17)

Additional independent samples *t* tests were run to test for possible differences in the number of perceptual, semantic and other errors that children made on the picture

analogy task (see Table 5.12.). There was no difference between groups in the proportion of errors that they made ( $t_s < .80$ ,  $p_s > .45$ ; the full results can be found in Appendix S5 Table 3).

Table 5.12. *Proportion of perceptual, semantic and other errors across groups on the picture analogy tasks.*

	Perceptual	Semantic	Other
Autistic	.03 (.04)	.13 (.07)	.05 (.07)
Control	.02 (.03)	.14 (.09)	.04 (.06)

In order to further assess whether the autistic and control groups solved the picture analogy problems using similar strategies, an item analysis was performed. For each group, the mean percent correct was obtained for each of the 16 basic items, collapsing across the distractor and non-distractor versions. If there are reliable differences among items, and if both sets of participants solved the problems using similar strategies, then we would expect to find a robust correlation between the difficulty of the individual items across the groups. The relative difficulty of the picture analogy items was indeed highly correlated between the autistic and control children,  $r(14) = .81$ ,  $p < .001$  suggesting that both groups solved the problems using similar strategies.

Finally, Table 5.13. depicts mean performance on the scene analogy task. We performed a 2x2x2 mixed ANOVA on the scene analogy task to examine the effects of relational complexity (one or two relations: within-subjects), distraction (present/absent: within-subjects), and group (autistic/control: between-subjects). This analysis revealed a main effect of relational complexity ( $F(1,53) = 22.88$ ,  $p < .001$ ,  $\eta_p^2 = .31$ ), a main effect of distraction ( $F(1,53) = 28.77$ ,  $p < .001$ ,  $\eta_p^2 = .36$ ), and a relational complexity by distraction interaction ( $F(1,53) = 10.23$ ,  $p < .01$ ,  $\eta_p^2 = .17$ ). This pattern reflected the fact that children performed worse on problems with higher relational complexity or in the presence of distraction. As is apparent in Table 5.13., the profile of accuracy across conditions was

very similar for autistic and control children, and the performance of autistic children was again statistically indistinguishable from that of the control group.

Table 5.13. *Proportion of correct responses on the scene analogy tasks.*

	Autistic M (SD)	Control M (SD)
One relation, non-distractor	.89 (.10)	.88 (.18)
One relation, distractor	.61 (.35)	.66 (.24)
Two relations, non-distractor	.67 (.22)	.71 (.24)
Two relations, distractor	.60 (.26)	.64 (.22)

We also compared the groups with respect to the number of perceptual, relational and other errors that they made on the scene analogy task (see Table 5.14.), using independent samples *t* tests. There was no difference between groups in the proportion of errors that they made ( $t_s < .92$ ,  $p_s > .36$ ; the full results can be found in Appendix S5 Table 3).

Table 5.14. *Proportion of perceptual, relational and other errors across groups on the scene analogy tasks.*

	Perceptual M (SD)	Relational M (SD)	Other M (SD)
autistic	.13 (.13)	.13 (.10)	.05 (.06)
control	.10 (.08)	.13 (.07)	.05 (.09)

Finally, we also performed an item analysis similar to that reported above for the picture analogy task. For each group, we obtained the mean percent correct for each of the 20 basic pairs of scenes, collapsing across the four conditions. The relative difficulty of the scene analogy items was highly correlated between the autistic and control children,  $r(18) = .71$ ,  $p < .001$ , again suggesting that the two groups used similar strategies to perform the analogy task.



*Relationships between the measures of individual differences in age, cognitive capacity and analogical reasoning performance in the autistic group*

The following analyses will only concern the autistic group. The same analyses were run in Experiment 3 with a (bigger) group of typically developing children, and those results can be used as reference for comparing the underlying processes in the autistic and the typically developing group. As we did in Experiment 3, we created an analogical reasoning composite score by adding up the z scores of each measure of analogical reasoning (the APM, performance on the distractor and non-distractor trials of the picture analogy task, and performance at the different levels of relational complexity and distraction on the scene analogy task). The Cronbach's alpha for the analogy measure was .57 for the autistic group (a Cronbach's alpha above .5 is considered to be an acceptable level of reliability – see Rust and Golombok, 1999).

Table 5.15. *Correlations between age and the cognitive ability measures in the autistic group (the correlations for the typically developing sample in Experiment 3 are displayed in brackets).*

	1.	2.	3.	4.	5.	6.
1. age	--					
2. counting span	.42 (.18)	--				
3. Corsi forwards	-.03	.32	--			
4. Corsi backwards	.02	.62**	.59**	--		
5. analogy composite	.29 (.28*)	.36 (.37**)	.26	.51*	--	
6. WISC composite	.25 (.42**)	.39(.42**)	.23	.67**	.43 (.59**)	--

Table 5.15. displays the correlations between age and the cognitive ability measures in the autistic group. Although the sample size was very small, there were some strong correlations between measures that reached significance. As expected, analogical

reasoning was correlated with the Corsi backwards score, which is a measure of non-verbal working memory and executive functioning. Interestingly, the Corsi backwards score was also strongly correlated with performance on the verbal working memory task, although verbal and non-verbal working memory scores are supposed to be at least partly independent (for a comparison, the corresponding correlation in the control group was  $r(29)=-.02$ , n.s.). As well as being a measure of non-verbal working memory, the backwards version of the Corsi blocks task is also a measure of executive functioning. In Experiment 4 there was also a significant correlation between verbal working memory and set-shifting ability (which is also related to executive functioning). In the control group these correlations were small in size, and non-significant. It has been suggested (e.g., Miyake et al., 2004) that inner speech is an important retrieval aid in cueing task goals, and goal retrieval is disrupted by articulatory suppression. As verbal working memory is often impaired in autism, it is possible that, at least in some cases, executive function deficits arise as a result of a problem with verbal processing.

#### *Relationships between reasoning performance across tasks in the autistic group*

We also investigated the relationship between reasoning performance on the different tasks (see Table 5.16.). Based on the results of the previous experiments (apart from Experiment 1) it can be expected that participants giving heuristic responses on one task will not necessarily give heuristic responses to another task. We can also expect both positive and negative associations between susceptibility to heuristics on the different tasks. In line with this expectation, there was no significant relationship between heuristic responding across tasks (although there was a medium-sized negative correlation between heuristic responding on the engineers and lawyers task with a representative description, and between the positive/negative recency effect). These results need to be taken with caution, as the sample size was very small ( $n=20$ ). Nevertheless, the pattern of correlations appears to be different from what was found in the typically developing group. Thus,

unlike in the case of the analogical reasoning tasks, there was some indication that the processes underlying performance on these tasks differed between groups. (However, note that the differences between correlations across groups were not even close to being significant,  $ps > .27$  as indicated by Fisher  $r$  to  $z$  transformations.)

Table 5.16. *Correlations between susceptibility to different biases in the autistic group (correlations for the typically developing group from Experiment 3 are displayed in brackets).*

	1.	2.	3.	4.
1. Engineers and lawyers problem, repr.	--			
2. Engineers and lawyers problem, non- repr.	.05 (.34**)	--		
3. Positive/negative recency	-.30 (.08)	.004 (.08)	--	
4. Conjunction fallacy	-.18 (.07)	.08 (-.20)	-.19 (-.35**)	--

*Relationships between performance on the reasoning tasks and age, cognitive capacity and analogical reasoning*

Our final analysis concerned the relationships between susceptibility to biases and the measures of age, cognitive capacity, and analogical reasoning in the autistic sample (the correlations are displayed in Table 5.17.). Based on the correlations between the different measures of analogical reasoning, we created an analogy composite score (see above). In addition, as a measure of cognitive capacity we used the combined  $z$  scores of the WISC, the counting span task, and the backward version of the Corsi blocks task. These measures were strongly associated with each other in the case of the autistic group (see Table 5.15.). The Cronbach's alpha for this measure in the autistic group was .56. As before, the susceptibility scores were created by computing the difference between normative responding on the non-conflict and conflict versions of each task. This was done in a way so that greater susceptibility to a bias resulted in a higher score. As we highlighted

before, the correlational analyses relating to the autistic sample have to be interpreted with caution due to the small sample size ( $n=20$ ).

Table 5.17. *Correlations between age, the cognitive capacity measures and susceptibility to different biases in the autistic group. (Correlations that were significant at the  $p < .10$  level are marked with †, correlations for the typically developing group from Experiment 3 are displayed in brackets)*

	age	Cognitive capacity	Analogical reasoning
Engineers and lawyers problem, repr.	<b>-.46*</b> (.04)	<b>-.42<sup>†</sup></b> (.12)	<b>-.31</b> (.36**)
Engineers and lawyers problem, non-repr.	<b>-.12</b> (.15)	<b>.38<sup>†</sup></b> (.04)	<b>.25</b> (.17)
Positive/negative recency	<b>-.07</b> (.16)	<b>.08</b> (.13)	<b>.03</b> (-.06)
Conjunction fallacy	<b>.30</b> (-.22)	<b>.48*</b> (-.15)	<b>.33</b> (-.14)

Based on the results of Experiment 4, and also on the weak central coherence theory, we expected that the relationship between susceptibility to biases and cognitive ability will be either positive (indicating that producing a heuristic response requires conscious effort) or there will be no relationship (indicating either that participants with autism do not engage in the contextualisation process, or that they are able to contextualise without effort). In line with this prediction, there was evidence that susceptibility to the conjunction fallacy, and to the equiprobability bias (on the engineers and lawyers problem with a non-representative description) were positively related to cognitive capacity. On the other hand, susceptibility to the representativeness heuristic (in the case of the engineers and lawyers problem with a representative description) was negatively correlated with cognitive capacity. After controlling for age all these correlations were non-significant at the  $p < .05$  level. However, the strength of correlations remained at a medium level;  $r(17) = -.32$  ( $p > .10$ ) for the representativeness heuristic and cognitive ability,  $r(17) = .45$  ( $p < .10$ ) for

cognitive ability and the equiprobability bias, and  $r(17)=.42$  ( $p<.10$ ) for cognitive ability and the conjunction fallacy. The positive/negative recency effect was unrelated to children's cognitive capacity.

Our final analysis concerned the use of social stereotypes in reasoning. In order to do this, we compared the performance of the two groups in the case of the engineers and lawyers problem with a representative description. Specifically, we looked at whether autistic and control participants were equally likely to choose the representative response in the case of non-conflict tasks with both social and non-social content. An ANOVA on the representative responses with group (autistic/control) as a between-subjects factor, and content (social/non-social) as a within-subjects factor indicated no effect of conflict or group, and no conflict by group interaction ( $F_s<.35$ ,  $p_s>.55$ ). We also compared the proportion of participants in each group who rated the representative item as the most likely option on the non-conflict conjunction fallacy problems. The purpose of this analysis was to evaluate whether both groups identified the representative item as the most likely to the same degree. This would indicate that similar knowledge was activated and employed in making judgments about the simple options. The analysis revealed no significant difference between the proportion of participants from the control and the autistic groups ranking the representative item as the most likely either for tasks with a non-social content (82% vs. 61%,  $\chi^2(1)=2.27$ , n.s.) or for tasks with a social content (90% vs. 83%,  $\chi^2(1)=.79$ , n.s.).

## Discussion

The results of this study were mostly in line with the findings of Experiment 4. Probably the most important result is that once again we found evidence for both less and more bias in the autistic group, as compared to controls. More specifically, autistic children were less susceptible to the equiprobability bias in the case of the engineers and lawyers problem with a non-representative description, and to the conjunction fallacy than controls.

On the other hand, autistic children were more susceptible to the negative recency effect (at least in the case of tasks with a non-social content) than controls. These results suggest that “heuristic” responding (that is, producing responses that are traditionally considered to be heuristic) on the different tasks have different processing requirements. Instead of heuristic responses being produced by a different type of process than normative responses, based on the present data it seems more likely that producing both heuristic and normative responses requires a certain amount of cognitive effort. Overall it seems that reasoning performance differs between groups, although the differences are hard to pin down. In the case of certain problems autistic participants seem to show a developmental delay (i.e., they perform similarly to a younger typically developing group) and they tend to produce fewer of those responses that require conscious effort in the case of the control group. Additionally, it is important to note that there was no main effect of group on normative responses in the case of any of the tasks, and the performance of autistic and typically developing children on non-conflict problems was very similar. This suggests that autistic participants did not show a general tendency for normative, rule-based (or logical structure-based) reasoning on all tasks, rather they exhibited a distinct pattern of responding on conflict tasks only, where contextualised responding played a major role. This is in contrast with the predictions of the extreme male brain theory, but it is in line with the weak central coherence account.

Looking at the content effects, it seems unlikely that the differences stem from autistic children’s having specific problems with reasoning within the social domain (see, for example, Adolphs, 1999). In fact, when there was a difference between groups in how they reasoned about tasks with a certain content, it affected reasoning about non-social, rather than social problems (such as in the case of the negative recency effect). Recent evidence suggests that stereotypes are activated and used as readily for autistic as typical populations (Hirschfeld, et al., 2007). In order to check if this is the case, we compared the performance of the two groups in the case of the engineers and lawyers problem with a

representative description. These results indicated that autistic children used stereotype information for drawing simple inferences as readily as typically developing children. That is, whatever was the basis of the differences between reasoning performance across groups, had little to do with the content of the problems. On the other hand, autistic children were more susceptible to the negative recency effect than the control group (which is a bias that affected reasoning about objects). Finally, we only found evidence for a general content effect in the case of the positive/negative recency task, which is in line with the literature (see e.g., Ayton & Fischer, 2004).

Another interesting aspect of the present study is the data regarding executive functioning in autism. One important finding is that in the case of these high-functioning children with autism, there was little evidence for a general deficit in executive functioning. In fact, the only aspect of executive functioning that was impaired in both experiments was verbal working memory. On the other hand, in both experiments there was evidence that verbal working memory capacity was related to other aspects of executive functioning for the autistic group. In Experiment 4, there was a significant correlation between verbal working memory and set-shifting scores, and in Experiment 5 verbal working memory correlated with performance on the backward version of the Corsi blocks task (which is a measure of both non-verbal working memory and central executive resources). By contrast, in the control groups verbal working memory and other aspects of executive functioning were not significantly related. These findings are in line with Benetto et al. (1996) who found correlations between memory and executive functioning in autism, and with McCroy et al. (2007), who reported that memory retrieval in autism was closely related to executive functioning, whereas there was no such relationship in the control group. According to Miyake et al. (2004) verbal working memory, and inner speech play a strategic role in the retrieval of relevant background knowledge and task goals. Thus, it is possible that autistic children with a working memory deficit will be

disadvantaged in the strategic use of knowledge, which is a central aspect of executive functioning.

Another aim of this study was to investigate analogical reasoning in autism, and its relationships to heuristic reasoning. We compared the performance of autistic and typically developing children on the Raven Advanced Progressive Matrices test and two knowledge-based tests of analogical reasoning. We found no difference between the two groups on any of the analogical reasoning tasks. Performance of the autistic and the typically-developing children was equally impaired by increased relational complexity and by the presence of perceptual and semantic distractors.

Our results for these analogy tasks confirmed patterns observed previously with non-autistic populations (Krawczyk et al., 2008; Richland et al., 2006), but provided no evidence of differential performance between the autistic and non-autistic groups. Moreover, we replicated the finding of Dawson et al. (2007) that the relative difficulty of the Raven items was highly correlated between the autistic and control children, arguing against the possibility that the two groups use substantially different strategies to solve the problems. In addition, the difficulty of individual items on both the picture analogy and scene analogy tests correlated highly between the two groups. Although relatively superior performance of autistic people on the Raven test (as compared to general intelligence as measured by the WISC) could be related to enhanced perceptual processing, it seems implausible to attribute normal performance on the picture analogy and scene analogy problems to group differences in perceptual abilities, as the relationally-matched source and target items in these problems were visually dissimilar. There was also no evidence for superior autistic performance on the Raven test in the case of our sample.

Our findings support the hypothesis that the ability to reason analogically is intact in autism, not only for abstract problems such as the Raven test (Dawson et al., 2007), but also for knowledge-based problems (the picture analogy and scene analogy tasks). The present findings are consistent with those of previous studies that found intact analogical



reasoning performance in autistic people (e.g., Dawson et al., 2007; Scott & Baron-Cohen, 1996). These results suggest boundary conditions on the hypothesis that autistic people have problems with using their knowledge spontaneously and flexibly in novel situations (Peterson & Bowler, 2000), insofar as basic analogical reasoning would seem to require some ability to use knowledge flexibly. Our results are also inconsistent with the hypothesis that autistic people exhibit reduced processing of similarity between stimuli and situations (Plaisted, 2001), as analogical transfer crucially depends on recognising similarities. Contrary to predictions based on the weak central coherence theory (Frith & Happé, 1994), autistic children were able to create systematic mappings between stimuli (which requires processing of stimuli in context and establishing coherence), even in the face of perceptual distractors that encourage use of a perceptual matching strategy. Altogether, the present findings paint a picture very different from Kanner's original description of autism, which stated that "If the slightest ingredient [of a situation, performance or sentence] is altered or removed, the total situation is no longer the same and therefore is not accepted as such..." (Kanner, 1943, p. 246).

The prerequisites for successful analogical reasoning include access to relevant real-world knowledge, as well as intact working memory and executive functions. Autistic participants in the present study were well-matched to controls on executive functioning, apart from a marginal difference in verbal working memory scores. As autistic people sometimes have impairments of executive functions when compared to IQ-matched controls (for reviews see Hill, 2004; Rajendran & Mitchell, 2007), it is possible that some autistic people will have problems with analogical reasoning when they are matched to controls on IQ only, but not on measures of executive functioning. (See Pellicano, Maybery, Durkin & Maley, 2006, for evidence concerning the relationship between executive function deficits and problems with integrating information.) Autistic people may also have difficulty when analogical reasoning requires the processing of complex

verbal material (Lopez & Leekham, 2003) or picking up and integrating cross-modal cues (Pexman, 2008).

In summary, the present findings indicate that autism *per se* does not imply an inability to reason analogically. It is important to note that all the analogy tasks used in the present study involved pictorial rather than verbal presentation of the problems. Moreover, the problems were either purely formal (APM), or required highly constrained knowledge largely provided directly by the pictorial stimuli (picture analogies and scene analogies). According to the Task Support Hypothesis (Bowler et al., 2004), autistic individuals show greater difficulty in retrieving and integrating background knowledge with a problem context when retrieval is not directly cued by a task. Our findings leave open the question of whether a difference in reaction times would be found between the autistic and control groups when solving the present tasks (indicating a difference in efficiency of processing), as well as the question of whether analogical reasoning ability in autism would also be intact for tasks that provide less support for memory retrieval.

Given that the same autistic participants whose performance on the analogical reasoning problems was indistinguishable from controls, showed distinct patterns of reasoning performance on the heuristics and biases tasks, we can expect that (in line with the Task Support Hypothesis - Bowler et al., 2004) autistic participants would show impairments in reasoning about problems that require the spontaneous activation of relevant background knowledge.

### 5.3. *General discussion*

The aim of the studies that we described in this chapter was twofold. One aim was to test some of the claims of dual-process theorists using a special population. The other aim was to find out more about reasoning in autism, especially in the domains of using reasoning heuristics, and making analogies. Although we tried to make some predictions regarding the expected patterns based on the prominent theories of autism, this was not

easy, because higher order reasoning in autism has not been extensively researched. The only relevant existing study was an investigation into the framing effect in risky choices by De Martino and colleagues (2008). This study found increased logical performance (and less sensitivity to framing effects) in autism. This is in line with the weak central coherence account, which predicts reduced contextualisation in autism, or, alternatively, a tendency to produce contextual responses through conscious effort, instead of a spontaneous appreciation of context.

Table 5.18. summarizes our findings in the two experiments. We used seven different tasks, and one of them (the conjunction fallacy) was included in both experiments. When comparing the autistic and the control groups, we found that in the case of the conjunction fallacy task, and in the case of the engineers and lawyers problem with a non-representative description, autistic children tended to give less heuristic/more normative responses, and they seemed to produce heuristic responses effortfully. Although the results concerning the belief bias are seemingly in contrast with this, the possible explanation for less belief bias in the typically developing group might be that they benefitted more from the fantasy context than the autistic group (see Leever & Harris, 1999). Thus, although the outcome is the opposite in the case of this task, it is possible that this difference was also the result of reduced contextualisation in the autistic group. In the case of the if-only bias, although there was no difference between groups in terms of their responses, heuristic responding seemed to be based on effortful reasoning in the case of autistic children. Finally, autistic children gave less normative responses in the sunk cost task, and in the negative recency task. In the sunk cost task they also seemed to exhibit a tendency to give reverse sunk cost responses (i.e., to abandon a situation more quickly if they made previous investments), thus exhibiting a different bias from the control group. Although we did not find direct evidence for this, it is possible that the increase in biased responding in the autistic group in these tasks might be associated with a reduced ability to control automatic responses, or maybe these responses were not produced completely

automatically, which resulted in less metacognitive awareness of these responses (or a restricted availability of cognitive resources to manipulate these representations), compared to the control group. This could have led to a reduced ability to resist these fallacies.

Table 5.18. *Summary of the results of Experiments 4 and 5.*

<b>Type of heuristic</b>	<b>Experiment</b>	<b>Any difference between groups?</b>	<b>Relationship between heuristic responding and cognitive capacity</b>
<i>Belief bias</i>	1	autistic children gave more biased responses	none
<i>Sunk cost fallacy</i>	1	autistic children gave less normative responses	none
<i>If-only bias</i>	1	none	positive correlation with WISC scores
<i>Conjunction fallacy</i>	1 and 2	autistic children gave less heuristic responses (but the effect disappeared in Exp. 1 after controlling for vocabulary)	positive correlation with WISC scores in Exp. 1., positive correlation with cognitive capacity in Exp. 2
<i>Engineers and lawyers problem, representativeness</i>	2	none	negative correlation with cognitive capacity
<i>Engineers and lawyers problem, equiprobability</i>	2	autistic children gave more normative responses	positive correlation with cognitive capacity
<i>Positive/negative recency</i>	2	autistic children gave less normative responses to negative recency tasks	none

In summary, although most of these findings can be understood in terms of reduced contextualisation / contextualisation through effort in autism, this is clearly not the whole story. This is reflected in the fact that in the case of some tasks, autistic participants were more susceptible to fallacies than typically developing children. Although our autistic participants did not show impairments in their executive functioning skills, they might be less able to manipulate and control verbal representations on-line due to a reduced ability to activate these representations automatically. According to Bowler et al.'s (2004) task support hypothesis autistic individuals often show difficulty in retrieving and integrating their background knowledge with a problem context when the task provides little support

for retrieval (i.e., when concepts are not directly presented or cued, or when no explicit instructions are given to search for a concept – see also McCroy et al., 2007).

Moving on to the issue of executive functioning in autism, we measured several aspects of executive functioning in our samples (verbal working memory, inhibition and set-shifting in Experiment 4, and verbal and non-verbal working memory, and central executive resources in Experiment 5). Additionally, the measures of analogical reasoning can be considered as indices of executive functioning or “decoupling”, especially the trials that require participants to resist interference while performing mapping between representations. Importantly, apart from an impairment in verbal working memory, we found no evidence for executive functioning deficits in the autistic sample. However, in both experiments there was a relationship between verbal abilities and measures of executive functioning in the autistic sample (which was not found or was much weaker in the control group), which indicated that the organization of cognitive abilities in the two samples might be different. A possible explanation for these findings is that children with autism do not benefit from a spontaneous organization of verbal materials, because they have verbal working memory deficits (see Miyake et al., 2004), and as a result they have to rely on their general executive resources during memory retrieval (see e.g., McCroy et al., 2007; see also Section 4.2.2. for a review on memory retrieval in autism). However, when retrieval is supported by external cues (such as pictures or explicit questioning) individuals with autism are able to retrieve memory contents as easily as typically developing individuals (see Bowler et al., 2004). This would explain why children with autism in our sample had no problem with analogical (i.e., contextualised) reasoning about pictorial stimuli, whilst they showed a distinct pattern of reasoning about heuristics and biases tasks. If children with autism need to rely on effortful processes when they retrieve memory contents, then when they reason about contextualised problems they essentially perform a dual task. This could explain why the correlations between giving particular responses to the heuristics and biases tasks and cognitive capacity were different between groups,

although the general cognitive resources of autistic children seemed to be intact. This also explains why autistic children reasoned about these problems in a way which was similar to the performance of “lower ability” (i.e., younger) typically developing children. These findings are also in line with Lopez and Leekam (2003) who reported that contextual processing was unimpaired in autism, apart from when it required the integration of complex verbal materials. The results of Experiment 5 also clarified that the differences between groups were not restricted to reasoning about the social domain. In addition, children with autism only showed a distinct pattern of performance in the case of conflict tasks (which required decontextualisation), but not in the case of non-conflict tasks (which could be solved by relying both on contextualised and decontextualised reasoning). Finally, as Table 5.18. demonstrates, children with autism were sometimes more biased on conflict tasks, and sometimes they were less biased (we found this in both experiments, using a range of different tasks). This clearly shows that the differences between groups cannot be explained in terms of a tendency for one group to rely more on logic/rules/statistical information. It is also not possible to describe the differences between groups as restricted to Type 1 or Type 2 processes.

When interpreting these findings it is important to keep in mind some methodological issues. First of all, autism is a diverse disorder. Certain behavioural symptoms of autism are similar across individuals – however, the origins of these symptoms might differ. Although researchers of autism usually consider their autistic sample as a homogeneous group, it makes little sense to generalize certain characteristics to all individuals with autism. There are theories that aim to describe the autistic cognitive profile (such as the weak central coherence account, the executive dysfunction theory, the theory of mind hypothesis and the extreme male brain theory). However, there is evidence from several studies which shows that there are many autistic individuals who do not suffer from executive function or theory of mind deficits (see Sections 4.1.1. and 4.1.3. for references). Similarly, the claims of the extreme male brain theory are very controversial

(see Section 4.1.4.). Probably the weak central coherence account of autism is based on the strongest evidence. Nevertheless, even this theory has been severely criticised (see e.g., Mottron et al., 1999; Pellicano et al., 2005). Although many studies found evidence for an autistic advantage on the block design task (most famously, Shah & Frith, 1993 who also showed that the autistic advantage disappears when the patterns are pre-segmented) in the present studies we found no evidence for an autistic advantage on this task. In fact, in both experiments our autistic participants scored lower than the control group (although the differences were very small, and far from being statistically significant). It could be argued that this might be an evidence for an inappropriate matching of samples in the present studies (see next paragraph for a discussion on the issue of matching). However, we are not alone in finding no difference between groups on this task (see White, O'Reilly & Frith, 2009 for a review). In fact, besides many studies reporting no difference between groups, one study found *impaired* rather than enhanced performance in autism on three visuospatial tasks (Burnette, Mundy, Meyer, et al., 2005). In recent years it has been suggested by a couple of researchers that the autistic advantage on these tasks might be carried by a significant minority of the autistic sample. For example, White et al. (2009) suggested that maybe only those with macrocephaly (i.e., increased head/brain size – which affects approximately 20% of all autistic individuals) exhibit weak central coherence.

Now we turn into the issue of matching the autistic and typically developing samples. There are several ways of matching samples. In developmental research, and when participants have a learning disability one of the most commonly used matching strategies is to have two control groups: a cognitive ability and age-matched group with a learning disability, and an ability-matched (younger) typically developing group (see e.g. Scott & Baron-Cohen, 1996). Another possible strategy, which is often used in the case of high-functioning autistic participants (like our present sample) is to match the samples individually (see e.g., Bowler et al., 2004). Although this strategy makes it possible to have

two samples that are fairly matched on a number of important measures, the problem with this method is that the typically developing sample is not a random sample, and as such it is likely to be *not* representative to the typically developing population. Another possible problem with this strategy is that it might under-or overestimate the abilities of the autistic sample depending on the nature of the tasks that are used for matching (see Mottron, 2004). One way to overcome the problem of having a non-representative control group is to have a relatively large comparison sample; to measure the abilities of both samples on a number of important dimensions, and to control for any pre-existing differences between groups statistically. In our present sample we did this by performing analyses of covariance. The problem with this method, however, is that it assumes certain statistical properties of the data (see also Jarrold & Brock, 2004). One issue is that if there is a substantial difference between groups on the covariate in question, observed means might be adjusted for spurious reasons (i.e., due to the aggregation of data points into “group” clusters without there necessarily being a relationship between covariate and the factor to be explained at the level of individuals in each group – see Robinson, 1950). Another problem is that for an analysis of covariance to be meaningful, both groups have to show the same pattern of relationship between covariate and performance. These issues highlight that using ANCOVAs to control for differences between groups is not the perfect method. However, we only used ANCOVAs in the case of tasks where we found a difference between groups using a simple ANOVA analysis. This way the risk we took was to disregard some meaningful differences between groups (that is, the method was overly conservative). On the other hand, we were not taking the risk of finding differences between groups due to chance (i.e., as an artefact of our statistical analyses). That is, using this method increases the chance of Type 1, but not Type 2 errors.

A final issue is the question of how informative our results with autistic children really are for dual-process theorists. In a number of influential publications, Karmiloff-Smith and colleagues (e.g., Karmiloff-Smith, Scerif & Ansari, 2003; Thomas & Karmiloff-



Smith, 2002) criticized the idea that certain “cognitive modules” can be selectively impaired, and double dissociations might occur in developmental disorders. This is because although there are specialised areas or cognitive modules in the adult brain, these emerge from a developmental process of modularization. This process is based on some innately specified starting points, but at the early stages of development, these are only domain-relevant, rather than domain specific. The modularization process unfolds during development, and also as a result of specific environmental interactions. In the case of individuals with developmental disorders the effects of genetic mutation during early development are likely to be widespread across the developing system. Instead of neatly segregated patterns of intact and impaired cognitive modules, it is more likely that all domains develop atypically, although some domains may be more affected than others due to their particular problem space. As a result of the dynamic nature of development (for example, the changes in the regulation of gene expression over time), and the effects of environmental interactions, it is also likely that the patterns of strengths and weaknesses within the same developmental disorder change over time. That is, based on this argument it is possible that although there is no evidence for separate Type 1 and Type 2 reasoning systems in developmental samples, these might be present in the case of adults.

Additionally, finding no evidence for the claims of dual-process theories in the autistic sample certainly cannot be interpreted as evidence against the existence of two systems/two types of process in the case of the general population.

What does all this tell us about dual-process theories? In Chapter 3 we described three studies that showed that giving responses that are traditionally considered to be heuristic is effortful for young children, and in some cases even for adolescents as well. In the present studies we found that for the autistic group, committing the conjunction fallacy was positively correlated with measures of cognitive capacity in both experiments. Additionally, in Experiment 4 committing the if-only fallacy was positively correlated with WISC scores, and in Experiment 5 susceptibility to the equiprobability bias was positively

correlated with autistic children's cognitive capacity. At the same time there was little evidence for a negative relationship between susceptibility to biases and cognitive capacity or analogical reasoning ability. These results clearly contradict the predictions of dual-process theories, at least the versions advocated by Evans (e.g., 2006a) and Stanovich and West (e.g., 2008). Although developmental dual-process theories (such as that of Klaczynski, 2009), anticipate an increase in heuristic processing with age, they also do not account for the positive correlation between cognitive capacity and "heuristic" responding that we found in many cases. Thus, dual-process theories do not describe the reasoning performance of children and adolescents in an adequate way. However, it is still possible that dual-process theories are applicable in the case of adults. Also, one of the latest conceptualisations of the theory (Stanovich & West, 2008) places a great emphasis on the role of relevant knowledge in reasoning performance. In order to test whether the theory applies to adults, and how reasoning performance is affected by knowledge, in Chapter 6 we are going to describe three experiments which were conducted with university students at different stages of their education. The students also differed in the discipline that they studied and the country where they lived. In the next chapter we are also going to review the educational approaches to heuristic reasoning.

## Chapter 6: The role of education and knowledge in heuristic reasoning (Experiments 6-8)

### *6.0. Introduction to Chapter 6.*

In Chapter 3 we described three studies which showed that giving responses that are traditionally considered to be heuristic is effortful for young children, and in some cases even for adolescents as well. In Chapter 5 we additionally found that producing a number of different heuristics required effortful reasoning in the case of autistic participants. Although we did not find this in the case of all problems, there was also no evidence for a negative relationship between susceptibility to biases and cognitive capacity. These results are inconsistent with the predictions of dual-process theories, at least in relation to the versions advocated by Evans (e.g., 2006a) and Stanovich and West (e.g., 2008). Although developmental dual-process theories (such as that of Klaczynski, 2009), anticipate an increase in heuristic processing with age, they also do not account for the positive correlation between cognitive capacity and “heuristic” responding that we found in the case of both typically developing and autistic children. Overall it seems that dual-process theories cannot predict children’s and adolescents reasoning performance on some of the most popular tasks in the heuristics and biases literature. However, it is still possible that dual-process theories are able to account for the reasoning performance of adults.

The main issue that we addressed in the present chapter is the role of knowledge (and education) in reasoning performance. This question was examined by comparing the performance of students at different stages of their education, and students studying different disciplines at university level, and also through comparing the performance of students in the UK and at an Italian university. Stanovich and West (2008) have suggested that cognitive capacity and thinking styles will only be reliable indicators of reasoning capacity if relevant knowledge is available (see more on this in Chapter 1). That is, one

might expect education to decrease heuristic use. By contrast, educational theorists (e.g., Fischbein, 1997) propose that education can actually increase heuristic use in some cases. Because of these conflicting predictions, and because the role of knowledge and education in heuristic use has been largely ignored in reasoning research so far (apart from some studies on the effect of training) it seems an interesting question to look at. A corollary of the idea that education can increase heuristic use is that some heuristics might appear at later stages of development, well beyond childhood.

In Chapter 3 we presented evidence which suggested that rather than being basic and effortless, reasoning heuristics appear during mid-childhood, and they are related to effortful contextualisation processes (thus, they are initially produced by Type 2, rather than Type 1 processes – see Experiment 1). At a later stage (i.e., during adolescence – see Experiments 2 and 3) some of these heuristics become effortless, and they can even be suppressed by effortful reasoning. However, there are also some heuristics which continue to require conscious effort, or which seem to emerge or increase during this stage (such as the conjunction fallacy). This also results in inconsistent correlational patterns between heuristics (with both positive and negative correlations, or no relationship between heuristic reasoning across certain tasks). Finally, the studies involving adolescents with autism (Experiments 4 and 5) suggest that producing heuristic responses sometimes requires conscious effort for this group, or they fail to give certain heuristic responses. Overall, these results suggest that most (or possibly all) reasoning heuristics require Type 2 contributions, at least initially. If reasoning plays a role in the development of heuristics, it is also possible that explicit knowledge can contribute to the development of certain heuristics. Moreover, this process should not be restricted to childhood or adolescence. Instead it should depend on the time when participants acquire the new knowledge that leads to the development of a heuristic. In order to test this prediction, the following three experiments were conducted with university students. In this chapter we also review the

literature on the role of education and knowledge in reasoning heuristics (both as it is described by dual-process and educational theorists).

There are a number of reasons to consider the claims and predictions of educational theorists alongside the ideas of dual-process theorists. One reason is that educational theorists have been interested in the role of intuitions in reasoning for a long time (see e.g., Fischbein, 1987, or our review of the intuitive rules approach in Section 2.1.4.). Moreover, educational theorists are also familiar with the inconsistent developmental patterns of heuristics (that heuristics can increase, decrease or remain stable with age and education), and they also acknowledge the inconsistent relationships between cognitive capacity and heuristic use (see e.g., Osman & Stavy, 2006). In the next section we introduce the ideas of educational theorists regarding probabilistic reasoning, and contrast their predictions with the claims of dual-process theorists. Then we describe three experiments which were designed to test these predictions.

### *6.1. Heuristics in probabilistic reasoning and their relationship with cognitive ability, thinking styles and educational background.*

Probabilistic reasoning consists of drawing conclusions about the likelihood of events based on available information, or personal knowledge and beliefs. The heuristics and biases literature (e.g., Gilovich, Griffin & Kahneman, 2002) mostly describes problems where a rule of probability is disregarded in favour of making decisions on the basis of vivid information or personal beliefs. Educational theorists have also been aware for a long time that the notion of probability is notoriously hard to grasp. One reason for this is that the concept of probability incorporates two seemingly contradictory ideas: that the individual outcomes of events are unpredictable. However, on the long run there is a regular pattern of outcomes. According to educational theorists (e.g., Metz, 1998) a failure to integrate these two aspects of probability (or randomness) leads to either interpreting

probabilistic events in an overly deterministic way, or to disregarding the pattern and focusing entirely on the uncertainty aspect. That is, people either overrate or underrate the information given. Depending on whether people under- or overrate the unpredictability of probabilistic events, educational theorists discriminate between two main categories of bias: errors based on the representativeness heuristic (which are the result of overly deterministic judgements), and errors based on the equiprobability bias (which stem from disregarding deterministic information).

Educational theorists have been interested in the role of intuitions in probabilistic reasoning for a long time. For example, according to Fischbein (1975) intuition plays a very important role in the domain of probabilistic reasoning, probably more so than in other domains of mathematics. He defined the concept of intuition as self-evident, holistic cognitions that appear to be true without the need for any formal or empirical proof. Fischbein distinguished between two sources of intuition: primary intuitions which are based on individual experiences, especially on interactions with, and on adaptations to the environment. Fischbein (1987) also claimed that these intuitions do not disappear with education, but they continue to influence judgment, even after formal instruction in a particular area. So far these ideas are very similar to the concept of Type I processes. However, in Fischbein's view there are also secondary intuitions, which are formed by scientific education at the school, and which are partly independent from cognitive development in general.

Developing useful primary intuitions about probability is not easy. Although people have lots of experience with situations involving chance in their everyday lives, these experiences are in general quite "messy". In contrast to arithmetic where  $2 + 2$  yields the easily testable result of 4, the outcomes of probabilistic events are much harder to evaluate. For example, the low chances of winning at the national lottery are in apparent contrast with the fact that people win every week (Borovcnik & Bentz, 1991). At the level of

personal experiences, “bad” decisions (in terms of probabilities) are sometimes followed by positive outcomes, or vice versa. For example, high risk gambles can occasionally result in big prizes, whereas low risk gambles can result in losses. Of course, it is counterintuitive (or even culturally unacceptable) to evaluate decisions irrespective of their consequences. It is even more unusual for people to imagine a series of equivalent events happening many times, so that they can see some sort of a pattern emerging on the long run (Borovsniak & Peard, 1996). As a result of probabilistic events being hard to conceptualise, people’s everyday experiences can actually lead to inappropriate intuitions about the nature of probability.

The heuristics and biases literature (e.g., Gilovich, Griffin & Kahneman, 2002; see also Chapter 1) describes the representativeness heuristic as a tendency for people to base their judgement of the probability of a particular event on how much it represents the essential features of the parent population or of its generating process. The representativeness heuristic often manifests itself in the belief that small samples will “look” exactly the same, and they will also contain the same proportion of outcomes, as the parent population. That is, when relying on the representativeness heuristic, people put too much confidence in small samples. An example for this is the *gambler’s fallacy* (see also Section 1.1.1.). For example, when tossing a fair coin, after a series of heads people have the feeling that a tails should follow, because this corresponds more to their expectation of having a mix of heads and tails, rather than a long sequence of just heads.

The representativeness heuristic is also at work when we base our probabilistic judgments about people on how much they resemble a prototypical member of a certain category. For instance the use of social stereotypes often leads to the neglect of relevant statistical rules and base rates. Demonstrations of this heuristic involve the *conjunction fallacy* (Tversky & Kahneman, 1983), and the *engineers and lawyers problem with a representative description* (Kahneman & Tversky, 1973 – see also Section 1.1.1.).

In contrast to the representativeness heuristic which is based on the content of problems, and leads to overly deterministic judgments, the other bias that we looked at in the present studies is based on the structure of probability problems, and it has the opposite effect: people focus entirely on the uncertainty and unpredictability aspect of probabilistic events, and disregard the patterns in the outcome. The equiprobability bias (which is very similar to the same A-same B rule, in the intuitive rules program, see Section 2.1.4., but it is restricted to cases when people reason about random outcomes) was first described by Lecoutre (1985). It is a tendency for individuals to think of random events as “equiprobable” by nature, and to judge outcomes that occur with different probabilities as equally likely. This bias emerges as the result of formal education in probability (thus, in Fischbein’s terms it is a secondary intuition), and it is based on a misunderstanding of the concept of randomness. An increase in equiprobability bias in the course of education was found in the case of both secondary school and university students (see e.g., Batanero, Serrano & Garfield, 1996; Lecoutre, 1985).

Equiprobability responses are usually based on the following (incorrect) argument: the results to compare are equiprobable, because random events are equiprobable “by nature”. This sort of reasoning is especially common when there is no single effect that could account for the outcome, or the effect is hard to identify and conceptualise (Callaert, 2004). For example, when playing at the national lottery, people do not believe that they have a 50% chance of winning the jackpot, although there is randomness involved, and there are only two possible outcomes for any person (ie., they either win or they do not win). By contrast, when students have to predict the sum of two dice, they often declare that no total is harder or easier to obtain than any other, because the dice are individually fair, and they cannot be controlled (Pratt, 2000). Another example for this bias is the *engineers and lawyers problem with a non-representative description* (Kahneman &



Tversky, 1973 – see also Section 1.1.1.). In line with the findings of educational theorists, in Experiment 3 we found that this bias increased with age in an adolescent population.

The theory of naïve probability (Johnson-Laird et al., 1999) offers a different explanation as to the origin of incorrect equiprobability responses. According to this model, individuals who are unfamiliar with the probability calculus construct mental models of what is true in the various possibilities, and each model represents an equiprobable alternative unless individuals have knowledge or beliefs to the contrary, in which case they will assign different probabilities to different models. Thus, in this approach equiprobability is assumed by default, unless individuals have beliefs to the contrary, and this effect is independent of education in statistics, or should actually decrease with education.

Fischbein and Schnarch (1997) investigated the relationship between education, age and probabilistic reasoning ability using seven different tasks. Participants were secondary school children between the age of 10 and 17, and college students specialising in mathematics. Similarly to the findings of Experiments 2 and 3, the results were mixed, with some misconceptions increasing, some decreasing, and some remaining stable across age groups. In general, representativeness-based responses (the representativeness heuristic in a lottery scenario, the negative recency effect, and the conjunction fallacy) decreased with age. However, other misconceptions, such as the neglect of sample sizes (see Section 1.1.1.), the availability heuristic (see e.g., Section 1.3.3.) and the time-axis fallacy (the erroneous belief that the knowledge of an event's outcome cannot be used to determine the probability of a previous event, because later events cannot retrospectively affect earlier events) increased with age, except in the case of college students who generally gave the correct response. The equiprobability bias was stable across ages in one scenario (rolling two dice and getting 5 and 6 vs. getting two 6s) and increased in the case of another

(neglecting sample size when predicting the probability of an unusual ratio of girls and boys being born the same day in a small versus in a big hospital).

According to Fischbein and Schnarch (1997) these results are based on the interplay between students' intuitions and the structure of the problems. In general, the impact of intuitions increases with age and with education. When a problem is easy to conceptualise, and the relevant rule is readily available, misconceptions will diminish with age. However, if the task is hard to conceptualise, the effects of misconceptions activated by some irrelevant aspect of the task will increase. Thus, according to Fischbein, the context of the tasks, personal experiences and education all have an effect on reasoning about probabilistic events.

As we already discussed earlier (e.g., in Chapter 1), dual-process theories of reasoning (e.g., Evans & Over, 1996) also discriminate between the effect of heuristics (i.e., intuitions) and the ability to conceptualise problems. An idea shared by all dual-process theories is that heuristics are automatically activated by certain aspects of the problem content (the representativeness heuristic is activated by a stereotypical description of a person, for example). Once a heuristic response is activated, it depends on a person's cognitive capacity and personal motivation, whether they override this initial intuition by conscious, effortful (i.e., Type 2) reasoning, and give a response based on the logical structure of the problem, or if they go with their initial intuitions (cued by Type 1 processes). For an overview of supporting evidence for dual-process theories see Section 1.2.2.

Recently, Stanovich and West (2008) proposed that people with higher cognitive ability might not always reason more normatively than lower ability people (see also Section 1.3.1.). This is usually because they do not possess the necessary "mindware" (i.e., relevant knowledge) to solve the problem. It is also possible that they know the relevant rule, but they do not recognise the need to apply it. People with higher need for cognition

(i.e., people who are inclined to think harder about problems) are more likely to recognise the need to apply the appropriate rule. Moreover, instructing people to think hard/logically about problems also sensitises them to potential conflicts between logic and intuitions (e.g., Epstein, Lipson, Holstein & Huh, 1992; Ferreira, Garcia-Marques, Sherman, & Sherman, 2006; Klaczynski, 2001), and this can lead to more normative performance, given that people possess the relevant knowledge to solve the task.

On a theoretical basis (see e.g., Stanovich & West, 2008) one can expect that relevant knowledge plays an important role in reasoning performance. However, participants' knowledge is usually not assessed in empirical studies of heuristic reasoning. On the other hand, there is some evidence that studying different disciplines at university level affects certain aspects of reasoning abilities in distinguishable ways. For example, Lehman, Lempert and Nisbett (1988) found that psychology and medical training increased performance on statistical and conditional reasoning problems (e.g., there was a marked increase in students' relying on the law of large numbers when reasoning about everyday situations involving uncertainty, and their ability to recognize the effect of confounding variables also improved), whereas law students got better only on conditional reasoning during their undergraduate years. On the other hand, chemistry training had no effect on any of the types of reasoning studied. Although there is clearly a self-selection effect (i.e., students with different interests and strengths enrol in different courses) which may explain some of these differences, because of the longitudinal nature of the study (the same students were tested in the first and third year of their course) it was clear that education in a specific discipline also had a distinguishable effect on students' reasoning, depending on their subject of study.

Lehman et al. (1988) proposed that psychology and medical students' statistical reasoning ability increased because these are probabilistic sciences, where statistical reasoning is a very important aspect of education. By contrast, chemistry is a non-

probabilistic (i.e., deterministic) science and law is a non-scientific discipline. They also proposed in line with an earlier study (Fong, Krantz, & Nisbett, 1986) that these changes in statistical reasoning ability not only affect the way students think about their own discipline, but also have an effect on how students of different disciplines reason about everyday-life events.

Heller, Salzstein and Caspe (1992) compared first, second and third year paediatric residents' reasoning about medical and non-medical problems. Some effects that they studied were consistent across training level and problem content. However, first year residents made use of base-rate information more than third year residents, whereas, first year residents were influenced by the way medical data were presented to them, but second and third year residents were not. On the non-medical problems there was no effect of training level. By contrast, higher year residents relied more on representativeness. This was attributed to the fact that medical training consists of practising the recognition of disease conditions and syndromes. Thus, just like the students in Fischbein and Schnarch's (1997) study, they got worse in one aspect of probabilistic reasoning, and got better in another. Unlike the changes identified by Fong et al. (1986) and Lehman et al. (1988) these effects were not general, but specific to medical problems.

The purpose of the following experiments was to investigate the changes in the representativeness heuristic (according to Fischbein, a primary intuition which leads to a deterministic view of probability), and the equiprobability bias (a secondary intuition which results in the overrating of the uncertainty aspect of probability) with statistical education.

Reasoning theorists (e.g., Stanovich & West, 2008) predict that biases in general should decrease with education. By contrast, educational research (e.g., Fischbein & Schnarch, 1997) showed that different heuristics can follow different developmental

patterns, and the overall impact of heuristics on reasoning increases with experience.

However, the use of the representativeness heuristic was found to decrease with education.

According to Batanero et al. (1996) and Lecoutre (1985) the equiprobability bias should increase with statistics education. On the other hand, Johnson-Laird et al. (1999) and Stanovich and West (2008) would predict that the equiprobability bias should decrease with statistics education. Moreover, dual-process theories (e.g., Stanovich & West, 1999) claim that students with higher need for cognition and higher cognitive ability will be less biased, at least when they possess the necessary “mindware” to solve the tasks (Stanovich & West, 2008). In order to test this prediction we included need for cognition as a control measure. Cacioppo and Petty (1982) described the need for cognition as individual differences in the tendency to engage in and enjoy effortful cognitive activity. The need for cognition is positively related to academic performance and course grades (Leone & Dalton, 1988; Sadowski & Gulgoz, 1996). Students with a high need for cognition are able to comprehend material requiring cognitive effort better (Leone & Dalton, 1988), and they are also more effective information processors (Sadowski & Gulgoz, 1996).

As a measure of cognitive abilities we included a short form (Arthur & Day, 1994) of the Raven Advanced Progressive Matrices (APM). The APM is a nonverbal measure of fluid intelligence with a low level of culture-loading which made it appropriate to use with students from different countries (we used students from different countries in these studies in order to test the robustness of our findings across different educational settings).

We wanted to address the question whether we find a change in statistical reasoning ability with education, and whether it is a general effect (as in Fong et al., 1986, and Lehman et al., 1988) or if it is specific to the subject of study (see Heller et al., 1992). To do this, we investigated the effects of problem content. We used problems with both social content (based on the idea that education in psychology might affect the way we reason about people), and we compared students’ performance on these tasks with their

reasoning about tasks with the same underlying logic, but with a less individualised content where the tasks were mostly about objects and natural phenomena.

The main aim of the present investigation was to test the (often conflicting) predictions of reasoning and educational theorists within one study, and to potentially integrate these approaches. In order to separate out the effects of individual differences, the educational system, and education in a certain discipline, we investigated the performance of different groups of students on the same problems. In Experiment 6 we compared the performance of undergraduate psychology (probabilistic science) and marine biology (deterministic science) students from a UK university. The samples included both first year students who have had no education in statistics, and higher year students at different stages of statistics education. In Experiment 7 we examined the performance of Italian psychology students (i.e., students studying the same discipline at a different educational setting) on the same problems. Finally, in Experiment 8 we manipulated the amount of cognitive effort that students invested in solving the problems by instructing them to reason logically vs. intuitively. The aim of this manipulation was to distinguish between reasoning errors that stemmed from a shallow processing of information (i.e., serial associative cognition – see Stanovich, 2009 and Section 1.1.1.), and errors that were the result of the lack of relevant knowledge (i.e., “mindware gaps” – see Stanovich & West, 2008, and Section 1.3.1).

### *6.2. Experiment 6: The role of education, discipline, cognitive abilities and cognitive styles in heuristic reasoning.*

This study investigated the following questions. Is there any effect of psychology education on the representativeness heuristic and on the equiprobability bias? If so, do these biases increase or decrease with education? Are these changes in psychology students' probability judgments only present when they are reasoning about people, or do

they also affect their reasoning about objects and natural phenomena? Are these effects present in both the psychology and the biology student groups, or only in psychology students (as education in probabilistic reasoning is more central to their education)? Finally, in addition to education-related changes, do we also find any effect of individual differences on probabilistic reasoning?

In order to address these questions we used 12 tasks to assess undergraduate psychology and marine biology students' probabilistic reasoning ability (see Appendix E).

## Method

### *Participants*

Participants were 40 first, 40 second and 31 third year psychology students, and 31 first, 21 second and 22 third year biology students from the University of Plymouth. The experiment was run in the first term, so first year students were at the beginning of their statistics course, second year students had a year's education in statistics, and third years had two years of statistics education behind them.

For the purpose of analyses we combined the second and third year groups (group 2 - those students who have had education in statistics), and compared them with the year 1 students (group 1 - no education in statistics). The mean age of the students in group 1 was 19.2 years (age range 18-28 years) for the psychology students ( $n=40$ , 29 females), and 19.7 years (age range 18-31 years) for the biology students ( $n=31$ , 21 females). In group 2 the mean age of the psychology students ( $n=70$ , 64 females) was 22.5 years (age range 19-52 years). The mean age of the biology students ( $n=42$ , 32 females) was 21.2 years (age range 19-37 years). Psychology students took part in the experiment for ungraded course credits, biology students participated as volunteers.

### *Probabilistic reasoning tasks*

We used 12 probabilistic reasoning tasks (see Appendix E for the actual tasks, the instructions, and the normative and heuristic response for each problem). There were 6 different types of task: 3 problems measuring the representativeness heuristic (problems 1, 2 and 4 in each set), and 3 problems measuring the equiprobability bias (problems 3, 5 and 6 in each set).

All the problems were presented twice (once in Set 1, and once in Set 2). The logical structure of the corresponding problems in the two sets was equivalent, but the content of the problems was different. In Set 1 the presentation and content of the problems were similar to probabilistic problems that are usually included in statistics textbooks (everyday context). In these tasks the randomness of the processes involved was more salient (i.e., the processes were related to activities which are known to generate random outcomes, such as throwing a dice, or taking part in a lottery) and the stories were usually about objects (e.g., flipping a coin, drawing a marble, etc.) or natural phenomena (e.g., the weather). Set 2 required probabilistic reasoning about people. These problems also contained information about the probabilities of the different outcomes, but the randomness of the generating process was less salient than in the object content tasks (e.g., when we know that most students with a learning disability are dyslexic, we probably still do not think that the type of learning disability an individual suffers from is randomly assigned to them). These problems resembled the way probabilistic data are usually presented in psychology papers and textbooks.

In each task participants chose from three different response options. One corresponded to the representativeness heuristic or the equiprobability bias, one was the normative response (i.e., a response that corresponds to the statistical rule that is required to solve the task), and the third one neither corresponded to the main misconception nor was normative.



An example for an everyday content task that measured the representativeness heuristic is the following:

*A fair coin is flipped five times, each time landing with tails up; TTTTT. What is the most likely outcome if a coin is flipped a sixth time?*

- a. *Tails*
- b. *Heads* (heuristic response; representativeness)
- c. *Tails and a heads are equally likely.* (normative response)

This task is measuring the *negative recency effect/gambler's fallacy*; to expect heads after a series of tails in order to make the pattern more representative to a random sequence of heads and tails (i.e., the alternation of the two outcomes). Students were given a representativeness point if they chose b) and they were given a normative point if they chose c).

An example for a social content task measuring the equiprobability bias is the following.

*The two most common causes of learning difficulties among university students are dyslexia (specific problems in learning to read, write and spell) and dyscalculia (specific problems in learning arithmetical concepts and procedures). Out of 15 university students with learning difficulties approximately 9 are dyslexic, and 6 have dyscalculia. Joe is a student with a learning difficulty. Which of the following is the most likely?*

- a. *Joe is dyslexic* (normative response)
- b. *Joe has dyscalculia*
- c. *Both are equally likely* (heuristic response; equiprobability)

Students were given an equiprobability point if they chose response c.) and they were given a normative point if they chose response a.).

The problems were presented in a fixed random order, and the order of the problems within both sets was the same. The order of the presentation of the two sets was counterbalanced across participants.

For each problem students were given either a representativeness/equiprobability, a normative or an “other” point. The maximum number of representativeness and equiprobability responses was 3 per set (6 in total), and the maximum number of normative responses was 6 per set (12 in total).

*Measures of individual differences and educational achievement in mathematics and statistics.*

A 12-item short form of the Raven Advanced Progressive Matrices (APM) test (Arthur & Day, 1994) was used as a measure of cognitive capacity which was designed for adolescents and adults with above average intelligence. The students solved 3 practice items taken from the Raven Standard Progressive Matrices (Raven, 1938) before they solved the actual tasks. Participants also filled in the 18-item need for cognition questionnaire (Cacioppo, Petty, Feinstein, & Jarvis, 1996). The Raven test is a measure of fluid intelligence, whereas need for cognition is an indicator of how much an individual is inclined to engage in effortful cognitive processing. These measures are independent (i.e., uncorrelated) but both are known to be related to individual differences in the tendency to avoid relying on automatic, heuristic processing. In this study we used the combined  $z$  scores of these two measures as an indicator of participants' *normative potential*, that is, a person's ability and propensity to engage in effortful reasoning. According to dual-process theories (see e.g., Kokis et al., 2002) this should be a good predictor of normative reasoning.

Students had to choose their highest qualification in mathematics from a list (see Appendix E). Second and third year students also had to report their marks on the previous

year's statistics module. Although both psychology and biology students have education in statistics, and they learn about the same statistical tests ( $t$  tests, ANOVA, Chi-square test, etc.), statistics and methodology is not as central to biology as to psychology education. Biology students have fewer lectures in a year (e.g., in the first year they have 5 lectures a year as opposed to 20 statistics lectures in the case of the psychology students). Additionally, psychology students, but not biology students, have education in critical thinking (10 lectures in the first year, where, among other things, they learn about reasoning heuristics and Bayes' Theorem).

### *Procedure*

Psychology students were tested in groups of 5-10, working on the problems individually. The tasks were presented in a paper and pencil version, and they were given in the same order for each participant (apart from the two sets of probabilistic reasoning problems which were presented in a counterbalanced order). The students worked through the probabilistic reasoning problems first, then they filled in the need for cognition questionnaire, and finally they solved the APM. The session took about 25 minutes.

Biology students solved the problems in their year groups (groups of 20-30). The procedure was the same, except that they did not do the APM (the reason for leaving out the APM was to reduce the time of test administration, as the students took part in the experiment during their tutorial time). The sessions took about 15 minutes.

### **Results**

In our first analysis we examined the individual differences measures across subject and year groups, in order to check for any pre-existing differences. A 2x2 ANOVA with subject of study (psychology/biology) and year group (first year/higher year) as between subjects variables showed a main effect of subject of study ( $F(1, 179)=6.39, p<.05, \eta_p^2=.03$ ) on need for cognition, but no effect of year group and no interaction between subject

of study and year group. This indicated that biology students had significantly higher need for cognition scores than psychology students. However, there was no difference between first and higher year students within subject groups. An independent samples *t* test indicated that there was no difference between first and higher year psychology student groups in terms of cognitive ability ( $t(108)=.65$ , n.s.). We also compared the students' highest qualification in mathematics before starting their university course. A chi-square test indicated that the distribution of qualifications in mathematics was not significantly different between subject groups ( $\chi^2(2, n=183) = 3.75$ ).

We then computed reliability scores for our measures of reasoning heuristics. The Cronbach's alpha for the representativeness tasks was .51, and for the equiprobability tasks was .59. A correlation of over  $r=.50$ , is considered a satisfactory level of reliability for group measurement (Rust & Golombok, 1999). Performance on the two scales was uncorrelated ( $r(181)=-.001$ ).

The next set of analyses concerned the number of representativeness-based, equiprobability and normative responses across contents (everyday/psychology) and groups (see Tables 6.1. and 6.2. for the descriptive statistics in percentages, see Appendix S6 Table 1 for full results). The purpose of these analyses was to find out whether statistical reasoning changed with education, whether these changes were present in general or they were restricted to the field of study, and whether students studying different disciplines showed the same pattern of changes. In order to investigate these questions we ran a series of 2x2x2 mixed ANOVAs with content (everyday/psychology) as a within-subjects factor, and subject group (psychology/biology) and year group (first year / higher year) as between-subjects factors on the representativeness, equiprobability and normative responses. As we found a difference in need for cognition between subject groups, and we expected this to have an effect on students' performance on the tasks, we included need for cognition as a covariate in the analyses.

Table 6.1 . *The means and standard deviations (in brackets) of the representativeness and equiprobability responses across groups and problem types as a percentage of all responses.*

group	Representativeness		Equiprobability	
	Everyday	Psychology	Everyday	Psychology
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
First year psychology	25 (16)	29 (19)	9 (15)	23 (25)
Higher year psychology	15 (18)	13 (19)	23 (29)	30 (32)
First year biology	17 (16)	20 (21)	9 (15)	27 (29)
Higher year biology	22 (20)	20 (20)	11 (23)	20 (22)

The ANOVA indicated a main effect of year group ( $F(1, 178)=6.39, p<.05, \eta_p^2=.04$ ) on the representativeness responses. There was no effect of content and subject group. However, there was a significant subject group by year group interaction ( $F(1, 178)=11.61, p<.01, \eta_p^2=.06$ ). This indicated that representativeness responses decreased with education, but only in the psychology student group. Need for cognition also had a significant effect ( $F(1, 178)=4.12, p<.05, \eta_p^2=.02$ ).

In the case of the equiprobability responses the ANOVA showed a significant subject group by year group interaction ( $F(1, 179)= 4.36, p=.05, \eta_p^2=.02$ ), indicating that higher year psychology students gave more equiprobability responses than first year psychology students. There was no difference between year groups within the biology student group. Need for cognition also had a significant effect ( $F(1, 178)=6.95, p<.01, \eta_p^2=.04$ ).

The ANOVA on the normative responses indicated a significant effect of need for cognition ( $F(1, 178)=13.24, p<.01, \eta_p^2=.07$ ). No other effects or interactions reached significance.

Table 6.2. The means and standard deviations (in brackets) of the normative responses across groups and contents as a percentage of all responses.

group	Normative responses	
	Everyday <i>M (SD)</i>	Psychology <i>M (SD)</i>
First year psychology	80 (15)	71 (17)
Higher year psychology	79 (19)	76 (20)
First year biology	84 (13)	73 (18)
Higher year biology	83 (16)	77 (16)

In order to see if individual differences (the students' *normative potential*) and the students' achievement in statistics had an effect on the number of representativeness and equiprobability responses that they gave, we ran a series of regression analyses (see also Appendix S6 Table 2.). Based on Stanovich and West (2008) we expected that students higher in *normative potential* would give more normative responses, given that they possess the necessary mindware to solve the task (which is more likely in the case of higher year students, and students with a more intensive training in statistics).

In the first year psychology student group there was no relationship between their *normative potential* and their giving heuristic responses. In the higher year psychology student group statistics marks alone were significant negative predictors of both the representativeness responses ( $\beta=-.25, p<.05$ ) and the equiprobability responses ( $\beta=-.32, p<.01$ ). When we also entered *normative potential* into the equation, after controlling for

the effect of statistics marks, we found that it was a significant negative predictor of giving equiprobability responses ( $\beta=-.27, p<.05$ ).

In the biology student group neither statistics marks, nor need for cognition was a significant predictor of giving heuristic responses in the first or in the higher year student group.

### Discussion

In this study we compared the performance of first year and higher year psychology and marine biology students from the same university on a number of probabilistic reasoning problems. The results indicated distinctive patterns of relationship between statistics education and the two reasoning biases, and we also found a difference between subject groups (in line with Heller et al., 1988). Amongst psychology students, representativeness responses decreased, and equiprobability responses increased with education. In the case of the biology student group there was no difference between year groups in the number of representativeness and equiprobability responses that they gave.

There was also no relationship between biology students' statistics marks or their need for cognition, and their probabilistic reasoning performance. This indicates that their education in statistics and biology did not have an effect on reasoning about the problems used in this study. Although dual-process theories of reasoning (e.g., Stanovich & West, 1999) predict a relationship between thinking styles and the use of reasoning heuristics, more recently Stanovich and West (2008) proposed that this relationship only holds if an individual has the necessary "mindware" (i.e., knowledge of relevant rules and procedures which lead to the normative response). This is a possible explanation for the lack of relationship between heuristic responding and need for cognition in the biology student group, and also amongst the first year psychology students.

In the psychology student group participants gave fewer representativeness and more equiprobability responses with education. These changes were present in the case of

both the psychology and the everyday content tasks. This result replicates the findings of other studies that found simultaneous increases and decreases in the use of different heuristics within the same group of students (e.g., Fischbein & Schnarch, 1997, Heller et al., 1992).

The fact that there was no increase in normative responding in either subject group is in line with earlier findings by Kahneman, Slovic and Tversky (1982) who compared heuristic-based errors made by those with and without training in statistics. They only found small and infrequent differences between the two groups.

In order to see if this pattern of increase in the equiprobability bias and decrease in the representativeness heuristic is generally present in psychology students we replicated the experiment with a group of Italian psychology students. As there was no effect of education in the case of biology students we have not included this comparison group in Experiment 7.

### *6.3. Experiment 7. The replication of Experiment 6 with Italian psychology students.*

This study was a partial replication of Experiment 6 with a group of Italian psychology students. The students in this sample had similar education in statistics as the psychology students in the UK, but they did not learn critical thinking. As in Experiment 6 first year students with no previous education in statistics (group 1) were compared with students from the same course who already attended statistics classes (group 2). We also investigated the individual differences (cognitive abilities and cognitive styles) correlates of the students' use of reasoning heuristics.



## Method

### *Participants*

Participants were 129 psychology students from the University of Florence. Two groups were created: group 1 ( $n=47$ , 33 females; students who had not done any statistics courses yet), and group 2 ( $n=82$ , 36 females; students who had had education in statistics already). The mean age of the students in group 1 was 21.3 years (age range 18-32 years). In group 2 the mean age was 24.8 years (age range 19-45 years). They took part in the experiment as volunteers.

### *Materials and procedure*

Students worked through the same 12 probabilistic reasoning problems that were used in Experiment 6. They also filled in the Italian version of the need for cognition questionnaire (Chiesi & Primi, 2008), and the short form of the APM. As in Experiment 6, we used the combined  $z$  scores of the two measures as an indicator of students' *normative potential*.

Students were also asked to report their high school final grade in mathematics. Higher year students also had to report their marks on the previous year's statistics module.

Participants were tested in groups of 5-10 (working on the tasks individually). They started with the two sets of probabilistic reasoning problems (the order of sets was counterbalanced across participants). Then they filled in the need for cognition questionnaire, and finally they solved the APM. The session took about 25 minutes.

## Results

The first analyses concerned the individual differences measures (APM scores, and need for cognition). The purpose of these analyses was to confirm that there were no pre-existing differences between groups. Independent samples  $t$  tests indicated that there was no difference between groups in terms of need for cognition ( $t(127) = -.63$ , n.s.) or

cognitive ability ( $t(126)=1.26$ , n.s.), or high school final grades in mathematics ( $t(122)=1.18$ , n.s.).

The next analyses concerned the differences in probabilistic reasoning ability between first and higher year students, and we also investigated the effects of task content. Based on the findings of Experiment 6, we expected that educational level would affect reasoning performance. We ran a series of 2x2 mixed ANOVAs with content (everyday / psychology) as a within-subjects factor, and group as a between-subjects factor on the representativeness, equiprobability and normative responses (see Tables 6.3. and 6.4. for descriptive statistics, a supplementary table for these analyses is Appendix S7 Table 1.). In the case of the representativeness responses, the ANOVA indicated no effect of group or content, nor a content by group interaction ( $F_s < 1$ ). The results showed that representativeness responses remained stable across year groups.

Table 6.3. *The means and standard deviations (in brackets) of the representativeness and equiprobability responses across groups and problem types as a percentage of all responses.*

group	Representativeness		Equiprobability	
	Everyday	Psychology	Everyday	Psychology
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
First year psychology	21 (18)	23 (22)	7 (13)	17 (18)
Higher year psychology	20 (18)	21 (19)	10 (16)	24 (22)

In the case of the equiprobability responses the ANOVA showed a main effect of content ( $F(1, 126)= 27.65$ ,  $p < .01$ ,  $\eta_p^2 = .18$ ), and a main effect of group ( $F(1, 126)= 5.63$ ,  $p < .05$ ,  $\eta_p^2 = .04$ ). There was no interaction between content and group. Thus, students gave more equiprobability responses to the psychology content tasks than to the everyday

content tasks, and, as in Experiment 6, the number of equiprobability responses increased with education.

Moreover, the ANOVA indicated that normative responding did not change with education, but there was a main effect of problem content ( $F(1, 124)= 14.01, p<.001, \eta_p^2 =.10$ ). There was no interaction between content and group. This indicated that students gave fewer normative responses on the psychology content tasks, and correct responding did not change with education.

Table 6.4. *The means and standard deviations (in brackets) of the normative responses across groups and problem types as a percentage of all responses.*

group	Normative	
	Everyday <i>M (SD)</i>	Psychology <i>M (SD)</i>
First year psychology	74 (14)	70 (13)
Higher year psychology	75 (13)	67 (15)

We also examined the relationship between students' *normative potential*, their statistics marks and the number of representativeness and equiprobability responses that they gave (see also Appendix S7 Table 2). We expected that normative potential would only be related to normative responding in the case of students who possess the relevant mindware to solve the tasks (i.e., higher year students). In fact, in the first year student group *normative potential* did not significantly predict their giving representativeness or equiprobability responses. Amongst higher year students there was no relationship between statistics marks and heuristic responding. However, students' *normative potential* was negatively related to their susceptibility to the equiprobability bias ( $\beta=-.24, p<.05$ ).

## Discussion

Experiment 7 was a partial replication of Experiment 6 with Italian psychology students as participants. As with the British psychology students in Experiment 6, the equiprobability bias increased with education. We also found that students gave more equiprobability responses to the psychology content tasks, which indicates that problem content might be important in the case of the equiprobability bias. Specifically, when the task is about people who have individual properties (as opposed to objects and natural phenomena, which are interchangeable with each other) students can “justify” their intuition of equiprobability by arguing (similarly to Gigerenzer et al., 1999) that base rates are irrelevant when making predictions about a single, individual outcome, disregarding the fact that in the present tasks individuals were randomly selected from a sample with known probabilities of the different possible outcomes.

Johnson-Laird et al. (1999) suggested that individuals who are unfamiliar with the probability calculus will construct mental models of all possible outcomes and they assume by default that these outcomes are equiprobable, unless they have knowledge or beliefs to the contrary. Experiments 6 and 7 showed that students without education in statistics were to some extent susceptible to the equiprobability bias (they gave erroneous equiprobability responses about 10-20 % of the time). However, this susceptibility increased significantly with education. On the other hand, the bias was negatively correlated with students’ *normative potential*. Stanovich (2008) used the term *serial associative cognition with a focal bias* to describe a reasoning process which is not purely associative and effortless (that is, not completely heuristic), but consists of the justification of a heuristically cued response through conscious, but low-effort processing (basically, a rationalization process) which lacks a real attempt to consider alternatives. Instead it is looking for easy ways of justifying the most obvious answer (a similar idea was proposed by Evans, 2006a, which he called the *satisficing principle* – see Section 1.3.1). It is possible that students with a

poor understanding of the concept of randomness use it to justify their initial intuitions about the equiprobability of different outcomes which results in the counterintuitive increase in equiprobability responses with education.

According to previous studies (e.g., Epstein, et al., 1992; Ferreira, et al. 2006; Klaczynski, 2001; see also Section 1.2.2.) instructing participants to reason analytically or logically (e.g., “take the point of view of a perfectly logical person” - Denes-Raj & Epstein, 1994) increases the mental effort that they invest in the task, and as a result, participants give more normative responses. If the equiprobability bias is partly the result of shallow, low-effort processing, we should be able to reduce the bias by giving logical instructions, which encourages a more thorough and effortful processing.

Going back to the results of Experiment 7, we found (similarly to the biology student group in Experiment 6) that the representativeness heuristic did not diminish as a result of education in statistics. It is possible that the decrease in the representativeness heuristic in the case of the British psychology students was not related to their education in statistics, but their education in critical thinking, where they learnt about the conjunction fallacy and other classic heuristics and biases tasks. This is also confirmed by the fact that in this group the representativeness heuristic was negatively correlated with their marks on the statistics and critical thinking module, whereas biology students' and Italian psychology students' achievement on their statistics courses was unrelated to their using the representativeness heuristic. Stanovich and West (2008) emphasised the role of the knowledge of relevant normative rules as a prerequisite of resisting reasoning heuristics. Obviously, if we are unaware that the response we give is based on some mistaken intuition we will not be able to correct it through conscious effort. Thus, increasing the mental effort that we invest into solving a problem will not have an effect on these types of error.

On the basis of the above we predicted that representativeness responses would not be affected by instructing students to think logically, whereas the equiprobability bias would be reduced by logical instructions, as compared to the intuitive condition.

*6.4. Experiment 8: The effect of instructions on susceptibility to the representativeness heuristic and to the equiprobability bias.*

Drawing on the findings of Experiments 6 and 7 we proposed that the representativeness heuristic and the equiprobability bias might be based on different reasoning processes (i.e., the former being the result of pure heuristic processing, whereas the latter resulting from an initial heuristic representation, followed by a low-effort, conscious justification phase). Although, as we have seen in Experiment 6, students who are explicitly taught about the representativeness heuristic (i.e., who possess the necessary “mindware”) are able to resist this response. However, without explicit tutoring, students continue to commit this fallacy, at least in the case of tasks where they do not recognize the need to employ a relevant normative principle. This idea is consistent with Stanovich and West (2008) who reported that a number of cognitive biases were independent of cognitive ability. They explained these findings by proposing that people with high cognitive ability often do not recognize the need for a normative principle any better than people with lower cognitive ability. On the other hand, even when a person possesses the necessary knowledge, they can still come to an erroneous conclusion. As Evans (2006b) noted, just because a reasoning process is conscious and capacity demanding it does not necessarily lead to correct, logical reasoning.

One possible way of distinguishing between reasoning errors that arise from the lack of the necessary “mindware”, and those that are the result of effortful but sloppy reasoning is to use different instructional conditions. Giving logical instructions sensitises participants to the potential conflict between logic and intuitions (Stanovich & West,

2008), and makes it more likely that if participants have the relevant knowledge, they will use it (see more on the effects of instructions in Section 1.2.2.). However, if the person lacks the knowledge that would lead to the normative solution, instructions should not make a difference.

Finally, based on a recent finding by Evans, Handley, Neilens, Bacon and Over (*in press*) it was hypothesized that people with higher cognitive ability would better comply with instructions. That is, we expected a larger effect of instructions in the case of participants with higher cognitive ability.

## Method

### *Participants*

Participants were 119 psychology students from the University of Florence who already had education in statistics, thus they were expected to possess relevant knowledge of the concept of randomness. The mean age of participants was 24.9 years (SD= 2.9, age range 23-45 years), there were 14 males and 105 females in this group. They took part in the experiment as volunteers.

### *Materials and Procedure*

The same 12 tasks as in Experiments 6 and 7 were administered to measure students' probabilistic reasoning ability. As before, half of the problems had a social content and the other half had an everyday content. The problems were presented in a fixed random order, and the order of the problems within each set was the same. The order of the presentation of the two sets was counterbalanced across participants. Students had to solve the tasks individually.

Two instruction conditions were used. In the "intuitive condition" the experiment was introduced as a study of human intuition. The purpose of this instruction was to encourage heuristic (Type 1) processing. The instructions were the following: "*The present study's goal is to evaluate personal intuition when one has to make choices on the basis of*

*some information. The following pages consist of multiple-choice questions. Please answer the questions on the basis of your intuition and personal sensitivity”.*

In the “logical condition” the experiment was introduced as a study on human rationality. The purpose of this instruction was to elicit logical (Type 2) processing. The instructions for this group were the following: “*The present study’s goal is to evaluate personal reasoning capacity when one has to make choices on the basis of some information. The following pages consist of multiple-choice questions. Please answer on the basis of your rational and reflective thinking. When answering the questions, take the perspective of a perfectly logical person”.*

About half of the participants ( $n= 57$ ) were randomly assigned to the intuitive condition and the other half ( $n= 62$ ) to the logical condition. After solving the tasks, the students also worked through the short form of the APM, and they filled in the need for cognition questionnaire.

## Results

First we compared the two instruction groups on their need for cognition and APM scores. There was no difference between groups in their cognitive abilities ( $M= 8.81$  and  $M=8.52$  for the intuitive and rational groups, respectively;  $t(117)=.79$ , n.s.) and cognitive styles ( $M=65.47$  and  $M=65.97$  for the intuitive and rational groups, respectively,  $t(117)=.30$ , n.s.). Then we computed the  $z$  scores of students’ APM scores and their need for cognition. As in the previous experiments, we used the combined  $z$  scores of these two measures as an estimate of students’ *normative potential*. In order to investigate the effects of *normative potential*, we created a high ( $n=59$ ) and a low ( $n=60$ ) *normative potential* group by dividing the sample at the median. The *normative potential* of the two groups was significantly different ( $t(117)=16.26$ ,  $p<.01$ ).

We ran a series of  $2 \times 2 \times 2$  mixed ANOVAs with content (everyday/psychology) as a within-subjects factor, and instruction and *normative potential* group (high/low) as a



between-subjects factor on the representativeness, equiprobability and normative responses (descriptive statistics are displayed in Tables 6.5. and 6.6., for additional results see Appendix S8 Table 1). The purpose of these analyses was to see whether instructions had an effect on students' giving representativeness-based and equiprobability responses, and if there was any difference between students with high and low normative potential in how they complied with instructions. More specifically, it was expected that students possessing relevant knowledge about randomness should be able to resist the equiprobability bias when instructed to think logically. However, because students did not have training in recognising the representativeness heuristic, we expected that instructions should not affect this bias.

A 2x2x2 ANOVA on the representativeness responses indicated no effect of problem content, instructions or *normative potential*, or any interaction between these.

Table 6.5. *The means and standard deviations (in brackets) of the representativeness and equiprobability responses across normative potential (NP) groups and problem types as a percentage of all responses.*

group	Representativeness		Equiprobability	
	Everyday <i>M (SD)</i>	Psychology <i>M (SD)</i>	Everyday <i>M (SD)</i>	Psychology <i>M (SD)</i>
Intuitive, low NP	12 (16)	19 (17)	12 (19)	20 (23)
Intuitive, high NP	22 (20)	24 (17)	9 (15)	22 (22)
Rational, low NP	23 (22)	20 (17)	3 (14)	26 (23)
Rational, high NP	19 (19)	20 (19)	7 (16)	14 (19)

In the case of the equiprobability responses the ANOVA showed a main effect of content ( $F(1, 114) = 40.72, p < .001, \eta_p^2 = .26$ ), and there was also a significant interaction

between content, instructions and *normative potential* ( $F(1, 114)= 5.89, p<.05, \eta_p^2=.05$ ).

No other effects or interactions were significant.

We analysed this effect further by running a 2x2 ANOVA with content (psychology/everyday) as a within subjects factor, and instructions (intuitive/rational) as a between subjects factor, separately for the high and low *normative potential* groups (see additional results in Appendix S8 Table 2.).

Table 6.6. *The means and standard deviations (in brackets) of the normative responses across normative potential (NP) groups and problem types as a percentage of all responses.*

group	Normative responses	
	Everyday <i>M (SD)</i>	Psychology <i>M (SD)</i>
Intuitive, low NP	82 (14)	71(22)
Intuitive, high NP	83 (14)	72 (15)
Rational, low NP	79 (17)	74 (16)
Rational, high NP	82 (17)	79 (16)

Within the high ability group there was a significant effect of content ( $F(1, 57)=36.58, p<.001, \eta_p^2=.39$ ) and also a significant content by instruction interaction ( $F(1,57)=8.94, p<.01, \eta_p^2=.14$ ). This was because students in the intuitive condition gave more equiprobability responses to the psychology than to the everyday content tasks ( $t(28)=5.63, p<.001$ ), but this difference was significantly smaller between the two contents in the rational condition ( $t(28)=2.54, p<.05$ ) as a result of students' giving fewer equiprobability responses to the psychology content tasks (but not the everyday content ones). The same analysis within the low ability group showed only a significant effect of content ( $F(1, 57)= 11.12, p<.01, \eta_p^2=.16$ ).

The ANOVA on the normative responses (see Table 6.6. for descriptives) indicated that normative responding did not change with instructions, but there was a main effect of problem content ( $F(1, 115) = 17.74, p < .001, \eta_p^2 = .13$ ). Students generally gave more correct responses on the everyday content tasks, and logical instructions did not increase the number of normative responses.

### Discussion

Experiment 8 replicated and extended some of the findings of the previous experiments. As in Experiments 6 and 7, we found that the representativeness heuristic was unrelated to students' *normative potential*. This heuristic was also resistant to the effect of instructions. This supports the idea that students without explicit knowledge about the representativeness heuristic (or at least some typical reasoning errors based on this heuristic) are not very much able to resist this response, regardless of their cognitive abilities.

As in Experiment 7, students gave more equiprobability responses to the psychology than to the everyday content tasks. Throughout Experiments 7 and 8 this effect was very strong. Although we did not find the same effect in Experiment 6, this was because in that experiment we included need for cognition as a covariate in the analysis of responses. Indeed, a simple ANOVA on the equiprobability responses indicated a main effect of content ( $F(1, 179) = 31.34, p < .01, \eta_p^2 = .15$ ). Similarly, in Experiment 8, the content effect in the case of the equiprobability bias only disappeared in the case of high ability students when they were given instructions to think like “a perfectly logical person” (that is, when they were instructed to think hard, or to think like somebody with a high need for cognition).

This suggests that when the equiprobability tasks have a social content, the bias is harder to resist (i.e., it requires more conscious effort). This is probably because we think about people as agents and we attribute higher uncertainty and unpredictability (which is

the basis of the equiprobability bias – see Callaert, 2004) to people than to objects, which we see as more determined. It is likely that when they make inferences about people, students expect individuals to not necessarily behave exactly like other members of their group would (as opposed to a flip of one fair coin that would produce the same result as that of another). Also, as Gigerenzer et al. (1999) noted probabilities do not help much in predicting single outcomes, and making probabilistic judgments about an individual seems like doing exactly this, although the individuals in the tasks were randomly selected from a bigger population, thus students should have recognized the relevance of population base rates. It is also important to note that there was no evidence for a similar content effect in the case of the representativeness heuristic.

In Experiment 8, as in Experiments 6 and 7, the equiprobability bias was more pronounced in the case of students with lower *normative potential*. Together with the fact that this bias increased with education, at least in the psychology student groups, this lends support to the notion that this bias stems from a misapplication of concepts or rules that students learn as part of their education. Based on this argument, we expected that instructions would have an effect on giving equiprobability responses, at least in the case of higher ability students (see Evans et al., *in press*). This is what was found, although it was only apparent on the psychology content tasks. However, given the small number of equiprobability responses on the everyday content tasks, it is possible that the lack of difference on these tasks was the result of a floor effect. This differential effect of instructions on the representativeness heuristic (which was unaffected by instructions) and on the equiprobability bias (which decreased under logical instructions) lends support to the claim that these two biases are based on different types of reasoning processes. This also helps in clarifying the distinction between primary and secondary intuitions, proposed by Fischbein (1987). Primary intuitions seem less accessible to consciousness, and thus we are less able to control them, whereas secondary intuitions are more accessible to

consciousness (and thus more controllable) but only if we engage in effortful cognitive processing.

### *6.5. General discussion*

The purpose of Experiments 6-8 was to examine the role of education in heuristic reasoning, and to test the contrasting predictions of educational (Batanero et al., 1996; Callaert, 2004; Fischbein & Schnarch, 1997; Lecoutre, 1985; Pratt, 2000) and reasoning theorists (Evans & Over, 1996; Johnson-Laird et al., 1999; Kahneman & Frederick, 2002; Stanovich & West, 1999) by looking at psychology students' probabilistic reasoning ability. In particular, these studies investigated students' use of the representativeness heuristic and the equiprobability bias. We combined the methods of the educational and reasoning approaches by simultaneously investigating changes with educational level, and the effects of individual differences in cognitive ability and need for cognition. Using an instruction manipulation it was also possible to explore the cognitive processes underlying primary and secondary intuitions.

According to dual-process theories (e.g., Stanovich & West, 2008) individuals with higher cognitive ability, higher need for cognition, and more domain-specific knowledge tend to give more normative responses to reasoning problems. More specific theoretical accounts of probabilistic reasoning assume that education and training will facilitate more accurate representations of probabilistic information. Johnson-Laird et al. (1999), for example, have proposed that the equiprobability bias will be most evident in the case of individuals who are not familiar with the probability calculus and, consequently, represent outcomes as equiprobable mental models.

In contrast, according to educational theorists, heuristics with different origins can develop differently, and some of them increase, rather than decrease with education (e.g., Fischbein & Schnarch, 1997). This literature predicts a counterintuitive increase in the

equiprobability bias as a result of statistics education which can be related to the students' misunderstanding of the concept of randomness (Batanero et al., 1996, Lecoutre, 1985).

Our results supported the claim that different heuristics can show different patterns with education, and that it is possible that one heuristic increases while another one decreases the same time. In Experiment 6 higher year British psychology students gave fewer representativeness responses and more equiprobability responses than first year students. We found the same counterintuitive increase in equiprobability responses in the case of Italian psychology students in Experiment 7. This suggested that (contrary to Johnson-Laird et al.'s, 1999, claim, and in line with Batanero et al., 1996, and Lecoutre, 1985) the equiprobability bias does not diminish with statistics education, rather, it increases in the case of students who misunderstand the role of randomness in determining the outcome of probabilistic events.

We also investigated the effects of problem content. The purpose of this was to see if the observed changes in probabilistic reasoning were specific to the subject of study or if they were general across contents. We found that all changes that we observed were present in both the psychology and in the everyday content tasks. We proposed that the fact that students gave more equiprobability responses when they had to make judgements about probabilistic events concerning people, was likely to be the result of students generally reasoning about people and objects differently. Namely, we argued that students are more inclined to ignore base rates when they reason about people, because they see them as individuals, rather than as interchangeable items which are mere representatives of the probabilistic properties of their parent population.

We also wanted to address the question of whether there are differences in how groups of students who study different subjects think about probabilities (see Heller et al., 1992, Lehman et al., 1988). We only found an effect of statistics education in the case of psychology students. Although biologists also study statistics, in psychology education

statistics is more central (first year psychology students have four times as many lectures in statistics as biology students). It is likely that it is not studying psychology *per se*, but rather the fact that students had more education in statistics that led to the differences between subject groups.

As dual-process theories of reasoning (e.g., Evans & Over, 1996; Kahneman & Frederick, 2002; Stanovich, 1999) predict, students with higher cognitive abilities and with higher need for cognition were less susceptible to biases. At least incorrect equiprobability responses were mostly given by students with lower cognitive abilities. In addition, students with higher need for cognition (Experiment 6) and with higher normative potential and given logical instructions (Experiment 8) were less sensitive to content effects. This suggests that students with higher normative potential are more able to extract the logical structure of problems, regardless of problem content. Nevertheless, in the case of students who did not learn about the representativeness heuristic cognitive abilities were unrelated to correct responding. Overall these findings are in line with the new model of Stanovich and West (2008) with the restriction that although learning about normative principles is usually beneficial, knowledge can also be “dangerous” sometimes (i.e., misunderstanding normative principles can increase the number of incorrect responses).

Instructions, which are considered to constrain the intentional level (by increasing the cognitive effort that students invest into solving tasks, and by sensitising them to the possible conflicts between logic and beliefs) somewhat improved reasoning performance, but only on the tasks measuring the equiprobability bias, and only in the case of higher ability students. In sum, although cognitive abilities, cognitive styles and instructions have an effect on the use of reasoning heuristics, education and relevant knowledge play a very important role too (although sometimes it can be a negative one!)

The main focus of this study was the changes in the representativeness heuristic and the equiprobability bias with education in statistics. Nevertheless, it is worth

emphasizing that normative responding did not increase in any of the groups in this study throughout their undergraduate years. Students in each group solved around 80% of the tasks correctly. This might sound impressive, but given that the tasks only required simple computations and the correct responses were based on basic statistical principles, this means that many students struggle with understanding the basics of probabilistic reasoning. In each group only around 10 % of the students solved all the tasks correctly. This highlights the need for helping students to learn to identify the possible sources of error and bias, and to understand the principles of probabilistic reasoning better.

From the educational point of view, the most important findings of this study were the increase in the equiprobability bias with education, and the fact that this bias was more prominent on psychology (social) content tasks. On the positive side, there is some indication that teaching students about the representativeness heuristic successfully decreased the number of representativeness-based responses in the case of the British psychology students. Although students with high need for cognition and high cognitive capacity tended to give less heuristic responses, in an educational setting it is very important that students with different levels of ability should be equally able to produce normative responses. It is clearly not enough to teach normative statistical rules when many students are unable to recognize the need to apply them, or they actually use them to justify their incorrect intuitions. It is also clear that learning about a certain type of heuristic does not affect the use of another heuristic (e.g., students who learnt to resist the representativeness heuristic continued to give erroneous equiprobability responses).

In summary, the present findings are in line with the predictions of educational theorists: heuristic use can increase or decrease with education, and these patterns can be present simultaneously in the same population. This is similar to our findings with adolescent populations (see Experiments 2 and 3) where the developmental patterns of different heuristics were very mixed. In contrast with the predictions of dual-process



theorists (e.g., Stanovich, 1999) overall normative responding did not increase with education, and giving logical instructions did not have an effect on the number of normative responses. This is again, in line with our previous findings, where cognitive abilities were mostly unrelated to normative responding, and normative responding did not tend to increase with age. On the other hand, students with higher need for cognition and higher cognitive capacity were less susceptible to the equiprobability bias, and we also found evidence that giving logical instructions had a greater effect in the case of higher ability students (which is in line with Evans et al., *in press*). The fact that normative potential was a good predictor of reasoning performance only in the case of students with a higher level of education (who presumably possessed relevant knowledge) in Experiments 6 and 7 is also in agreement with the predictions of Stanovich and West (2008). Overall the present findings confirm some of the ideas of dual-process theories, but also highlight the fact that these predictions can only be applied under certain conditions (e.g., normative potential only predicts performance in the case of participants with relevant knowledge; logical instructions only have an effect on high ability participants' performance, etc.). These results also provide further evidence that heuristics are not always basic, and they can emerge/increase at later stages of development, well beyond childhood.

In the final chapter I am going to discuss further the implications of these findings, and that of the other studies reported in Chapters 3 and 5, for dual-process theories. I will also summarize the most important results of the empirical studies, discuss some issues that still need to be resolved, and (based on the above) outline some possible future directions of investigation.

## Chapter 7: A review and interpretation of the experimental findings, and recommendations for future research.

### *7.0. Introduction to Chapter 7.*

The main aim of this chapter is to review and interpret the experimental findings reported in Chapters 3, 5 and 6, and to relate these findings to the theoretical accounts that I have introduced in the previous chapters (i.e., dual-process theories, the cognitive theories of autism, and educational theories of heuristic reasoning). I also identify some issues that still need to be resolved, especially regarding the cognitive underpinnings of heuristic reasoning, and the development of reasoning heuristics.

Section 7.1. reviews our most important empirical findings regarding the development of heuristics, heuristics and biases in autism, and the role of education in the emergence and suppression of reasoning heuristics. Besides the effects of cognitive capacity and personal experiences, this section also discusses the role of thinking styles, education and instructions in heuristic reasoning. I also discuss the compatibility of our findings with dual-process theories of reasoning, and with some prominent cognitive theories of autism. Another important aim of this section is to try and identify the driving forces behind the development of reasoning heuristics.

Section 7.2. explores the question of whether dual-process theories offer the best explanation for our findings, or whether single process theories (such as the intuitive rules model – see Section 2.1.4.) could account for the present results just as well. In this section I also introduce a recent theoretical paper (Keren & Schul, *in press*) which criticized dual-process theories, and relate the issues that they raised to the findings of our empirical studies. The main point here is that dual-process theories cannot account for many empirical findings, especially regarding the development of reasoning heuristics. Although the methods developed/implemented by dual-process theorists (for example, investigating

the effects of cognitive capacity on reasoning, using instruction manipulations or dual-task paradigms, etc.) are extremely useful in studying reasoning processes, dual-process theorists' claims about the underlying cognitive processes of reasoning still remain controversial. To say the least, it seems that single process theories can explain our empirical findings just as well as dual-process theories.

Finally, Section 7.3. discusses the rationality of heuristics and biases. Although avoiding heuristics is supposed to be rational (see e.g., Stanovich, 2009) sometimes people who could be expected to be more rational are more susceptible to biases. For example, we have found that older children are more susceptible to biases than younger children, typically developing children are (often) more susceptible to biases than children with autism, and people with education in statistics are more susceptible to (a certain) bias than people without education in statistics. In this section I also outline some possible future directions of research, and I summarize my conclusions.

### *7.1. Experimental findings regarding the development of heuristics and biases in autism and in typical development, and the role of education in heuristic reasoning.*

The main aim of this section is to review and interpret the experimental findings reported in Chapters 3, 5 and 6, and to relate these findings to the predictions of dual-process theories and the cognitive theories of autism. I also review our empirical findings regarding the role of education in the emergence and suppression of reasoning heuristics.

#### *7.1.1. A summary of the predictions of dual-process theories regarding the development of heuristic reasoning.*

In Section 1.2.1. I reviewed the evidence from studies with adults that supports the idea of two separate types of processes (i.e., quick, automatic and effortless vs. slow and effortful processes that are demanding of cognitive resources). This evidence includes the

fact that on some problems (notably on conflict problems where Type 1 and Type 2 processes are expected to cue conflicting responses) people who have higher cognitive capacity, people who are dispositionally inclined to reason carefully or who are explicitly instructed to reason rationally are more likely to produce normative responses to problems. In addition, people who are given little time to think, or people whose working memory is burdened with preload are less likely to give normative responses to the same problems. By contrast, performance on non-conflict problems is unaffected by these manipulations. Evidence from experiments using verbal and non-verbal process-tracing methods also indicates that participants' attention is drawn to certain aspects of problems very quickly and they are more likely to process these details more carefully, neglecting other (sometimes crucial) aspects. All of these findings are in line with the predictions of traditional, default-interventionist dual-process theories (e.g., Evans & Over, 1996; Stanovich, 1999 – see also Section 1.2.1.). Although these claims are almost exclusively based on research with adult participants, some dual-process theorists (e.g., Stanovich, Toplak & West, 2008) maintain that all these should equally apply to children as well.

Both Stanovich and West (2008) and Evans (2006a) modified their theories recently to offer a more complex idea of human cognition than their initial conceptualisations of dual processes (see Section 1.3.1.). One idea that both recent models emphasize is that cognitive effort does not always guarantee a normative solution, or even an abstract, logical representation of a task. Hypothetical thinking, and mental simulation of alternative outcomes might not be possible without Type 2 thinking, but Type 2 thinking is also no guarantee for an abstract, logical representation (see Evans, 2006a; Stanovich, 2009). Stanovich and West (2008) also emphasize the role of relevant knowledge in reasoning and they specify the ways in which knowledge, cognitive capacity and cognitive styles interact in shaping reasoning performance (see Figure 1.4. for a summary of the model). In Section 1.3.2. I also summarized some concerns about the claims that the original dual-process

theories of reasoning (i.e., for example, Evans & Over, 1996 and Stanovich, 1999) make. There is evidence, for example, that the belief bias (a prototypical Type 1 process) arises in frontal brain areas (which are normally associated with Type 2 processes – although dual-process theorists do not make strong claims about the localization of Type 1 and 2 processes). Another important observation is that although Type 2 processes *sometimes* override Type 1 reasoning, most of the time Type 1 reasoning is in control of people's behaviour. A final issue is that both Type 1 and Type 2 processes seem to include a number of different components. In the case of Type 1 processes the following components have been identified: implicit learning, overlearned associations, and the modular processes for solving adaptive problems (Stanovich, 2008). Type 2 processes also include some (partially) independent components, such as working memory/fluid intelligence and executive function resources (i.e., the algorithmic mind) and thinking dispositions (i.e., the autonomous mind - see Stanovich, 2009).

Moving on to the topic of the development of reasoning heuristics, I started Chapter 2 by summarizing the existing empirical evidence regarding age-related changes in heuristic use. Although there is plenty of evidence that logical competence, rule use, and normatively correct responding generally increase with development, heuristic use can also increase or remain stable across age groups, which results in “messy” developmental trajectories in children's performance on heuristics and biases tasks. Then (in Sections 2.1.2. and 2.1.3.) I described three developmental dual-process accounts which have attempted to identify the driving forces behind these changes: that of Stanovich and colleagues (e.g., Kokis et al., 2002), Klaczynski and colleagues (e.g., Klaczynski & Cottrell, 2004), and the fuzzy trace theory (e.g., Brainerd & Reyna, 1992). Although these theories show marked differences, there is still some consistency in the claims that they make (other than proposing that two separate processes form the basis of human reasoning). First of all, these theorists agree that both types of process are available at all

points of development, although they go through changes, especially between early childhood and adolescence. These theorists are also in agreement that heuristics are independent of cognitive capacity, and that they work outside conscious control. In addition, they propose that Type 1 (i.e., quick, contextualised, automatic) processing is the default, but slow and effortful Type 2 processes can override or suppress them. Finally, they all agree that Type 2 reasoning requires the decontextualization of problems (by means of decoupling, metacognitive operations, or by relying on the actual stimuli, and suppressing automatic inferences and personal interpretations of problems).

In Section 2.1.2. I also described a taxonomy of heuristics and biases proposed by Stanovich et al. (2008; see also Table 2.1.). Based on this taxonomy, it is possible to make predictions about the relationships between performance on the actual tasks that we used in our experiments, and participants' cognitive capacity and thinking styles. According to Stanovich et al. these predictions should equally hold in the case of children and adults. The first two categories of bias that they identified were the "cognitive miser" and "override failures". These include the tendency to default to effortless (Type 1) processing, to focus on some salient (but often irrelevant) aspects of problems, or to fail to engage in sustained decoupling in order to apply a normative rule. Examples for these include the representativeness heuristic (i.e., the positive/negative recency effect, and the engineers and lawyers problem with a representative description), framing effects (such as the sunk cost and the if-only fallacy), and the belief bias in syllogistic reasoning. In the case of all of these biases (that is, virtually all heuristics and biases that we looked at in our studies) we can expect that people with higher cognitive ability and higher need for cognition (or people who are instructed to reason logically) should respond more normatively. By contrast, in the case of biases which stem from "mindware gaps" (i.e., missing probability knowledge), such as the conjunction fallacy, there should be no relationship between people's performance and their cognitive abilities and the instructions they are given.

Finally, in Section 2.3. I summarized the predictions of these theorists in a table (Table 2.2.) that I reproduce here in Table 7.1.

Table 7.1. *Developmental trajectories predicted by developmental dual-process theories, and the cognitive foundations of these changes (a reproduction of Table 2.2.)*

Developmental theories of reasoning	Normative responding	Heuristic responding	Development is driven by...
Stanovich et al.	Steadily increases with age, or remains stable	Decreases or remains stable with age	increases in cognitive ability although thinking dispositions and relevant knowledge play a role as well
Klaczynski et al.	Can increase, decrease or remain stable with age, although it is expected to increase during adolescence	Can increase, decrease or remain stable with age	the development of metacognitive abilities, and thinking dispositions, together with real-life experiences
Reyna et al.	Increases with age, or U-shaped pattern	inverted U-shaped pattern	changes in representations as a result of experiences; increasing reliance on gist processing

### 7.1.2. *Contrasting our experimental findings with the predictions of dual-process theories.*

I consider the evidence for age-related changes in heuristic reasoning first, and then I move on to discuss the possible driving forces behind these changes. It is important to note that, as our results show, even in the case of the same heuristic, different developmental patterns can be expected depending on the age groups that we look at. Nevertheless, it seems that initially (that is, in early and mid-childhood) the use of reasoning heuristics steadily increases. In fact, in Experiment 1 we found this to be the case for the conjunction fallacy, the if-only fallacy, and the belief bias in syllogistic reasoning in the case of children between the age of 5 and 11. In Experiment 2 we administered the same tasks to a group of adolescents between the age of 11 and 16. In this experiment we found little evidence for age-related changes in any of the heuristics, apart from an increase in the

conjunction fallacy. However, if we look at the changes between the age of 5 and 16, and we collapse overall heuristic, normative and “other” responding across tasks, there seem to be some recognizable patterns in the way children respond (see Figure 7.1.).

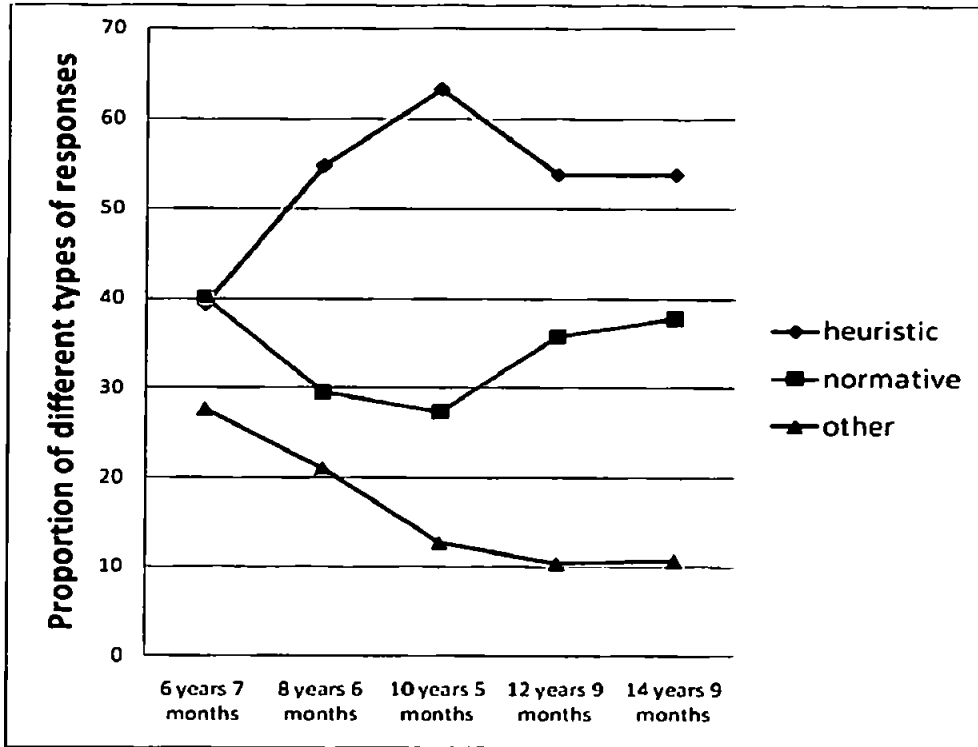


Figure 7.1. Age-related changes in the overall proportion of heuristic, normative and “other” responses on four typical heuristics and biases tasks between the age of 5 and 16 (based on Experiments 1 and 2.)

First of all, heuristic responding is dominant in all age groups (as predicted by all dual-process theories), apart from in the youngest age group where giving normative responses is as likely as giving heuristic responses. Although this could be interpreted as a sign of random responding, looking at the response patterns on the individual tasks shows that even the youngest children have preferred responses on all of these problems (see Table 3.3.). Going back to the developmental patterns, there seems to be a clear trend for



“other” responses to decrease with age. By contrast, normative responding shows a U-shaped pattern, and heuristic responding shows an inverted U-shaped pattern, just as the fuzzy-trace theory would predict. The point in time where heuristic responding starts to decrease and normative responding starts to increase appears to be around age 10 or 11 which roughly corresponds to the Piagetian notion (e.g., Piaget, 1976) that this is the time when children reach the formal operational stage. Nevertheless, it is also the time when heuristic reasoning is at its highest. It is also worth mentioning that after an initial increase, very soon the level of normative responding reaches a plateau, or at least the increase seems to slow down. In addition, even at the age of 15 normative responding is at the same level as at 7 years, which is clearly not what traditional theories of cognitive development would predict.

These age-related changes are also informative regarding the literature on the development of reasoning heuristics. Whereas some authors have reported an age-related increase in heuristic use (e.g., Davidson, 1995), others have found the opposite pattern (e.g., Kokis et al., 2002), and yet others have found no change (e.g., Klaczynski, 2001). Looking at the patterns in our sample, we can find evidence for all of these patterns, depending on which age groups we compare (we could actually expect these patterns to also differ for the individual heuristics that we examined in these studies, but I am not going to discuss this here). Interestingly, our general pattern fits well with the findings reported by the above authors, although there are some differences as well. Davidson (1995) found an increase in heuristic responding during the primary school years, and we have found the same. Kokis et al. (2002) reported that heuristic responding generally decreased between the age of 10 and 12, and we have found a similar pattern between 11 and 13 years (an independent samples *t* test indicated a significant decrease in heuristic responding between age groups 3 and 4 on Figure 7.1.;  $t(56)=3.51, p<.01$ ). However, in contrast with Kokis et al. (2002) who reported a positive relationship between normative

responding and cognitive capacity, we have found no relationship between general intelligence (as indexed by the short form of the WISC) and normative or heuristic responding in this age group. On the other hand, heuristic responding was *positively* correlated with verbal working memory scores ( $r(56)=.50, p<.001$ ). Finally, looking at the early to mid-adolescent age groups (i.e., age groups 4 and 5), we have found (similarly to Klaczynski, 2001) no change in heuristic responding, although, in contrast with the results of Klaczynski (2001) there was no overall increase in normative responding. However, this might be because in our experiment children did not get strong instructions to reason logically.

So what could be the driving forces behind these age-related changes? The steady decrease in “other” responses (which are atypical responses that do not seem to make sense) in Experiments 1 and 2 could *not* be directly related to the steady increase in cognitive capacity with development, although they are probably related to children’s increasing knowledge (however, this was not measured separately). These responses should probably be considered as random noise.

In Experiment 2 we employed tasks to measure the three main aspects of executive functioning (i.e., working memory, inhibition and set-shifting). All of these tended to increase with age. Although working memory was unrelated to heuristic responding in the case of adolescent participants (at least when we controlled for the effects of age), there was a negative correlation between heuristic responding and inhibition in the case of the belief bias and the sunk cost problems. By contrast, the if-only fallacy was positively correlated with inhibition skills. Finally, overall normative responding was positively correlated with participants’ set-shifting ability. In Experiment 3 we employed an index of analogical reasoning ability as a measure of executive functioning, and we used a new set of heuristics and biases problems. Analogical reasoning, especially when it requires resistance to interference (which can be semantic, perceptual or relational) is a pure

measure of the ability to decouple representations. It is noteworthy that analogical reasoning ability was mostly unrelated to susceptibility to heuristics, or if there was a relationship, this was positive (in the case of the engineers and lawyers problem with a representative description). The correlations between the measures of cognitive capacity and heuristic reasoning were also mixed, with more evidence for positive than for negative relationships. These results are clearly in contrast with Stanovich et al. (2008) who predict a negative relationship between heuristic reasoning and cognitive capacity, apart from in the case of the conjunction fallacy where they expect no relationship. Note that although in these experiments we did not take into account the effects of relevant knowledge, according to Stanovich et al. (2008) knowledge has no effect on reasoning in the case of the tasks that we employed here, apart from the conjunction fallacy task, which is attributed to mindware gaps in probability knowledge.

Those researchers who report inverted U-shaped developmental patterns (e.g., Brainerd et al., 2004; De Neys & Everaerts, 2008) tend to attribute this to the interaction of two separate processes, where the later developing process inhibits/suppresses the output of the primary process. For example, Brainerd et al. (2004) claim that gist (i.e., contextualised) processing increases most profoundly between the age of 7 and 11, whereas verbatim processing (i.e., decontextualised, fact-based processing) shows a marked increase between the age of 11 and 14. This fits well with the age trends that we found, and also with the idea that the increase in heuristic processing is attributed to the primary (contextualised) system, whereas the decrease in heuristic processing is related to the secondary (decontextualised) system. However, there are still some intriguing details. For example, the primary (i.e., gist-based) system is best indexed by working memory (and probably the ability to retrieve stored knowledge - see also Section 2.2. for the evidence from the area of contextual reasoning). However, as I noted above, it is hard to identify the decontextualisation process with any measure of cognitive capacity or executive

functioning due to the mixed relationships between these and heuristic/normative reasoning. The explanation that fuzzy-trace theorists give for these mixed correlations is that the secondary system is a memory system that cannot be indexed by any single measure of cognitive capacity. If this conjecture is right, future research on the development of heuristics and biases should focus more on the relationships between memory performance and reasoning. The phenomena that could be interesting in this respect involve the ease of generating associates for words, and the frequency of source monitoring errors (e.g., attributing self-generated word associates to others; for example, when participants mistakenly believe that a stereotype which was activated by task content was included in the description of a task). A potential way of investigating these questions could be to apply the DRM (Deese, 1959; Roediger & McDermott, 1995) paradigm. The relationship between children's susceptibility to false memories (i.e., their ability to automatically activate the critical concept on the DRM task) and their tendency to reason heuristically could be investigated.

### *7.1.3. Our findings regarding heuristics and biases, and executive functioning in autism*

Probably the most relevant cognitive theory of autism for the investigation of heuristic reasoning is the Weak Central Coherence (WCC) theory (Frith & Happé, 1994; Happé, 1999 –see also Section 4.1.2). The WCC theory proposes that typically developing individuals tend to engage in global processing, building up a gist-based representation, whereas autistic individuals engage in more detailed, local or piecemeal processing. As a result, autistic people are less able to process information in context (e.g., Jolliffe & Baron-Cohen, 1999). From this it seems to follow that autistic children should be able to avoid some reasoning heuristics. According to Stanovich (1999, 2006) decoupling or decontextualisation is probably the most important function of the analytic system. If

contextualisation does not take place in autism (or it takes place to much less extent than in typical development) we can expect that autistic people will perform better in situations when a decontextualised representation is required for a logical/normative response. A recent version of the WCC theory (Happé & Frith, 2006) proposes that individuals with autism might be able to contextualise representations as much as typically developing individuals. However, in this case the contextualisation process is expected to be effortful.

Looking at our results, although most of the findings regarding heuristic reasoning can be interpreted in terms of reduced contextualisation/contextualisation through effort in autism, this is clearly not the whole story. Although one of the strongest and most consistent pieces of evidence for the WCC theory is the superior performance of autistic participants on the block design task, we found no evidence for this in either of our studies (see Experiments 4 and 5). Additionally, although analogical reasoning requires contextualised processing, autistic participants showed no impairment on any of the measures of analogical reasoning. Moreover, the underlying processes of analogical reasoning also seemed to be the same across groups, as reflected by the high correlations between item difficulty for autistic and control participants. This shows that participants with autism were able to contextualise non-verbal materials, and they were also able to process complex non-verbal materials similarly to the control group. Finally, autistic participants were able to use their background knowledge when making judgments about the representativeness of a single item (based on a description). That is, they were clearly able to contextualise simple verbal materials. These results seem to suggest that problems with contextualisation in our sample were only evident in the case of complex verbal materials (see also Lopez & Leekam, 2003). Overall these findings are more in line with Bowler et al.'s (2004) task support hypothesis than the WCC theory. According to the task support hypothesis, autistic individuals often show difficulty in retrieving and integrating their background knowledge with a problem context when the task provides little support

for memory retrieval (i.e., when concepts are not directly presented or cued, or when no explicit instructions are given to search for a concept – see also McCroy et al., 2007).

Another relevant cognitive theory of autism is the executive dysfunction account (e.g., Ozonoff, Pennington & Rogers, 1991). This theory was based on the observation that some symptoms of autism (i.e., the need for sameness, the tendency to perseverate, the need to be prompted in order to initiate certain behaviours or to switch set, and autistic individuals' problems with impulse control) are very similar to those shown by individuals with Dysexecutive Syndrome (Baddeley & Wilson, 1988). In our studies we have employed a number of different measures of executive functioning (EF). In Experiment 4 these measures included verbal working memory, inhibition and set-shifting ability (which, according to Miyake et al., 2000 tap into the three main aspects of EF). In Experiment 5 we investigated both verbal and non-verbal working memory capacity, as well as analogical reasoning ability. We examined analogical reasoning using three different non-verbal analogy tests, which differed in terms of the abstractness of the materials, the complexity of the relations that had to be mapped and the presence/absence of distraction.

Although we have found a consistent deficit (or at least a tendency for this) in the verbal domain (in terms of vocabulary, and verbal working memory scores) autistic participants performed just as well as the control group on all other measures of cognitive capacity and executive functioning. This suggests that EF deficits might not be as common in the case of high functioning autistic participants as one might expect based on the EF dysfunction account. Another interesting finding is the relationship between verbal working memory scores and other aspects of EF which were only present in the autistic sample. In Experiment 4 set-shifting scores were significantly (moderately) correlated with verbal working memory scores, and in Experiment 5 there was a significant (strong) correlation between verbal working memory and the backward version of the Corsi blocks task (which is a measure of both non-verbal working memory and central executive

resources – see Vandierendonck et al., 2004). Both the set-shifting task and the backward recall version of the Corsi blocks task are non-verbal measures. However, both tasks require the strategic organization and possibly the labelling of materials in working memory (see e.g., Miyake et al., 2004). If this process breaks down due to verbal working memory impairments, we can expect problems in autism in the case of certain EF skills (whether verbal or non-verbal), such as mental flexibility and strategic retrieval of knowledge. A better understanding of these processes would be useful, and could be the subject of future studies. This could also lead to a better understanding of the performance of autistic individuals on higher order reasoning tasks, such as the heuristics and biases tasks we investigated in the present experiments.

In summary, although our studies have not yielded a clear conclusion as to whether autistic participants are more or less biased than typically developing individuals, this is no surprise if we consider the complexity of the development of reasoning heuristics in the general population. It is more important that we have found consistent evidence that complex verbal processing is atypical in autism, and there is also some indication that this might be related to some aspects of executive functioning (such as the strategic retrieval of procedures and the organization of knowledge). We have also found convincing evidence that the basic ability to reason analogically is intact in autism (which is in contrast with the predictions of both the EF deficit theory and the WCC theory).

#### *7.1.4. The role of knowledge and thinking dispositions in reasoning.*

In Experiments 6, 7 and 8 we investigated the role of knowledge (i.e., education) in probabilistic reasoning. We also looked at the effect of thinking styles and instructions on reasoning. Participants with a higher need for cognition (Cacioppo & Petty, 1982) are dispositionally inclined to think harder about problems, and to enjoy effortful cognitive activity in general. Instructing participants to reason logically, however, reduces the effect

of individual differences in thinking styles, resulting in a higher level of mental effort in all participants.

The fact that the available knowledge base changes rapidly with age and education is recognized by dual-process theories, and this has also been used to explain why cognitive capacity is not a good indicator of children's reasoning performance in many cases (see Sections 2.1.2. and 2.1.3). For example, according to Klaczynski and colleagues (e.g., Jacobs & Klaczynski, 2002; Klaczynski, 2001; Klaczynski, 2009) social, motivational, and affective influences, as well as prior beliefs, greatly affect the way people reason about problems. Klaczynski and colleagues also emphasize that an increasing number of heuristics is acquired over the course of development, and the use of these heuristics becomes more prevalent with age. Another developmental dual-process account that emphasizes the central role of knowledge and experiences is the fuzzy-trace theory (e.g., Brainerd & Reyna, 2001; Reyna & Brainerd, 1992; Reyna & Ellis, 1994; Reyna & Farley, 2006). The "gist" memory system is considered to be the primary (default) system which generates automatic inferences based on the relational and semantic properties of problems. This system flexibly changes with age, as individuals extract new meanings and structures all the time from the stimuli they encounter. Importantly, not only the contents of long-term memory change with development, but also the reasoning processes based on these fuzzy traces. This is why reasoning performance cannot be reliably linked to cognitive capacity.

As these examples show, developmental dual-process theories recognize the importance of knowledge in reasoning performance. However, they propose that performance can increase, decrease or remain stable with development, and no specific prediction can be made about the relationships between development, knowledge, cognitive ability and reasoning performance due to the great individual differences between people's experiences. Thus, none of the above theorists tried to incorporate the



role of knowledge in a detailed model, and no attempt has been made to explore systematically the interactions between knowledge and individual differences in age, cognitive capacity and metacognitive skills.

As we described in Section 1.3.1., in two recent theoretical papers Stanovich and colleagues (Stanovich & West, 2008; Stanovich et al., 2008) discussed the way in which knowledge (“mindware”) and participants’ abilities interact. These authors have suggested that cognitive capacity will only predict reasoning performance if the following conditions are satisfied. Participants have to have the necessary “mindware” (i.e., relevant knowledge), they have to detect the need to override the default heuristic response, and sustained inhibition or decoupling has to be necessary to solve the problem. If any of these is not present then there will be no relationship between people’s performance on a task, and their cognitive abilities. In addition, it is more likely that people will detect the conflict between heuristics and normative reasoning if they are given instructions to think logically, if there are cues in the task that make the conflict salient (e.g., when a within-subjects design is used as opposed to a between-subjects design), and when people are predisposed to reason carefully, and to invest mental effort into solving problems (see Figure 1.4. for a summary of the model).

Probabilistic reasoning is appropriate for the investigation of the role of knowledge in heuristic reasoning for two reasons. First of all, many theorists (e.g., Fischbein, 1987) emphasize that the rules of probabilistic reasoning are virtually impossible to infer from our everyday experiences which are hopelessly “messy”, or the actual patterns of probabilistic outcomes, given that observable sequences of events are always finite, resemble more what could be predicted based on the fallacies of probabilistic reasoning than on the relevant normative rules (similarly to the situation when somebody is waiting at a bus stop and instead of three buses coming at five-minute intervals, all three arrive at the same time, after 15 minutes – see Hahn & Warren, 2009). Thus, we can expect that the

normative rules of probability are acquired through education. Moreover, most traditional heuristics and biases tasks examine the conflict between a rule of probabilistic reasoning, and the effect of vivid, distracting information. Thus, heuristic reasoning on these tasks usually implies the violation of a rule of probability.

Looking at the results of the experiments in Chapter 6, the data do seem to show a good fit with *some* of the predictions of traditional dual-process theories (e.g., Evans & Over, 1996; Stanovich, 1999). Students with higher cognitive abilities and with higher need for cognition were less susceptible to biases. At least incorrect equiprobability responses were mostly given by students with lower cognitive abilities. In addition, students with higher need for cognition (in Experiment 6) and with higher normative potential and given logical instructions (in Experiment 8) were less sensitive to content effects. This suggests that students with higher normative potential are more able to extract the logical structure of problems, regardless of problem content.

Other aspects of our results, however, are in contrast with the predictions of dual-process theories. For example, the above theorists would not predict a simultaneous increase and decrease in heuristic responding, depending on the type of problem (in this case the representativeness heuristic and the equiprobability bias). There was also no relationship between participants' reliance on the representativeness heuristic and their cognitive abilities and cognitive styles. The latter result can be explained based on Stanovich and West (2008), if we assume that participants relying on the representativeness heuristic lack the knowledge of the relevant normative rules (i.e., they have mindware gaps). This interpretation is supported by the evidence that after learning about various tasks that elicit the representativeness heuristic as part of their education in critical thinking, British psychology students were able to resist this heuristic. On the other hand, the equiprobability bias increased with education. This shows that relevant knowledge can not only be used appropriately, but it can also be used to justify/rationalize

non-normative responses, leading to a counterintuitive increase in reasoning errors. However, this was most likely to happen in the case of participants with a lower “normative potential” (i.e., cognitive capacity and need for cognition), and when participants were not given strong instructions to reason logically. Nevertheless, neither education nor logical instructions increased the overall number of normative responses.

Dual-process theories (at least default-interventionist dual-process theories) do not predict simultaneous increase and decrease in heuristic responding within the same population, and do not expect heuristics to appear or significantly increase in adult populations. Although Stanovich et al. (2008) acknowledge that beliefs and knowledge can have both positive and negative effects on reasoning (i.e., by providing mindware and contaminated mindware) they do not discuss the possibility of how the *same* knowledge can form the basis of both mindware and contaminated mindware in the case of different individuals. Overall, these results show that the processes that underlie performance on these well-known heuristics and biases tasks are not very well understood even in the case of adults. In the next section I review the intuitive rules research programme, and I also contrast our findings with the claims of this approach.

## *7.2. Are two systems better than one?*

This section explores the question of whether dual-process theories offer the best explanation for our findings, or whether single process theories (such as the intuitive rules model – see Section 2.1.4.) could account for the present results as well.

### *7.2.1. Are the results of our studies in line with the intuitive rules approach?*

In Section 2.1.4. I described a single system account of reasoning development, the *intuitive rules research program* (Stavy & Tirosh, 2000). In two recent publications Osman (2004) and Osman and Stavy (2006) criticized dual-process theories on the basis that they

cannot account for typical patterns of reasoning performance in developmental samples. Additionally, these theorists claim that the representations and the procedures in reasoning, whether they are used implicitly, explicitly or automatically, are always the same. In this section I review our empirical findings, and contrast the explanations offered by dual-, and single-process theories.

The intuitive rules approach deals with the development of certain heuristics (i.e., intuitive rules) which are defined as self-evident and self-consistent cognitions (based on Fischbein, 1987). These rules are retrieved without conscious intention and they are activated by certain aspects of tasks. One of the characteristics of this theory is a focus on the importance of saliency, and a distinction between bottom-up and top-down saliency. Another interesting aspect of the intuitive rules approach is the distinction between implicit, explicit and automatic forms of reasoning. In dual-process theories implicit and automatic reasoning would both be categorised as Type 1 processes. By contrast, according to the intuitive rules approach implicit and automatic processing correspond to the two extremes of the same continuum, which is characterized by different degrees of representational strength and stability. Automatic representations are considered to be the final products of cognitive development, and they form the basis of skill-based reasoning, and mental flexibility. Although the representations change with individuals' experiences, the ways that they can be utilized remain the same, and for this reason implicit, explicit and automatic representations are attributed to a single system.

According to Osman and Stavy (2006), experiencing conflict during reasoning is not necessarily evidence that the competing responses are generated by two independent reasoning systems. For example, in the case of bottom-up saliency, reducing the discriminability between salient and non-salient features leads to conflict at the perceptual level. However, it is also possible to create a conflict at the level of explicit, top-down reasoning by presenting students with counterexamples. Finally, two intuitive rules can get

activated simultaneously by different aspects of the same problem, and these can cue conflicting responses. Additionally, in the case of many tasks there are U-shaped developmental patterns, where young children and adults respond similarly, although for different reasons. U-shaped developmental patterns are usually considered to provide evidence for the interaction between two systems (i.e., the gist and the verbatim system in the case of the fuzzy-trace theory, and automatic retrieval and effortful inhibition processes in the case of conditional reasoning – see Sections 2.1 and 2.2.). By contrast, according to the intuitive rules program, these patterns are explained by changes in children’s perceptions of tasks, and in their tendency to rely on particular intuitive rules which changes with age.

Finally, the application of intuitive rules can be both positively and negatively related to cognitive capacity. For example Babai and Alon (2004) reported that the application of the “more A-more B” rule decreased with cognitive capacity. On the other hand, the use of the “same A-same B” rule (i.e., the equiprobability bias) and the “everything can be divided” rule increased with cognitive capacity. Given that the application of these rules can lead to both correct and incorrect responses, depending on the task, there is no set relationship between cognitive capacity and normative responding. In addition, responding on a particular task can show U-shaped and inverted U-shaped patterns which is uninterpretable with regard to the monotonic increase of cognitive capacity throughout childhood and adolescence.

If we contrast these claims with our experimental findings, it seems that there is a much better fit for the intuitive rules approach, than for dual-process theories. Although the fuzzy trace theory can account for U-shaped and inverted U-shaped developmental patterns, they predict no interpretable relationships between cognitive ability and heuristic use, as the latter is considered to be related to the development of the gist memory system. By contrast, the intuitive rules program discusses the possibility of any sort of relationship

between heuristics and cognitive ability. Osman and Stavy (2006) also discuss the role of knowledge in reasoning, and how different intuitive rules can show different developmental trajectories (with simultaneous increase and decrease in heuristic use in different areas). These authors acknowledge that the application of the same rule can lead to both normative and non-normative responses, depending on the problem that students are trying to solve. Finally, these theorists also consider primary heuristics/bottom-up saliency (such as the “more A - more B” rule) and secondary heuristics/top-down saliency (such as the “same A – same B” rule; i.e., the equiprobability bias).

Although the intuitive rules research program was developed to account for students’ intuitions in the domain of learning about science (Stavy & Tirosh, 2000) many of the claims about intuitive rules seem to apply for the tasks typically used in the heuristics and biases research program. For example, the distinction between bottom-up and top-down heuristics could be applied to representativeness-based and equiprobability responses. Despite the obvious correspondences between the two approaches, in order to be able to make clear predictions regarding the development of heuristics and biases, a systematic investigation of the claims of the intuitive rules program would be necessary, using heuristics and biases problems. Thus, although our findings are in line with the general predictions of the intuitive rules program, a more thorough investigation would be necessary, and possibly a new taxonomy of heuristics and biases (for example, on the basis of which heuristics are triggered by bottom-up and which ones are triggered by top-down saliency). This could also provide us with some ideas on which aspects of these tasks are crucial in eliciting a heuristic response. For example, we could expect that presenting the ratio bias problem (e.g., Brecher, 2005) in a non-pictorial form should elicit different responses from when the task is presented in a verbal format. More specifically, whereas pictorial presentation is affected by the discriminability of the number of different categories of items (e.g., black vs. white counters), this should not play a role in

participants' judgments when they are presented with stimuli in a numerical format (e.g., 4 black counters out of 6). A numerical presentation might also increase the overall top-down contributions to reasoning. On the other hand, the effects of perceptual saliency (when participants are presented with stimuli in a pictorial format) should be present even when they are instructed to reason on the basis of normative considerations. By contrast, when a heuristic is based primarily on top-down saliency, interventions to reduce the availability of the dominant inappropriate rule, or to encourage students to think critically and to expect conflicts between their intuitions and normative considerations (such as when participants are instructed to think "like a perfectly logical person") should decrease the inappropriate use of rules. A direct test of these and similar claims based on the intuitive rules program could be another possible future line of research.

### *7.2.2. Some theoretical issues regarding dual-process theories of reasoning.*

Dual-process theories have not only been criticized on the basis that they cannot always account for empirical findings. There are also some other, more theoretical, considerations that I would like to briefly discuss. Keren and Schul (in press) raised a number of issues regarding the scientific merit, and theoretical soundness of two-system approaches to higher order cognition. These authors propose that complex higher order mental phenomena (such as reasoning) are unlikely to be based on two non-overlapping systems, because such systems are unlikely to be either isolable or complete. Keren and Schul (in press) claim that dual-process theorists commit a fallacy when they assume that all the mental processes that require minimal contribution of cognitive resources (and especially working memory capacity) are necessarily the part of a single system. This is similar to saying that the body consists of two organs: the liver and everything else (i.e., "not-the-liver" – see Bedford, 2003). If we consider our findings with children, where working memory capacity was positively correlated with "heuristic" responses the question

arises whether we should say that children produce heuristics effortfully (i.e., relying on Type 2 processes), or whether this should be considered as evidence that children are unable to use reasoning heuristics (at least when solving some typical heuristics and biases problems). However, considering the claim that heuristics are basic and primary processes which is implicitly or explicitly assumed by all dual-process theorists (apart from the fuzzy-trace theorists) both conclusions seem quite problematic.

According to Schachter and Tulving (1994) systems should be distinguishable based on the kind of information that they use, their rules of operation, and their neural substrates. In addition, to demonstrate that two systems exist, it is not enough to show that there are two feature sets (such as quick, automatic, effortless, vs. slow, conscious and effortful), but it is also necessary to demonstrate that other combinations do not tend to appear together. For example, the fact that conscious behaviour can be automatic seems very problematic for dual-process accounts. Additionally, a system should not depend on another system in carrying out its operations. For example, in the case of default-interventionist dual-process theories (e.g., Kahneman & Frederick, 2002), and the heuristic-analytic theory of Evans (2006) it is clear that System 2 could not operate without continuous input from System 1. Although this would still be reconcilable with the idea of two separate systems (if there was a strict hierarchy where one system provides the input for the other – see e.g., Carruthers, 2005) this idea is in conflict with recent findings regarding the role of goals and intentions on various automatic processes (e.g., Eitam, Hassin & Schul, 2008), for example the fact that emotional responses (such as sadness and anger) are not generated automatically (see Feldman, Barrett, Ochsner & Gross, 2007 for a review).

Keren and Schul (in press) also discuss the issue of conflict between competing responses. One possibility is that although reasoners experience simultaneity in the emergence of two conflicting responses, this does not mean that the two responses were



generated at the same time. We could also refer to our findings here with adolescents. Dual-process theorists, such as Stanovich and West (2008), for example, assume that when conflict arises between Type 1 and 2 processes reasoners try and inhibit Type 1 response tendencies, although this override process can fail. In Experiment 2 we found that participants' scores on a measure of inhibition skills (the stop signal task) was negatively correlated with their tendency to give heuristic responses on the sunk cost and syllogistic reasoning tasks. Although correlational evidence should be taken with caution, this seems to support Stanovich and West's (2008) notion of the relationship between inhibition and less reliance on heuristics. However, within the same sample, but using a different task, we have found evidence for the opposite pattern as well (i.e., a positive correlation between inhibition skills, and the if-only fallacy). If we consider inhibition as one of the indices of Type 2 reasoning capacity, these results seem to suggest that different components of Type 2 reasoning can function in conflict with each other (i.e., it is possible that inhibition was needed to suppress the normative response in the case of the if-only task), or that some "heuristic" responses are actually generated by Type 2 processes (which is what we actually found in Experiment 1).

In addition, contradictory mental states can be expected on a theoretical basis in any sufficiently complex system (see Kelso & Engstrom, 2006). For example, Evans and Curtis-Holmes (2005) found that the belief bias in syllogistic reasoning increases under time pressure. This could be taken as a supporting evidence for a dual-process model. However, it is also compatible with the idea that there is a single system, which can process representations based on different criteria (such as validity and believability). Evaluating logical validity might take longer than generating a belief-based response. This is, however, not even necessarily the case. For example, Evans, Handley and Bacon (2009) found that time pressure did not influence the tendency of participants to give belief-based responses to conditional reasoning problems. Handley and Newstead (in submission) even

found that producing logic-based responses to modus ponens inference took less time (and was less prone to error) than producing belief-based responses to the same problems. The bottom line is that even if the evidence for a default belief-based processing was stronger, this fact would still be compatible with both single and dual-process accounts.

Finally, although dual-process theorists like to emphasize the family resemblance across different dual-process models (see e.g., Stanovich & West, 2000), different theorists actually make very different claims about the nature of the proposed systems (as noted by Newstead, 2000). For example, whereas some theorists equate System 2 with general intelligence (e.g., Evans & Over, 1996; Stanovich & West, 2000) or working memory capacity (e.g., De Neys, 2006; Kokis et al., 2002), other theorists explicitly claim that there is no clear relationship between cognitive capacity and decontextualised/rational thinking (see Klaczynski, 2009; Pacini & Epstein, 1999). In his latest book Stanovich (2009) also claims that heuristics and biases tasks measure “the thinking (i.e., rational thinking) that IQ tests miss”. Newstead (2000) also points out that the correlations between different components or measures that are supposed to index the same system are often weak or even absent. (This is also evident from our findings, especially from the fact that heuristic reasoning on different tasks was unrelated, and that the relationship between heuristic reasoning and cognitive capacity was very mixed). The fact that dual-process theories are not very well defined also makes it hard to test their predictions (see also Gigerenzer, 2008 and Section 1.3.3.).

In spite of all these problems, dual-process theories are very popular, and they are used in many different areas of psychology, such as reasoning and decision-making, and social psychology (see Evans, 2008 for a review). As Keren and Schul (in press) note, thinking in dichotomies is simple, intuitively compelling, and dual-process theories offer a good story. Another characteristic of dual-process research is that theorists tend to try and collect supporting evidence for their claims, rather than attempting to disconfirm their own

theories. This is also evident in dual-process theorists' focus on the differences between their proposed systems, rather than on the similarities (or on the possibility that their dichotomies actually represent the extremes of a continuum). This is, by the way, also characteristic of the whole discipline of psychology where results are divided into real effects ( $p < .05$ ) and no effects ( $p > .05$ ), although there is no essential difference between an experimental result of  $p = .050$  and  $p = .051$  (see Loftus, 1996). Based on all these claims, Keren and Schul (in press) call for a more precise specification of the theoretical constructs of dual-process theories, and more rigorous empirical tests of these theories, which should allow not just for collecting supporting evidence for dual-process theories, but also for the falsification of the claims these theories make.

### *7.3. Concluding comments and possible future directions*

This section discusses the issue of the rationality of heuristics and biases. I also outline some possible future directions of research, and I summarize my conclusions.

#### *7.3.1. Is "rational thinking" more rational than heuristics and biases?*

One thing that most dual-process theorists are in agreement on is that heuristic reasoning, although often very useful, can lead to irrational behaviour in certain situations. Theorists, such as Evans and Over (1996) and Stanovich (2004) emphasize that heuristics are rational evolutionarily. However, they are not necessarily rational from the individual's point of view (for example, they might lead to bad decisions when we choose from different products or when we interpret legal documents). Other theorists, on the other hand, consider heuristics and contextualised reasoning to be examples of the most effective and mature (as well as the most economic) reasoning processes. One example of this is the concept of gist processing (see the review on the fuzzy-trace theory in Section 2.1.3). The gist of an event depends on people's previous experiences with situations which are similar

to their present experience, and it makes it possible for people to derive the meaning of situations using their relevant knowledge. Fuzzy-trace theorists also emphasize that conscious, fact-based thinking is slow, low-capacity and error-prone. Similarly, according to the intuitive rules approach (see Section 2.1.4.) automatic reasoning is the final stage of cognitive development, which forms the basis of mental flexibility. Gigerenzer and colleagues also emphasize that norms should not be content blind (i.e., decontextualised), and they also consider heuristics to be the constituents of the “adaptive toolbox” which enables individuals to reason in a fast and frugal way, which is also very effective.

Our results add some interesting details to the rationality debate. One important finding is that in the case of young children, contextualisation was effortful, which demonstrates that there is not necessarily a contextualised=quick and easy; decontextualised=slow and effortful correspondence. This relationship in the case of older children changed, yielding no relationship between intelligence/cognitive capacity and either contextualised or decontextualised thinking. Probably more importantly (at least regarding the issue of the rationality of heuristics) it is clear from Figure 7.1. that 6 year-olds were as “rational” (i.e., normative) on our set of tasks as 15 year-olds (with the age groups in between reaching even lower normative scores). Then, moving on to the results from the university student sample, we have found that learning new (relevant) concepts, such as the concept of randomness, decreased normative responding in the case of our sample (although it was more likely to happen in the case of low-ability students, with no effect of education in the case of high ability students). Then consider the finding that heuristic reasoning was less prevalent in our autistic sample than in the control group on a number of tasks, and higher ability autistic participants were more inclined to reason heuristically in some cases than lower ability autistic children. These results pose interesting questions about the adaptive value of reasoning heuristics.

One example of a heuristic that participants with autism were consistently less likely to use than the control group, and that younger children relied on less than older children is the conjunction fallacy. In the past 26 years around a hundred scientific papers have been published on the conjunction fallacy, and the “Linda problem” has been a key topic in the debate on human rationality. The reason for this is that, for most of us, the draw towards the conjunctive option is so powerful that even when we are aware that something is not quite right about our judgment, it is nevertheless very difficult to resist the powerfully compelling intuitive choice. Consider, for example, Gould’s (1992) famous comment on the task: “...*I know that the third statement is least probable, yet a little homunculus in my head continues to jump up and down, shouting at me—“but she can't just be a bank teller; read the description.”*”

As Hertwig and Gigerenzer (1999) in their famous critique of the conjunction fallacy noted, committing the fallacy is clearly not rational according to the rules of probability theory and logic. However, it can be perfectly rational in a social sense. In support of this claim they give the following example. According to the basic principle of internal consistency, the preference for one choice over another should be independent of the availability of other alternative choices (that is, the context in which a problem is presented). However, imagine that you are taking part in a dinner party, and it looks like there are fewer pastries than there are people. Although you would normally have dessert, in this situation you decide that you will not take the last remaining éclair from the tray, giving a chance to other people to take it. This behaviour is not rational according to probability theory. Nevertheless, it is polite and makes perfect sense socially. As this example illustrates, the relevance of contextual processing is not restricted to the domain of reasoning. In fact its importance is much more evident in everyday situations (and in social situations, especially). This example also gives a possible explanation for how problems with contextual processing could lead to serious problems in everyday social situations in

the case of individuals with autism. It would be interesting to study heuristic reasoning in autism in relation to the behavioural symptoms of autism (such as deficits in social interaction and social imagination) in the future. It is possible that although heuristics and biases tasks are unnatural in a certain sense, they could explain some of the problems that individuals with autism experience in their daily lives.

### *7.3.2. Possible future lines of research and conclusions*

I have already mentioned that looking at the relationships between the development of memory and reasoning performance could be a novel, theory-driven approach to reasoning development. Although memory retrieval, and its relationship with conditional reasoning has been studied before (see e.g., Markovits & Barouillet, 2002), this has not been done in the case of reasoning heuristics, not even by fuzzy-trace theorists who expect a close relationship between memory development and reasoning. Exploring the relationships between memory development and heuristic reasoning would be interesting to do in the case of both typically developing and autistic children. The links between memory retrieval and executive functioning/working memory capacity should also be investigated in both populations.

In the present experiments we have investigated the development of reasoning heuristics. Our youngest participants were 5 year-olds, and the oldest participants were young adults. This is obviously not the whole picture regarding the development of heuristics across the life span. We can expect that executive functioning and working memory capacity will be gradually developing from childhood into young adulthood, which should be followed by a decline of cognitive resources as people get older. If we accept that there is a close relationship between heuristic reasoning and working memory and executive functioning, we could expect age-related changes in heuristic use in adulthood as well. This is, however, a topic that has rarely been investigated so far. Not

surprisingly (at least, considering our own mixed findings with children) the two studies which have been conducted with older adults yielded opposing results. Strough, Mehta, McFall and Schuller (2008) have found that older adults were less susceptible to the sunk cost fallacy than younger adults. These authors have attributed this to the tendency of older adults to focus on positive information, and their mature skills in emotion regulation which helps them to maintain positive mood states for longer (see Cartensen, Mikels & Mather, 2006 for a review). By contrast, De Neys and Van Gelder (2008) reported that older adults (and children) were more susceptible to the belief bias in syllogistic reasoning than young adults, and their susceptibility to the belief bias was significantly correlated with their performance on some typical heuristics and biases tasks (i.e., the worse they performed on the heuristics and biases tasks, the less they could resist the belief bias). De Neys and Van Gelder attributed these patterns to the “rise and fall” of belief inhibition across the life span.

Another interesting question is whether there is a relationship between reasoning performance (and especially susceptibility to heuristics) and the behavioural symptoms of autism as measured by diagnostic tools (such as the Autism Diagnostic Interview or the Autism Diagnostic Observation Schedule). For example, Hertwig and Gigerenzer (1999) suggested that there is a close relationship between heuristic/contextualised reasoning and behaviour in social situations. There is also an interesting parallel between the development of reasoning heuristics and theory of mind (ToM). Although ToM reasoning (just as heuristics) is often considered to be a basic, or even modular mechanism, there is much evidence that ToM reasoning requires executive functioning resources (see e.g., Apperly, Samson & Humphreys, in press, for a review). However, it is possible that although beliefs are inferred automatically, it requires mental effort to use them appropriately in a particular situation (see e.g., Cohen & German, 2009). Similarly, it is possible that heuristics are generated automatically, but due to the fact that participants

have to respond to problems by choosing from multiple response options, there is some deliberation involved, which requires the involvement of cognitive resources. Back and Apperly (in press) make an interesting distinction between spontaneous and automatic inferences. The idea is that people might track other's beliefs spontaneously (i.e., without the need for being instructed) which leads to quick responses when they are unexpectedly asked to report others beliefs. This, however, does not imply that these inferences are automatic and effortless. Similarly, it might just be that the criteria that heuristic responses rely on (for example, using background knowledge to answer questions, instead of relying on probabilistic information presented in the task) are easier and more natural to apply. However, individuals (for example, individuals with autism) who find it hard to rely on their background knowledge might prefer to rely on other criteria (in this case, to use the probabilistic information presented in the task) as the basis of their response.

Another interesting parallel is that ToM reasoning seems to be related to the development of verbal skills, although it is only apparent in the case of young children, and in the case of participants with autism. This is similar to our findings in that verbal working memory was involved in heuristic reasoning in the case of young children and autistic participants, but not in the case of late adolescents and adults. Recent investigations into ToM could serve as a model for developing novel paradigms to examine heuristic reasoning. Developing non-verbal measures of heuristic reasoning (see Apperly, Riggs, Simpson, Samson & Chiavarino, 2006 for an example for a non-verbal ToM task) would be especially useful to find out whether children and individuals with autism need to invest cognitive effort because of the processing demands of verbal reasoning problems, or whether they do not generate heuristic responses spontaneously/automatically in the first place. (Although unrelated to the present investigation, it would be also interesting to study ToM performance in autism using these novel methodologies to understand better the sources of their problems.)



Going back to the research on reasoning heuristics, adults' responses could be investigated using more sensitive methods, such as reaction times, eye tracking or physiological measures. These techniques have already been used by various researchers, although without conclusive results (at least the results are hard to generalize across different tasks). One possible reason for this is that the underlying processes of heuristics are actually quite diverse (this is a reasonable assumption looking at the differences between the developmental trajectories of heuristics that we have found in our samples). Another reason for this might be that the underlying processes of heuristics are not well understood, or well specified to enable systematic investigations into the claims of dual-process theories. Unless we make falsifiable claims, it is not possible to rigorously evaluate the usefulness of dual-process theories for explaining reasoning performance. In this respect systematically contrasting the predictions of dual-process theories and alternative accounts of reasoning (such as the intuitive rules approach or the fuzzy trace theory) would be a very useful step forward, especially when looking at developmental samples. Claims about processing fluency/effort and conflict between different reasoning processes would also be interesting to examine using physiological measures.

In summary, in this series of studies we have examined three main questions: the development of reasoning heuristics in typical populations, the development of reasoning heuristics in autism, and the role of education and knowledge in heuristic reasoning. I have discussed our findings in relation to dual-process theories, the cognitive theories of autism, and educational theories of reasoning heuristics. These results can inform theoretical debates about the appropriateness of dual-process theories and about the rationality of reasoning heuristics. In addition, our findings highlighted the need for more rigorous investigations in certain areas, and I have also outlined some possible directions for future research.

## References

- Adolphs, R. (1999). The Human Amygdala and Emotion. *The Neuroscientist*, 5, 125-137.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*. 4th ed. (DSM-IV). Washington, DC: American Psychiatric Association.
- Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, 45, 153–219.
- Andrews, G. Halford, G.S., Bunch, K.M., Bowden, D. & Jones, T. (2003). Theory of mind and relational complexity, *Child Development*, 74, 1476–1499.
- Apperly, I. A. & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116, 953-970.
- Apperly, I. A., Riggs, K. J., Simpson, A., Samson, D., & Chiavarino, C. (2006). Is belief reasoning automatic? *Psychological Science*, 17, 841–844.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends In Cognitive Sciences*, 9, 572-577.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, 45, 190-201.
- Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, 125, 591–600.
- Arthur, W., Jr., & Day, D. V. (1994). Development of a short form for the Raven advanced progressive matrices test. *Educational and Psychological Measurement*, 54, 394–403.
- Ayton, P. & Fischer, I. (2004). The Hot Hand Fallacy and the Gambler's Fallacy: Two faces of Subjective Randomness? *Memory & Cognition*, 32, 1369-1378.

- Babai, R., & Alon, T. (2004). *Intuitive thinking, cognitive level or grade level: What predicts students' incorrect responses in science and mathematics*. Paper presented at the National Association of Research in Science Teaching Conference, Vancouver.
- Babai, R., Brecher, T., Stavy, R., & Tirosh, D. (2006). Intuitive interference in probabilistic reasoning. *International Journal of Science & Mathematics Education*, 4, 627-639.
- Back, E., & Apperly, I.A. (in press). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition*.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 1, 158-173.
- Baddeley, A., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition*, 7, 212-230.
- Baird, G., Cass, H. & Slonims, V. (2003). *British Medical Journal*, 327, 488-493.
- Ball, L. J., Lucas, E. J., Miles, J. N. V., & Gale, A. G. (2003). Inspection times and the selection task: What do eye-movements reveal about relevance effects? *Quarterly Journal of Experimental Psychology*, 56A, 1053-1077.
- Ball, L. J., Phillips, P., Wade, C. N., & Quayle, J. D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing models. *Experimental Psychology*, 53, 77-86.
- Barkow, J., Cosmides, L., & Tooby, J. (Eds.) (1992). *The adapted mind*. New York: Oxford University Press.
- Baron-Cohen, S. (1989). The autistic child's theory of mind: the case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30, 285-98.
- Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. MIT Press/Bradford Books.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6, 248-254.

- Baron-Cohen, S., Auyeung, B., Ashwin, E., Knickmeyer, R. (2009). "Fetal testosterone and autistic traits: A response to three fascinating commentaries". *British Journal of Psychology*, 100, 39-47.
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38, 813-822.
- Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, 21, 37-46.
- Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, 29, 407-418.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste Y. & Plumb, I. (2001). The 'Reading the Mind in the eyes' test revised version: A study with normal adults, and adults with Asperger Syndrome or High-Functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241-252.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. (2001). The autism spectrum quotient (AQ): evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5-17.
- Barrouillet, P., Markovits, H., & Quinn, S. (2002). Developmental and content effects in reasoning with causal conditionals. *Journal of Experimental Child Psychology*, 81, 235-248.

- Batanero, C., Serrano, L., Garfield, J. B. (1996). Heuristics and biases in secondary school students' reasoning about probability. In L. Puig & A. Gutiérrez (Eds.), *Proceedings of the 20th conference on the international group for the psychology of mathematics education* (Vol. 2, pp. 51–59). University of Valencia.
- Beck, S.R., Robinson, E.J., Carroll, D.J., & Apperly, I.A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development*, *77*, 413-426.
- Bedford, F.L. (2003). More on the not-the-liver fallacy: Medical, neuropsychological, and perceptual dissociations. *Cortex*, *39*, 170–173.
- Begeer, S., Meerum Terwogt, M., Lunenburg, P. & Stegge, H. (2009). Additive and subtractive counterfactual reasoning in children with high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 1593-1597.
- Belmonte, M., Cook, E., Anderson, G., Rubenstein, J., Greenough, W., Beckel-Mitchener, A., Courchesne, E., Boulanger, L., Powell, S., Levitt, P., Perry, E., Jiang, Y., DeLorey, T., Tierney, E. (2004). Autism as a Disorder of Neural Information Processing: Directions for Research and Targets for Therapy, *Molecular Psychiatry*, *9*, 646-663.
- Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development*, *67*, 1816-1835.
- Bernstein, D. M., Atance, C., Loftus, G. R., & Meltzoff, A. N. (2004). We saw it all along: Visual hindsight bias in children and adults. *Psychological Science*, *15*, 264 – 267.
- Bird, G., Catmur, C., Silani, G., Frith, C.D. & Frith, U (2006). Attention does not modulate neural responses to social stimuli in Autism Spectrum Disorders. *NeuroImage*, *31*, 1614-1624.

- Bjorklund, D.F. & Jacobs, J. W. (1985). Associative and categorical processes in children's memory: The role of automaticity in the development of free recall, *Journal of Experimental Child Psychology*, 39, 599–617.
- Blackmore, S. (1999). *The Meme Machine*, Oxford, Oxford University Press.
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77, B25-B31.
- Borovcnik, M., & Bentz, H.-J. (1991). Empirical research in understanding probability. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 73–105). Dordrecht: Kluwer.
- Borovcnik, M., & Peard, R. (1996). Probability. In J. Kilpatrick (Ed.), *International handbook of mathematics education* (pp. 371–401). Dordrecht: Kluwer.
- Boucher, J., & Warrington, E. K. (1976). Memory deficits in early infantile autism: some similarities to the amnesic syndrome. *British Journal of Psychology*, 67, 73–87.
- Bowler, D. M., Gardiner, J. M. & Berthollier, N. (2004). Source memory in Asperger's syndrome. *Journal of Autism and Developmental Disorders*, 34, 533-542.
- Bowler, D. M., Gardiner, J. M. & Grice, S. (2000). *Journal of Autism and Developmental Disorders*, 30, 305-316.
- Bowler, D. M., Matthews, N.J., & Gardiner, J. M. (1997). *Neuropsychologia*, 35, 65-70.
- Brainerd, C. (1981). Working memory and the developmental analysis of probability judgment, *Psychological Review*, 88, 463–502.
- Brainerd, C. J. (2004). Dropping the other U: An alternative approach to U-shaped developmental functions. *Journal of Cognition & Development*, 5, 81-88.
- Brainerd, C.J., Reyna, V.F. (1992). Explaining "memory-free" reasoning. *Psychological Science*, 3, 332-339.
- Brainerd, C.J., Reyna, V.F. (2001). Fuzzy-trace theory: Dual Processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behavior*, 28, 49-100.

- Brecher, T. (2005). *Application of the intuitive rule: "More A-more B" in a probability task. A reaction time study*. Unpublished master's thesis, Tel Aviv University, Israel (in Hebrew).
- Brock, J., Norbury, CF, Einav, S, Nation, K (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements, *Cognition*, 108, 896-904.
- Brock, J., Brown, C.C., Boucher, & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Developmental Psychopathology*, 14, 209-224.
- Brosnan, M. J., Scott, F. J., Fox, S., & Pye, J. (2004). Gestalt processing in autism: Failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry*, 45, 459-469.
- Bull, R., Phillips, L. H., & Conway, C. (2008). The role of control functions in mentalizing: Dual task studies of theory of mind and executive functioning. *Cognition*, 107, 663-672.
- Bumette, C. P., Mundy, P. C., Meyer, J. A., Sutton, S. K., Vaughan, A. E., & Charak, D. (2005). Weak central coherence and its relations to theory of mind and anxiety in autism. *Journal of Autism and Developmental Disorders*, 35, 63-73.
- Buss, D. M. (1999). *Evolutionary psychology: The new science of the mind*. Boston: Allyn and Bacon.
- Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31, 61-83.
- Byrne, R. M. J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, 6, 426-431.
- Byrne, R.M.J., (2007). Precis of The Rational Imagination: How People Create Alternatives to Reality. *Behavioral and Brain Sciences*, 30, 439 - 453.
- Byrne, R. M. J., Espino, O., & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language*, 40, 347-373.

- Cacioppo, T. J., & Petty, I. L. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.
- Callaert, H. (2004). In search of the specificity and the identifiability of stochastic thinking and reasoning. In M. A. Mariotti (Ed.), *Proceedings of the third conference of the European society for research in mathematics education*. Pisa: Pisa University Press.
- Capon, A., Handley, S. J. & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking and Reasoning*, *9*, 203-244.
- Carey, S. (2001). Evolutionary and ontogenetic foundations of arithmetic. *Mind and Language*, *16*, 37–55.
- Carlson, S. M., Moses, L. J. & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*, 672–691.
- Caron, M.-J., Mottron, L., Berthiaume, C., & Dawson, M. (2006). Cognitive mechanisms, specificity and neural underpinnings of visuo-spatial peaks in autism. *Brain*, *129*, 1789-802.
- Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404–431.
- Carroll, J.M. & Kin Yung, C. (2006). Sex and discipline differences in empathising, systemising and autistic symptomatology: Evidence from a student population, *Journal of Autism and Developmental Disorders*, *36*, 949–957.
- Carruthers, P. (2005). Dual-content theory: The explanatory advantages. In P. Carruthers (Ed.), *Consciousness: Essays from a higher order perspective* (pp. 98–115). Oxford, United Kingdom: Oxford Scholarship Online Monographs.



- Carstensen, L. L., Mikels, J. A., & Mather, M. (2006). Aging and the intersection of cognition, motivation and emotion. In Birren, J., & Schaie, K. W. (eds.), *Handbook of the Psychology of Aging* (sixth ed., pp. 343–362). San Diego: Academic Press.
- Chakrabarti, S., Fombonne, E., (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *American Journal of Psychiatry*, *162*, 1133-41.
- Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J. & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*, *46*, 500-513.
- Chiesi, F., Primi, C. (2008). Le proprietà psicometriche della versione italiana della need for cognition scale – short form. *Congresso Nazionale Sezione di Psicologia Sperimentale*, Padova, 18–20 Sep.
- Cho, S., Holyoak, K. J., & Cannon, T.D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition*, *35*, 1445-1455.
- Chomsky, N. (1980). *Rules and Representations*. Oxford: Basil Blackwell.
- Cleeremans, A. & Jimenez, L. (2002). Implicit Learning and Consciousness: A Graded, Dynamic Perspective. In Robert M. French & Axel Cleeremans (eds.), *Implicit Learning and Consciousness: An Empirical*. Psychology Press.
- Cohen, A. S. & German, T.C. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, *111*, 356-363.
- 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.

- Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: the evolution of functional organization. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 85–116). New York: Cambridge University Press.
- Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23, 646–658.
- Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making, *Merrill-Palmer Quarterly*, 41, 328–346.
- Dawkins, R. (1993). Viruses of the Mind. *Free Inquiry*, 34–41.
- Dawson, M., Soulieres, I., Gernsbacher, M.A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science*, 18, 657–662.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall, *Journal of Experimental Psychology*, 58, 17–22.
- De Martino, B., Harrison, N., Knafo, S., Bird, G., & Dolan, R. (2008). Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience*, 28, 10746-10750.
- Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, 66, 819–829.
- De Neys, W. (2003). *Memory search for stored counterexamples during conditional reasoning*. Unpublished doctoral dissertation, University of Leuven, Belgium
- De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, 17, 428-433.
- De Neys, W., & Everaerts, D. (2008). Developmental trends in everyday conditional reasoning: The retrieval and inhibition interplay. *Journal of Experimental Child Psychology*, 100, 252-263.

- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of reasoning. *Cognition, 106*, 1248-1299.
- De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: A test of the semantic memory framework. *Memory & Cognition, 30*, 908–920.
- De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition, 31*, 581–595.
- De Neys, W., Schaeken, W., d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning, 11*, 349-81.
- De Neys, W., & Van Gelder, E. (2008). Logic and belief across the life span: The rise and fall of belief inhibition during syllogistic reasoning. *Developmental Science, 12*, 123-130.
- De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science, 19*, 483-489.
- Diamond, A. (2006). The early development of executive function. In E. Bialystok & F.I.M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). Oxford: Oxford University Press.
- Eitam, B., Hassin, R.R., & Schul, Y. (2008). Non-conscious goal pursuit in novel environments: The case of implicit learning. *Psychological Science, 19*, 261–267.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49*, 709-24.
- Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62*, 328–339.
- Evans, J.St.B.T., (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75*, 451–468.

- Evans, J.St.B.T. (1995). Relevance and reasoning. In S.E. Newstead & J.St.B.T. Evans (Eds.), *Perspectives on thinking and reasoning*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Evans, J. St. B. T. (1998). Matching bias in conditional reasoning. Do we understand it after 25 years? *Thinking & Reasoning*, 4, 45–82.
- Evans, J.St.B.T. (2003). In two minds: dual process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454-459.
- Evans, J.St.B.T (2006a). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*, 13, 378-395.
- Evans, J. S. B. T. (2006b). Dual system theories of cognition: Some issues. In: *Proceedings of the 28th annual conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum. pp. 202-208.
- Evans, J. St. B. T. (2007). On the resolution of conflict in dual-process theories of reasoning. *Thinking & Reasoning*, 13, 321-329.
- Evans, J.St.B.T. (2008). Dual-processing accounts of reasoning, judgement and social cognition. *Annual Review of Psychology*, 59, 255-278.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, 11, 295–306.
- Evans, J.St.B.T., & Curtis-Holmes J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, 11, 382-389.
- Evans, J. St. B.T., Handley, S.J. & Bacon, A.M. (2009). Reasoning under time pressure: A study of causal conditional inference. *Experimental Psychology*, 56, 77-83.
- Evans, J. St. B. T., Handley, S. J., & Harper, C. N. J. (2001). Necessity, possibility and belief: A study of syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 54A, 935–958.

- Evans, J. St. B. T., Handley, S. J., Neilens, H., Bacon, A., Over, D. E. (*in press*). The influence of cognitive ability and instructional set on causal conditional inference. *The Quarterly Journal of Experimental Psychology*.
- Evans, J. St. B. T., Handley, S. J., & Over, D. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 321-335.
- Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove: Psychology Press.
- Evans, J.St.B.T. & Over, D.E., (2004). *If*. Oxford University Press.
- Evans, J. St. B. T., & Pollard, P. (1990). Belief bias and problem complexity in deductive reasoning. In J-P. Caverni, J-M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 131 –154). North Holland: Elsevier.
- Falter, CM, Plaisted, KC, and Davis, G (2008). Visuo-spatial processing in autism--testing the predictions of extreme male brain theory. *Journal of Autism and Developmental Disorders*, 38, 507-15.
- Feldman Barrett, L., Ochsner, K.N., & Gross, J.J. (2007). On the automaticity of emotion. In J.A. Bargh (Ed.), *Social psychology and the unconsciousness: The automaticity of higher mental processes* (pp. 173–219). New York: Psychology Press.
- Ferreira, M. B., Garcia-Marques, L., Sherman, S. J., & Sherman, J. W. (2006). Automatic and controlled components of judgment and decision making. *Journal of Personality and Social Psychology*, 91, 797–813.
- 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.
- Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. The Netherlands: Reidel, Dordrecht.

- Fischbein, E. (1987). *Intuition in science and mathematics*. The Netherlands: Reidel, Dordrecht.
- Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28, 96–105.
- Fischer, J.E., Steiner, F., Zucol, F., Berger, C., Martignon, L., & Bossart, W. (2002). Using simple heuristics to target macrolide prescription in children with community-acquired pneumonia, *Archives of Pediatric and Adolescent Medicine*, 156, 1005–1008.
- Fombonne, E. (1999). The epidemiology of autism: a review. *Psychological Medicine*, 29, 769-86.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18, 253–292.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17, 172-179.
- Frisch, D., (1993). Reasons for framing effects. *Organizational Behavior and Human Decision Processes*, 54, 399–429.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.
- Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Oxford: Blackwell.
- Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind". *Cognition*, 50; 115-132.
- Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, 2, 108-124.
- Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Journal of Developmental Psychology*, 1, 329-342.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10, 483–527.

- Garcia-Retamero, R., Takezawa, M. & Gigerenzer, G. (2006). How to learn good cue orders: When social learning benefits simple heuristics. In R. Sun & N. Miyake (Eds.) *Proceedings of the 28<sup>th</sup> Annual Conference of the Cognitive Science Society* (pp. 1352-1358). Mahwah, NJ: Erlbaum.
- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, 7, 155–170.
- Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences*, 1021, 77-85.
- Geier, D.A., Geier, M.R. (2006). A clinical and laboratory evaluation of methionine cycle-transsulfuration and androgen pathway markers in children with autistic disorders. *Hormone Research*, 66, 182–188.
- Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. New York: Oxford University Press.
- Gigerenzer, G. (2009). Surrogates for theory. *Observer*, 22, 10-13.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Postscript: Fast and frugal heuristics. *Psychological Review*, 115, 238-239.
- Gigerenzer, G., Todd, P.M., & the ABC Research Group (1999). *Simple Heuristics that Make Us Smart*. Oxford University Press, New York.
- Gilovich, T., Vallone, R. & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- Green, L. & Mehr, D.R. (1997). What alters physicians' decisions to admit to the coronary care unit?, *Journal of Family Practice*, 45, 219–226.
- Griggs, R., & Cox, J.R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73, 407-420.

- Goel V, Dolan RJ. (2003). Explaining modulation of reasoning by belief. *Cognition*, 87, B11-22.
- Gould, S.J. (1992) *Bully for Brontosaurus: Further Reflections in Natural History*, Penguin Books.
- Gronchi, G. & Sloman, S. A. (2008). Do causal beliefs influence the hot-hand and the gambler's fallacy? *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*, Washington, DC., 1164-1169.
- Hahn, U., & Warren, P. A. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review*, 116, 454–461.
- Halford, G.S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Lawrence Erlbaum.
- Handley, S., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. B. T. (2004). Working memory, inhibitory control, and the development of children's reasoning. *Thinking and Reasoning*, 10, 175–195.
- Handley, S. & Newstead (in submission) Logic, beliefs and instruction: A test of the default interventionist account of belief bias.
- Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129-154.
- Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66, 843-855.
- Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, 3, 216-222.
- Happé, F., Booth, R., Charlton, R.A. & Hughes, C. (2006). Executive function deficits in Autism Spectrum Disorders and Attention-Deficit/Hyperactivity Disorder: Examining profiles across domains and ages. *Brain & Cognition*, 61, 25-39.



- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism And Developmental Disorders*, *36*, 5-25.
- Happé, F. & Ronald, A. (2008). 'Fractionable Autism Triad': A Review of Evidence from Behavioural, Genetic, Cognitive and Neural Research. *Neuropsychology Review*, *18*, 287-304.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*, 243-259.
- Heller, R. F., Salzstein, H. D., & Caspe, W. B. (1992). Heuristics in medical and nonmedical decision-making. *The Quarterly Journal of Experimental Psychology*, *44*, 211-235.
- Hertwig, R., Benz, B., & Krauss, S. (2008). The conjunction fallacy and the many meanings of "and". *Cognition*, *108*, 740-753.
- Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*, 275-305.
- Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, *24*, 189-233.
- Hirschfeld, L., Bartmess, E., White, S. & Frith, U. (2007). Can autistic children predict behavior by social stereotypes? *Current Biology*, *17*, 451-452.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295-355.
- Hughes, C., Russell, J., Robbins, T.W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 477-92.
- Huizinga, M., Dolan, C. V., & van der Molen, Maurits W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017-2036.

- Hulme, C. & Roodenrys, S. (1995). Practitioner review: verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry*, 36, 373-398.
- Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally-plausible cognitive architecture: An overview of the LISA project. *Current Directions in Cognitive Science*, 14, 153-157.
- Hutchinson, J.M.C., Gigerenzer, G., (2005). Simple heuristics and rules of thumb, where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69, 97–124.
- Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology*, 35, 904–911.
- Jacobs, J. E., & Klaczynski, P. A. (2002). The development of decision making during childhood and adolescence. *Current Directions in Psychological Science*, 4, 145-149.
- Jacobs, J. E. & Potenza, M. (1991). The use of judgment heuristics to make social and object decision: A developmental perspective. *Child Development*, 62, 166-178.
- Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, 34, 81-86.
- Jarrold, C., Butler, D. W., Cottington, E. M., & Jimenez, F. (2000). Linking theory of mind and central coherence bias in autism and the general population. *Developmental Psychology*, 36, 126-138.
- Johnson-Laird, P. N., Legrenzi, P., Girotto, P., Legrenzi, M. S., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106, 62–88.
- Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures test? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38, 527-534.

- Jolliffe, T. & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome—is local coherence impaired? *Cognition*, *71*, 149–185.
- Just, M. A., Cherkassky, V. L., Keller, T. A., Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*, *127*, 1811-1821.
- Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex*, *17*, 951-961.
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics, *American Economic Review*, *93*, 1449-1475.
- Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. pp. 49-81 in T. Gilovich, D. Griffin, and D. Kahneman (eds). *Heuristics & Biases: The Psychology of Intuitive Judgment*. New York. Cambridge University Press.
- Kahneman, D. & Frederick, S. (2005). A Model of Heuristic Judgment. in K. J. Holyoak & R. G. Morrison (eds.) *The Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press. 267-293.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237-251.
- Kahneman, D. & Tversky, A., (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*, 263–291.

- Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology*, 28, 899–904.
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin & Review*, 9, 637-671.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, 2, 389-398.
- Karmiloff-Smith, A. Scerif, G. & Ansari, D. (2003). Double dissociations in developmental disorders? Theoretically misconceived, empirically dubious. *Cortex*, 39, 161-163.
- Kelso, J.A., & Engstrom, D.A. (2006). *The complementary nature*. Cambridge, MA: MIT Press.
- Keren, G. & Schul, Y. (in press) Two is not always better than one. A critical evaluation of two-system theories. *Perspectives on Psychological Science*.
- Klaczynski, P. A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development*, 71, 1347-1366.
- Klaczynski, P. A. (2001). Framing effects on adolescent task representations, analytic and heuristic processing, and decision making: Implications for the normative-descriptive gap. *Journal of Applied Developmental Psychology*, 22, 289-309.
- Klaczynski, P. A. (2005). Metacognition and cognitive variability: A two-process model of decision making and its development. In J. E. Jacobs & P. A. Klaczynski (Eds.), *The development of decision making in children and adolescents* (pp. 39-76). Mahwah, NJ: Erlbaum.
- Klaczynski, P. A. (2007). Education: Theory, practice, and the road less followed. *Journal of Applied Developmental Psychology*, 28, 80-83.

- Klaczynski, P. A. (2009). Cognitive and social cognitive development: Dual-process research and theory. J. B. St. T. Evans & K. Frankish (Eds.), *In two minds: Psychological and philosophical theories of dual processing* (pp. 265-292). Oxford, UK: Oxford University Press.
- Klaczynski, P. A., & Cottrell, J. E. (2004). A dual-process approach to cognitive development: The case of children's understanding of sunk cost decisions. *Thinking & Reasoning, 10*, 147-174.
- Klaczynski, P. A., & Daniel, D. (2005). Individual differences in conditional reasoning: A dual-process account. *Thinking & Reasoning, 11*, 305-325.
- Klaczynski, P. A., & Gordon, D. H. (1996). Goal-directed everyday problem solving: Motivational and general ability influences on adolescent statistical reasoning. *Child Development, 67*, 2873-2891.
- Klaczynski, P. A., & Narasimham, G. (1998). Representations as mediators of adolescent deductive reasoning. *Developmental Psychology, 5*, 865-881.
- Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*, 852 – 884.
- Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and asperger syndrome: The social attribution task. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*, 831-846.
- Kokis, J., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26-52.
- Krawczyk, D.C., Morrison, R.G., Viskontas, I., Holyoak, K.J., Chow, T.W., Mendez, M.F., Miller, B.L., & Knowlton, B.J. (2008). Distraction during relational reasoning: the role of prefrontal cortex in interference control. *Neuropsychologia, 46*, 2020-2032.
- Kuhn, D. (2001). How do people know? *Psychological Science, 12*, 1 – 8.

- Le Couteur, A., Baird, G. (2003). National Initiative for Autism: Screening and Assessment (NIASA). *National autism plan for children*. London: National Autistic Society. Available at [www.nas.org.uk/profess/niasa.html](http://www.nas.org.uk/profess/niasa.html)
- Lecoutre, M. P. (1985). Effect d'informations de nature combinatoire et de nature fréquentielle sur les jugements probabilistes. *Recherches en Didactique des Mathématiques*, 6, 193–213.
- Leevers, H. J. & Harris, P. L. (1999). Persisting effects of instruction on young children's syllogistic reasoning with incongruent and abstract premises. *Thinking and Reasoning*, 2, 145-173.
- Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, 43, 431–442.
- Leone, C., & Dalton, C. (1988). Some effects of the need for cognition on course grades. *Perceptual and Motor Skills*, 67, 175–178.
- Levisohn, L., Cronin-Golomb, A., & Schmahmann, J. D. (2000). Neuropsychological consequences of cerebellar tumour resection in children - cerebellar cognitive affective syndrome in a paediatric population. *Brain*, 123, 1041-1050.
- Ling, J., Burton, T. C., Salt, J. L., & Muncer, S. J. (2009). Psychometric analysis of the systemizing quotient (SQ) scale. *British Journal of Psychology*, 100, 539-552.
- Liu, I., Lo, K., & Wu, J. (1996). A probabilistic interpretation of "if-then." *Quarterly Journal of Experimental Psychology A*, 49, 828–844.
- Loftus, G.R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science*, 5, 161–171.
- Lopez, B., & Leekam, S.R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44, 285–300.

- Lutchmaya, S., Baron-Cohen, S. & Raggatt, P. (2002a). Foetal testosterone and eye contact in 12 month old infants. *Infant Behavior and Development*, 25, 327–335.
- Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002b). Foetal testosterone and vocabulary size in 18 and 24 month old infants. *Infant Behavior and Development*, 24, 418-424.
- Machery, E. (2008). Modularity and the flexibility of human cognition. *Mind & Language*, 23, 263-272.
- Manktelow, K. I. (1999). *Reasoning and thinking*. Hove, UK: Psychology Press.
- Markovits, H. (1995). Conditional reasoning with false premises: Fantasy and information retrieval. *British Journal of Developmental Psychology*, 13, 1–11.
- Markovits, H. (2000). A mental model analysis of young children’s conditional reasoning with meaningful premises. *Thinking & Reasoning*, 6, 335–347.
- Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review*, 22, 5–36.
- Markovits, H., Doyon, C. (2004). Information processing and reasoning with premises that are empirically false: Interference, working memory, and processing speed. *Memory & Cognition*, 32, 592-601.
- Markovits, H., & Dumas, C. (1999). Developmental patterns in the understanding of social and physical transitivity, *Journal of Experimental Child Psychology*, 73, 95-114.
- Markovits, H., Fleury, M., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child Development*, 69, 742–755.
- Markovits, H. & Potvin, F. (2001). Suppression of valid inferences and knowledge structures: The curious effect of producing alternative antecedents on reasoning with causal conditionals. *Memory and Cognition*, 29, 736-744.
- Markovits, H., & Quinn, S. (2002). Efficiency of retrieval correlates with “logical” reasoning from causal conditional premises. *Memory & Cognition*, 30, 696–706.

- Markovits, H., & Vachon, R. (1989). Reasoning with contrary-to-fact propositions. *Journal of Experimental Child Psychology, 47*, 398–412.
- Markovits, H., Venet, M., Janveau-Brennan, G., Malfait, N., Pion, N., & Vadeboncoeur, I. (1996). Reasoning in young children: Fantasy and information retrieval. *Child Development, 67*, 2857–2872.
- McKenzie, R., Evans, J. St. B. T. & Handley, S.J. (in press). Conditional reasoning in autism: activation and integration of knowledge and belief. *Developmental Psychology*.
- McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition, 102*, 179–218.
- McCloskey, M. (1983). Intuitive physics. *Scientific American, 248*, 122–130.
- McCrorry, E., Henry, L.A., Happé, F. (2007). Eye-witness memory and suggestibility in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry, 48*, 482-489.
- Mendel, N. (1998). *The intuitive rule “same of A, same of B”: The case of comparison of rectangles*. Unpublished master’s thesis, Tel Aviv University, Israel (in Hebrew).
- Metz, K. E. (1998). Emergent understanding and attribution of randomness: Comparative analysis of the reasoning of primary grade children and undergraduates. *Cognition and Instruction, 16*, 285–365.
- Minshew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry, 42*, 1095-1101.
- Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*, 123-142.



- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. (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. (1996) Control of mental processes. In: *Unsolved Mysteries of the Mind: Tutorial Essays in Cognition* (Bruce, V., ed.), pp. 93–148, Erlbaum.
- Morgan, B., Maybery, M. & Durkin, K. (2003). Weak central coherence, poor joint attention, and low verbal ability: Independent deficits in early autism. *Developmental Psychology, 39*, 646-656.
- Morris, A. K. (2000). Development of logical reasoning: Children's ability to verbally explain the nature of the distinction between logical and nonlogical forms of argument. *Developmental Psychology, 36*, 741–758.
- Morrison, R.G., Krawczyk, D.C., Holyoak, K.J., Hummel, J.E., Chow, T.W., Miller, B.L., & Knowlton, B.J. (2004). A neurocomputational model of analogical reasoning and its breakdown in Frontotemporal Lobar Degeneration. *Journal of Cognitive Neuroscience, 16*, 260–271.
- Moshman, D. (2000). Diversity in reasoning and rationality: Metacognitive and developmental considerations. *Behavioral and Brain Sciences, 23*, 689–690.
- Mottron, L. (2004) Matching strategies in Cognitive Research with Individuals with High-functioning Autism: Current Practices, Instrument Biases, and Recommendations. *Journal of Autism and Developmental Disorders*, special issue on methodology, *34*, 19-27.
- Mottron, L., Burack, J. A., Stauder, J.E.A., Robaey, P. (1999). Perceptual Processing among High-functioning Persons with Autism. *Journal of Child Psychology and Psychiatry, 40*, 203–211.

- Moutier, S., Plagne-Cayeux, S., Melot, A., & Houdé, O. (2006). Syllogistic reasoning and belief-bias inhibition in school children: Evidence from a negative priming paradigm. *Developmental Science, 9*, 166-172.
- Newstead, S.E. (2000). Are there two different types of thinking? *Behavioral and Brain Sciences, 23*, 690–691.
- Newstead, S.E., Handley, S. J., Harley, C., Wright, H. & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology, 57*, 33-60.
- O'Loughlin, C. & Thagard, P. (2000). Autism and coherence: A computational model. *Mind & Language, 15*, 375–392.
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin and Review, 11*, 988-1010.
- Osman, M., & Stavy, R. (2006). Intuitive rules: from formative to developed reasoning. *Psychonomic Bulletin & Review, 13*, 935-953.
- Ozonoff, S. (1995). Reliability and validity of the Wisconsin card sorting test in studies of autism. *Neuropsychology, 9*, 491–500.
- Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R., Klin, A., McMahon, W.M., Minshew, N., Munson, J.A., Pennington, B.F., Rogers, S.J., Spence, M.A., Tager-Flusberg, H., Volkmar, F.R., & Wrathall, D. (2004). Performance on CANTAB subtests sensitive to frontal lobe function in people with autistic disorder: Evidence from the CPEA network. *Journal of Autism and Developmental Disorders, 34*, 139-150.
- Ozonoff, S., & Jensen, J. (1999). Brief report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29*, 171-177.

- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals - relationship to theory of mind. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *32*, 1081-1105.
- Pacini, R. and Epstein, S., (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology* *76*, 972–987.
- Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: clues to developmental primacy. *Developmental Psychology*, *43*, 974-990.
- Pellicano, E., Maybery, M., & Durkin, K. (2005). Central coherence in typically developing preschoolers: Does it cohere and does it relate to mindreading and executive control? *Journal of Child Psychology and Psychiatry*, *46*, 533-547.
- Pellicano, E., Maybery, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: "weak" central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology*, *18*, 77-98.
- Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*, 337-344.
- Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child Development*, *73*, 752–767.
- Peterson, D., & Bowler, D. M. (2000). Counterfactual reasoning and false belief understanding in children with autism. *Autism*, *4*, 391-405.
- Pexman, P. M. (2008). It's fascinating research: The cognition of verbal irony. *Current Directions in Psychological Science*, *17*, 287-290.
- Piaget, J. (1976). *The grasp of consciousness*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1997). *How the mind works*. New York: Norton.

- Pijnacker, J., Geurts, B., van Lambalgen, M., Kan, C. C., Buitelaar, J. K., & Hagoort, P. (2009). Defeasible reasoning in high-functioning adults with autism: Evidence for impaired exception-handling. *Neuropsychologia*, *47*, 644-651.
- Plaisted, K. C. (2001). Reduced generalization in autism: An alternative to weak central coherence. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 149-172). Mahwah, NJ: Lawrence Erlbaum Associates.
- Politzer, G. , & Bourmaud, G. (2002). Deductive reasoning from uncertain premises. *British Journal of Psychology*, *93*, 345-381.
- Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education*, *31*, 602–625.
- Quinn, S., & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: Strength of association as a predictive factor for content effects. *Cognition*, *68*, B93-101.
- Quinn, S., & Markovits, H. (2002). Conditional reasoning with causal premises: Evidence for a retrieval model. *Thinking & Reasoning*, *8*, 179–191.
- Rajendran, G., & Mitchell, P. (2007). Cognitive Theories of Autism. *Developmental Review*, *27*, 224-260.
- Randall, L. (2005). *Warped passages*. New York: Ecco Press.
- Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: H. K. Lewis.
- Raven, J., Raven, J.C., & Court, J.H. (1998). *Manual for Raven's Advanced Progressive Matrices* (1998 edn.). Oxford: Oxford Psychologists Press.
- Reed, T. (1996). Analogical reasoning in subjects with autism, retardation, and normal development. *Journal of Developmental and Physical Disabilities*, *8*, 61–76.
- Reyna, V. F. (1996). Conceptions of memory development, with implications for reasoning and decision making. *Annual Review of Child Development*, *12*, 87-118.

- Reyna, V.F. (1996). Conceptions of memory development, with implications for reasoning and decision making. *Annals of Child Development*, R. Vasta, ed., 12, 87-118.
- Reyna, V.F., Brainerd, C.J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7, 1-75.
- Reyna, V.F. & Ellis, S.C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science*, 5, 275-279.
- Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making: Implications for theory, practice, and public policy. *Psychological Science in the Public Interest*, 7, 1-44.
- Reyna, V.F., Lloyd, F.J. & Brainerd, C.J. (2003). Memory, development, and rationality: An integrative theory of judgment and decision-making. In: S. Schneider and J. Shanteau, Editors, *Emerging perspectives on judgment and decision research*, Cambridge University Press, New York, 201-245.
- Richland, L.E., Morrison, R.G., & Holyoak, K.J. (2006). Children's development of analogical reasoning: insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94, 249-271.
- Riggs, K.J., Peterson, D.M., Robinson, E.J. & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 13, 73-90.
- Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., & Tonge, B. J. (2000). Atypical interference of local detail on global processing in high-functioning autism and asperger's disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 769-778.
- Roediger, H.L. & McDermott, K.M. (2000). Tricks of memory, *Current Directions in Psychological Science*, 9, 123-127.

- Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The windows task as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331-349.
- Rust, J., & Golombok, S. (Eds.). (1999). *Modern psychometrics: The science of psychological assessment* (2nd ed.). New York: Routledge.
- Roberts, M.J., & Newton, E.J. (2001). Inspection times, the change task, and the rapid-response selection task. *Quarterly Journal of Experimental Psychology, 54*, 1031-48.
- Rutter, M., Silberg, J., O'Connor, T., Simonoff, E. (1999). Genetics and child psychiatry: II empirical research findings. *Journal of Child Psychology and Psychiatry, 40*, 19-55.
- Sadowski, C. J., & Gulgoz, S. (1996). Elaborative processing mediates the relationship between the need for cognition and academic performance. *Journal of Psychology, 130*, 303-308.
- Saldaña, D., & Frith, U. (2007). Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology, 96*, 310-319.
- Sattler, J. M. (1992). *Assessment of children* (3rd ed.). San Diego: Jerome M. Sattler.
- Saxe, R., Schulz, L.E., Jiang, Y.V. (2006). Reading Minds versus Following Rules: Dissociating Theory of Mind and Executive Control in the Brain. *Social Neuroscience, 1*, 284-98.
- Schacter, D.L., & Tulving, E. (1994). What are the memory systems of 1994? In D.L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 1-38). Cambridge, MA: MIT Press.
- Schlottmann, A. (2000). Children's judgements of gambles: A disordinal violation of additive utility. *Journal of Behavioral Decision Making, 13*, 77-89.
- Schroyens, W., Schaeken, W., & Handley, S. J. (2003). In search of counter-examples: Deductive rationality in human reasoning. *Quarterly Journal of Experimental Psychology, 56A*, 1129-1145.

- Scott, F.J., & Baron-Cohen, S. (1996). Logical, analogical and psychological reasoning in autism: a test of the Cosmides theory. *Development and Psychopathology*, 8, 235–245.
- Sedlmeier, P., Hertwig, R., & Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 754–770.
- Shah, A., & Frith, U. (1983). An islet of ability in autistic-children - a research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 24, 613-620.
- Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 34, 1351-1364.
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal-lobe damage in man. *Brain*, 114, 727-741.
- Simoneau, M., & Markovits, H. (2003). Reasoning with premises that are not empirically true: Evidence for the role of inhibition and retrieval. *Developmental Psychology*, 39, 964–975.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.
- Spelke, E.S. (2000). Core knowledge. *American Psychologist*, 55, 1233–1243.
- Spelke, E.S. & Kinzler, K.D. (2007). Core knowledge. *Developmental Science*, 10, 89-96.
- Sperber, D., Cara, F. & Girotto, V. (1995). Relevance Theory explains the Selection Task. *Cognition*, 57, 31-95.
- Sperber, D. & Girotto, V. (2002). Use or misuse of the selection task? *Cognition*, 85, 277-290.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahweh, NJ: Erlbaum.

- Stanovich, K. E. (2003). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 291-342). New York: Cambridge University Press.
- Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: Chicago University Press.
- Stanovich, K. E. (2006). Fluid intelligence as cognitive decoupling. *Behavioral and Brain Sciences*, 29, 139-140.
- Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford: Oxford University Press.
- Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161-188.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-665.
- Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, 13, 225-247.
- Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, 94, 672-695.
- Stavy, R., & Tirosh, D. (2000). *How students (mis-)understand science and mathematics: Intuitive rules*. New York: Teachers College Press.
- Strough, J., Mehta, C. M., McFall, J. P., & Schuller, K. L. (2008). Do older and younger adults make different decisions about sunk costs? *Psychological Science*, 19, 650-652.
- Tentori, K.; Bonini, N., & Osherson, D. (2004). The conjunction fallacy: a misunderstanding about conjunction? *Cognitive Science*. 28, 467-477.



- Thomas, M. S. C. & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioral and Brain Sciences*, 25, 727-788.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22, 742-758.
- Thompson, V.A. (1996). Reasoning from false premises: The role of soundness in making logical deductions. *Canadian Journal of Experimental Psychology*. 50, 315-319.
- Toichi, M. & Kamio, Y. (2003). Long-Term Memory in High Functioning Autism: Controversy on Episodic Memory in Autism Reconsidered, *Journal of Autism and Developmental Disorders*, 33, 151-61.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind*, (pp. 19-136). New York: Oxford University Press.
- Torrens, D., Thompson, V. A., & Cramer, K. M. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking & Reasoning*, 5, 1-28.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105-110.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211, 453-458.
- Tversky, A. & Kahneman, D., (1982). Evidential impact of base rates. In: Kahneman, D., Slovic, P. and Tversky, A., Editors, 1982. *Judgment under uncertainty: heuristics and biases*, Cambridge University Press, New York, pp. 153-160.
- Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293-315.

- Vandierendonck, V., Kemps, E., Fastame, M.C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, *95*, 57–79.
- Venet, M. & Markovits, H. (2001). Understanding uncertainty with abstract conditional premises. *Merrill-Palmer Quarterly*, *47*, 74-99.
- Verschueren, N., De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Working memory capacity and the nature of generated counterexamples. *Proceedings of the Annual Conference of the Cognitive Science Society*, *24*, 914-919.
- Verschueren, N., Schaeken, W. & d'Ydewalle, G. (2005). Everyday conditional reasoning: a working memory-dependent tradeoff between counterexample and likelihood use. *Memory and Cognition*, *33*, 107-19.
- Viskontas, I.V., Morrison, R.G., Holyoak, K.J., Hummel, J.E., & Knowlton, B.J. (2004). Relational integration, inhibition and analogical reasoning in older adults. *Psychology and Aging*, *19*, 581–591.
- Waltz, J.A., Knowlton, B.J., Holyoak, K.J., Boone, K.B., Mishkin, F.S., de Menezes Santos, M., Thomas, C.R., & Miller, B.L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*, 119–125.
- Waltz, J.A., Lau, A., Grewal, S.K., & Holyoak, K.J. (2000). The role of working memory in analogical mapping. *Memory and Cognition*, *28*, 1205–1212.
- Wason, P., (1966). Reasoning. In: Foss, B.M., Editor, *New horizons in psychology*, Penguin, Harmondsworth.
- Wason, P. C. & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*, 141-154.
- Wason, P.C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, *23*, 63-71.
- Webley, P., & Plaisier, Z. (1998). Mental accounting in childhood. *Children's Social and Economics Education*, *3*, 55-64.

- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—third edition (WISC-III)*.  
San Antonio, TX: Psychological Corporation.
- Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development, 72*, 655–684.
- White, S., O'Reilly, H., & Frith, U. (2009). Big heads, small details and autism. *Neuropsychologia, 47*, 1274-1281.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-28.
- Wing, L. (1988). The continuum of autistic characteristics. In: Schopler E, Mesibov GB, eds. *Diagnosis and assessment in autism*. New York: Plenum Press.

**APPENDIX S1: Supplementary tables for Experiment 1.**

Table 1. Regression analyses on heuristic responding.

Predictors	Unique variance explained ( $R^2$ )	$\beta$ Weight	$t$ Value	Sig.
Age and cognitive ability entered simultaneously	.29	.05	.33	.75
1. Age	.19	.45	4.58	.0001
2. Cognitive ability	.29	.52	3.55	.001
1. Cognitive ability	.30	.56	6.07	.0001
2. Age	.29	.05	.33	.75

**APPENDIX S2: Supplementary tables for Experiment 2.**

Table 1. Correlations between the individual differences measures and the three response types across tasks.

		age	Cognitive ability	Inhibition	Set-shifting cost
Syllogistic reasoning	Heuristic responses	-.19	-.18	-.25 <sup>†</sup>	.07
	Normative responses	-.03	-.03	.09	.05
	“Other” responses	.30 <sup>*</sup>	.27 <sup>†</sup>	.24 <sup>†</sup>	-.17
Sunk cost fallacy	Heuristic responses	-.06	-.14	-.32 <sup>*</sup>	.08
	Normative responses	.30 <sup>*</sup>	.18	.22	-.27 <sup>†</sup>
	“Other” responses	-.30 <sup>*</sup>	-.07	.07	.10
If-only fallacy	Heuristic responses	.10	.18	.25 <sup>†</sup>	.08
	Normative responses	-.06	-.04	-.27 <sup>†</sup>	-.20
	“Other” responses	-.18	-.44 <sup>**</sup>	.00	.33 <sup>*</sup>
Conjunction fallacy	Heuristic responses	.29 <sup>*</sup>	.30 <sup>*</sup>	.15	.20
	Normative responses	-.29 <sup>*</sup>	-.30 <sup>*</sup>	-.15	-.20
Total	Heuristic responses	.07	.09	-.08	.21
	Normative responses	-.04	-.09	-.06	-.30 <sup>*</sup>
	“Other” responses	-.09	-.12	.15	.13

<sup>†</sup> correlation significant at the p<.10 level (Note that higher set-shifting cost indicates worse performance.)

Table 2. Intercorrelations between heuristic responding on the different reasoning and decision making tasks.

	1.	2.	3.	4.
1. Belief bias	--			
2. Sunk cost bias	.05	--		
3. If-only fallacy	-.02	-.18	--	
4. Conjunction fallacy	.07	-.03	.13	--

Table 3. Intercorrelations between normative responding on the different reasoning and decision making tasks.

	1.	2.	3.	4.
Syllogistic reasoning	--			
Sunk cost task	.06	--		
If-only task	-.09	-.01	--	
Conjunction fallacy task	.19	-.12	.06	--

Table 4. Regression analysis on normative responding.

Predictor	Unique variance explained ( $R^2$ )	$\beta$ Weight	$t$ Value	Sig.
Set-shifting	.11	-.33	4.37	.001



**APPENDIX S3: Supplementary tables for Experiment 3.**

Table 1. ANOVA analyses conducted on the judgment and decision-making tasks.

	F	Sig.	Partial Eta Squared
<i>Engineers and lawyers problem with a representative description</i>			
Conflict	149.10	.0001	.67
Conflict * age group	.06	.81	.001
Age group	.50	.48	.01
<i>Engineers and lawyers problem with a non-representative description /normative responses</i>			
conflict	30.51	.00	.35
conflict * age group	.20	.66	.003
age group	6.51	.01	.10
<i>Engineers and lawyers problem with a non-representative description / equally likely responses</i>			
conflict	16.04	.00	.22
conflict * age group	5.38	.02	.09
age group	.18	.67	.003
<i>Positive/negative recency effect</i>			
conflict	32.87	.00	.31
conflict * age group	1.27	.26	.02
age group	4.01	.049	.05
<i>Conjunction fallacy</i>			
conflict	80.49	.0001	.59
conflict * age group	2.13	.15	.04
age group	.37	.55	.01
<i>Syllogistic reasoning</i>			
conflict	72.25	.0001	.63
conflict * age group	1.84	.18	.04
age group	1.70	.20	.04

Table 2. ANOVA analyses conducted on the picture and scene analogy tasks.

	F	Sig.	Partial Eta Squared
<i>Picture analogy problems</i>			
distraction	100.95	.0001	.58
distraction * age group	.13	.72	.002
age group	.12	.73	.002
<i>Scene analogy problems</i>			
relational complexity	22.88	.0001	.24
relational complexity *			
age group	.16	.69	.002
distraction	48.92	.0001	.40
distraction* age group	.31	.58	.004
relational complexity *			
distraction	8.96	.004	.11
relational complexity *			
distraction * age group	1.75	.19	.02
age group	5.68	.02	.07

#### Appendix S4: Supplementary tables for Experiment 4

Table 1. ANOVA analyses on the heuristic, normative and “other” responses on the different reasoning tasks.

		F	Sig.	Partial Eta Squared
Syllogistic reasoning	<i>heuristic responses</i> group	6.58	.01	.09
	<i>normative responses</i> group	3.91	.052	.05
	<i>“other” responses</i> group	1.99	.16	.03
Sunk cost fallacy	<i>heuristic responses</i> group	.14	.71	.002
	<i>normative responses</i> group	4.89	.03	.07
	<i>“other” responses</i> group	18.13	.0001	.21
If-only fallacy	<i>heuristic responses</i> group	1.27	.26	.02
	<i>normative responses</i> group	.77	.38	.01
	<i>“other” responses</i> group	.67	.42	.01
Conjunction fallacy	<i>heuristic/normative responses</i> group	5.28	.03	.07

Table 2. Correlations between heuristic responding across tasks in the autistic group.

	1.	2.	3.	4.
1. belief bias	--			
2. sunk cost fallacy	.10	--		
3. if-only fallacy	-.05	-.08	--	
4. conjunction fallacy	.05	-.03	.15	--

Table 3. Correlations between normative responding across tasks in the autistic group.

	1.	2.	3.	4.
1. syllogistic reasoning	--			
2. sunk cost fallacy	.02	--		
3. if-only fallacy	-.07	.16	--	
4. conjunction fallacy	.15	.23	.13	--

**APPENDIX S5: Supplementary tables for Experiment 5.**

Table 1. ANOVA analyses conducted on the judgment and decision-making tasks.

	F	Sig.	Partial Eta Squared
<i>Engineers and lawyers problem with a representative description / base rate responses</i>			
Content	.33	.57	.01
Content * group	.89	.35	.02
Conflict	20.54	.0001	.29
Conflict * group	.30	.59	.01
Conflict * age	7.01	.01	.12
Content * conflict	.01	.91	.0001
Content * conflict * group	.01	.92	.0001
Group	.20	.66	.004
<i>Engineers and lawyers problem with a non-representative description / equally likely responses</i>			
Content	.98	.33	.02
Content * group	4.88	.03	.08
Conflict	3.51	.07	.06
Conflict * group	1.06	.31	.02
Content * conflict	.03	.86	.001
Content * conflict * group	.44	.51	.01
Group	.90	.35	.02
<i>Engineers and lawyers problem with a non-representative description / base rate responses</i>			
Content	.39	.53	.01
Content * group	.03	.87	.001
Conflict	15.47	.0001	.23
Conflict * group	4.45	.04	.08
Content * conflict	1.35	.25	.03
Content * conflict * group	.85	.36	.02
Group	1.76	.19	.03
<i>Positive/negative recency / normative responses</i>			
Content	.17	.68	.003
Content * group	.17	.68	.003
Conflict	16.12	.0001	.24
Conflict * group	.40	.53	.01
Content * conflict	3.73	.06	.07

	F	Sig.	Partial Eta Squared
<i>Engineers and lawyers problem with a representative description / base rate responses</i>			
Content	.33	.57	.01
Content * group	.89	.35	.02
Conflict	20.54	.0001	.29
Conflict * group	.30	.59	.01
Conflict * age	7.01	.01	.12
Content * conflict	.01	.91	.0001
Content * conflict * group	.01	.92	.0001
Group	.20	.66	.004
<i>Engineers and lawyers problem with a non-representative description / equally likely responses</i>			
Content	.98	.33	.02
Content * group	4.88	.03	.08
Conflict	3.51	.07	.06
Conflict * group	1.06	.31	.02
Content * conflict	.03	.86	.001
Content * conflict * group	.44	.51	.01
Group	.90	.35	.02
<i>Engineers and lawyers problem with a non-representative description / base rate responses</i>			
Content	.39	.53	.01
Content * group	.03	.87	.001
Conflict	15.47	.0001	.23
Conflict * group	4.45	.04	.08
Content * conflict * group	.81	.37	.02
Group	2.38	.13	.05
<i>Positive recency effect</i>			
Content	6.43	.01	.11
Content * group	1.67	.20	.03
Conflict	14.00	.0001	.22
Conflict * group	.84	.37	.02
Content * conflict	17.86	.0001	.26
Content * conflict * group	3.88	.06	.07
Group	.47	.50	.01

	F	Sig.	Partial Eta Squared
<i>Negative recency effect</i>			
Content	8.18	.006	.14
Content * group	4.25	.04	.08
Conflict	3.31	.08	.06
Conflict * group	.001	.98	.0001
Content * conflict	8.13	.01	.14
Content * conflict * group	.97	.33	.02
Group	2.98	.09	.06
<i>Conjunction fallacy / heuristic and normative responses</i>			
Content	.42	.52	.01
Content * group	1.35	.25	.03
Conflict	8.85	.004	.14
Conflict * group	4.36	.04	.08
Content * conflict	2.02	.16	.04
Content * conflict * group	.31	.58	.01
Group	.68	.42	.01



Table 2. ANOVA analyses conducted on the analogical reasoning tasks.

	F	Sig.	Partial Eta Squared
<i>Picture analogy tasks</i>			
Distraction	88.46	.0001	.63
Distraction * group	.56	.46	.01
Group	.01	.94	.0001
<i>Scene analogy tasks</i>			
Relational complexity	22.88	.0001	.31
Relational complexity * group	.12	.73	.002
Distraction	28.77	.0001	.36
Distraction * group	.40	.53	.01
Relational complexity * distraction	10.23	.002	.17
Relational complexity * distraction * group	.40	.53	.01
Group	.30	.59	.01

Table 3. Analyses of error rates on the analogical reasoning tasks.

		<i>t</i> (53)	Sig. (2-tailed)
Picture analogy test	perceptual	.18	.86
	semantic	.67	.51
	other	.75	.46
Scene analogy test	perceptual	.91	.37
	relational	.11	.91
	other	.06	.95

**APPENDIX S6: Supplementary tables for Experiment 6.**

**Table 1. ANOVA analyses on the representative, equiprobability and normative responses with need for cognition as a covariate.**

		F	Sig.	Partial Eta Squared
Representativeness heuristic	content	.09	.77	.0001
	content * need for cognition	.06	.81	.0001
	content * year group	2.37	.13	.013
	content * subject group	.02	.89	.0001
	content * year group * subject group	.003	.96	.0001
	need for cognition	4.12	.044	.02
	year group	6.39	.012	.04
	subject group	.001	.976	.0001
	year group * subject group	11.61	.001	.06
Equiprobability bias	content	.41	.52	.002
	content * need for cognition	.02	.89	.0001
	content * year group	3.05	.08	.02
	content * subject group	1.02	.31	.01
	content * year group * subject group	.01	.93	.0001
	need for cognition	6.95	.009	.04
	year group	1.43	.23	.01
	subject group	.79	.38	.004
	year group * subject group	4.36	.04	.02
Normative responses	content	1.18	.28	.007
	content * need for cognition	.20	.66	.001
	content * year group	3.80	.05	.02
	content * subject group	.53	.47	.003
	content * year group * subject group	.01	.92	.0001
	need for cognition	13.24	.0001	.07
	year group	.57	.45	.003
	subject group	.41	.52	.002
	year group * subject group	.0001	.99	.0001

Table 2. Regression analyses on the representativeness-based and equiprobability responses

Predictors		Unique variance explained ( $R^2$ )	$\beta$ Weight	$t$ Value	Sig.
<i>Analyses on the representativeness-based responses:</i>					
First year psychology students	Normative potential	-.02	-.09	.58	.56
Higher year psychology students	Statistics marks	.05	-.25	2.08	.04
	1. Statistics marks	.10	-.17	1.29	.20
	2. Normative potential		-.21	1.62	.11
First year biology students	Need for cognition	.03	-.24	1.34	.19
Higher year biology students	Statistics marks	-.03	-.08	.43	.67
	Need for cognition	-.01	-.13	.81	.42
<i>Analyses on the equiprobability responses:</i>					
First year psychology students	Normative potential	-.02	-.04	.26	.79
Higher year psychology students	Statistics marks	.09	-.32	2.69	.01
	1. Statistics marks	.14	-.22	1.70	.10
	2. Normative potential		-.27	2.12	.04
First year biology students	Need for cognition	.03	-.08	.44	.66
Higher year biology students	Statistics marks	.0001	-.01	.06	.95
	Need for cognition	.02	-.20	1.31	.20

**APPENDIX S7: Supplementary tables for Experiment 7.**

Table 1. ANOVA analyses on the representative, equiprobability and normative responses.

		F	Sig.	Partial Eta Squared
Representativeness	content	.06	.81	.0001
heuristic	content * year group	.30	.74	.01
	year group	.68	.51	.01
Equiprobability	content	31.63	.0001	.20
bias	content * year group	1.11	.33	.02
	year group	3.34	.04	.05
Normative	content	15.83	.0001	.12
responses	content * year group	.54	.59	.01
	year group	.32	.73	.01

Table 2. Regression analyses on the representativeness-based and equiprobability responses.

Predictors		Unique variance explained ( $R^2$ )	$\beta$ Weight	$t$ Value	Sig.
<i>Analyses on the representativeness-based responses:</i>					
First year psychology students	Normative potential	.01	10	.62	.54
Higher year psychology students	Statistics marks Normative potential	.01 .02	.12 .13	.86 1.14	.39 .26
<i>Analyses on the equiprobability responses:</i>					
First year psychology students	Normative potential	.01	-.19	1.28	.21
Higher year psychology students	Statistics marks Normative potential	.01 .05	-.11 -.24	.82 2.22	.42 .03

**APPENDIX S8: Supplementary tables for Experiment 8.**

Table 1. ANOVA analyses on the representative, equiprobability and normative responses with need for cognition as a covariate.

		F	Sig.	Partial Eta Squared
Representativeness heuristic	content	.82	.37	.01
	content * normative potential group	2.32	.13	.02
	content * instructions	.004	.95	.0001
	content * normative potential group * instructions	1.61	.21	.01
	normative potential group	.17	.68	.001
	instructions	.99	.32	.01
	normative potential group * instructions	2.707	.10	.023
	Equi probability bias	content	40.72	.0001
content * normative potential group	1.37	.25	.01	
content * instructions	1.71	.19	.02	
content * normative potential group * instructions	5.89	.02	.05	
normative potential group	.98	.32	.01	
instructions	.91	.34	.01	
normative potential group * instructions	.35	.55	.003	
Normative responses	content	17.74	.0001	.13
	content * normative potential group	.03	.87	.0001
	content * instructions	3.74	.06	.03
	content * normative potential group * instructions	.05	.83	.0001
	normative potential group	.89	.35	.01
	instructions	.30	.58	.003
	normative potential group * instructions	.55	.46	.01

Table 2. ANOVA analyses on the equiprobability responses for the high and low normative potential groups

		F	Sig.	Partial Eta Squared
Low normative potential	Content	11.12	.002	.16
	Content * instructions	.52	.48	.01
	Instructions	.06	.80	.001
High normative potential	Content	36.58	.0001	.39
	Content * instructions	8.94	.004	.14
	Instructions	1.22	.27	.02