



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

Copyright statement

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

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

Liability statement

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

Access Agreement

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

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

Why it happened and how it could have been different:

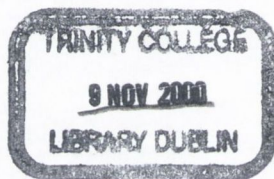
A comparison of causal and counterfactual thinking

Alice McEleney

Ph. D.

University of Dublin, Trinity College

2000



Thesis
5891

Declaration

- (a) The work contained in this thesis has not been submitted as an exercise for a degree at this or at any other University.
- (b) This thesis is the result of my own investigations, except where otherwise stated. The contributions of others are duly acknowledged in the text wherever included.
- (c) I give permission to the Library to lend or copy this thesis upon request.

Signed: Alice McElaney

Date: 20 / 09 / 2000

Acknowledgements

My biggest word of thanks goes to Ruth Byrne for being an outstanding supervisor. She has not only provided excellent advice and criticism from start to finish, she has also maintained an enthusiasm for the research which has been my constant source of encouragement. I could not have hoped for a better mentor.

Special thanks also go to Vittorio Girotto for his collaboration and support during and after my three month research visit to the University of Provence.

Thanks to everyone else at the Psychology Department, University of Dublin, and the Laboratory of Cognitive Psychology, University of Provence, especially Rachel McCloy, Guylaine Molinaire and Clare Walsh for being perfect officemates, and all the other postgraduates for their support; Jean-Paul Caverni, Emma Riggs and Sandrine Rossi for practical assistance; Lisa Cullen for technical assistance; Abdessadek El Ahmadi, Michael Gormley, Mick O'Connell, Jean Quigley and Howard Smith for help with statistics; Orlando Espino, Juan García-Madruga, Yevgeniya Goldvarg and Michel Gonzalez for discussions of experiments; and Hugh Garavan, Fiona Newell, Shane O'Mara and Ian Robertson for helpful advice.

Thanks to everyone else who showed an interest in the research, especially Fintan Costello, Aidan Feeney, Phil Johnson-Laird, Mark Keane, Paolo Legrenzi, David Over, Antonio Rizzo, Maria Sonino-Legrenzi, Rosemary Stevenson, Valerie Thompson and Carl Vogel.

Thanks to everyone who took part in the experiments and pilot studies, which were run at the University of Dublin, University College Dublin, the University of Provence, and the University of Trieste.

Thanks to my sources of funding: the University of Dublin, Enterprise Ireland, and the Government of Ireland Humanities and Social Sciences Council.

Finally, thanks to my friends and housemates for their encouragement, especially Deirdre Bonini, Sarah Clarke, Deirdre Feeney, Kay Maunsell, Bryan McEleney and Karen Smyth.

The thesis is dedicated to the memory of my father, James Patrick McEleney.

Table of Contents

	Page
Chapter 1 Introduction	
Causal and counterfactual thinking	1
The relation between causal and counterfactual thinking	3
The activation of causal and counterfactual thoughts	14
The content of causal and counterfactual thoughts	22
The consequences of causal and counterfactual thoughts	41
Understanding and reasoning from causal and counterfactual assertions	51
Summary and aims of the thesis	64
 Chapter 2 The Activation and Content of Causal and Counterfactual Thoughts	
Introduction	67
Experiment 1	70
Experiment 2	82
Experiment 3	89
General discussion	99
 Chapter 3 The Consequences of Causal and Counterfactual Thoughts	
Introduction	104
Experiment 4	106
Experiment 5	112
Experiment 6	121

(h) Experiment 8	260
(i) Experiment 9	265
Appendix 2 Sample causal and counterfactual thoughts from Experiment 1	271
Appendix 3 Experiment 6 t-tests	272
Appendix 4 Experiment 7 content effects	277
Appendix 5 List of publications	278

List of tables

Table	Title	Page
Table 1.1	The four main conditional inferences, illustrated for the conditional, “If John studies hard, then he gets good marks”	59
Table 2.1	Frequencies (with means per participant in parentheses) of causal and counterfactual thoughts by condition in Experiment 1	77
Table 2.2	Percentages of causal and counterfactual thoughts in each content category in Experiment 1	80
Table 2.3	Frequencies (with means per participant in parentheses) of causal and counterfactual thoughts by condition in Experiment 2	88
Table 2.4a	Percentages of participants who rated the covarying or non-covarying event and the controllable or uncontrollable event as more causal in Experiment 3	96
Table 2.4b	Percentages of participants who rated the time or the route as more causal for each scenario in Experiment 3	96
Table 2.5	Mean causal ratings and percentages of counterfactual mutations of events in Experiment 3	97

Table 2.6a	Percentages of participants whose counterfactuals mutated the covarying or non-covarying event and the controllable or uncontrollable event earlier in Experiment 3	97
Table 2.6b	Percentages of participants whose counterfactuals mutated the time or the route earlier for each scenario in Experiment 3	98
Table 3.1	Mean ratings of preventability, controllability and predictability following different sentence completion tasks in Experiment 4	110
Table 3.2	Mean predictability ratings and likelihood judgements following different sentence completion tasks in Experiment 5	118
Table 3.3	Mean preventability, predictability, future likelihood and emotion scores following different thought listing tasks in Experiment 6	130
Table 4.1	Ten interpretations of the conditional “if p then q” (adapted from Johnson-Laird & Byrne, 2000)	141
Table 4.2	Percentages of ten interpretations of causal, counterfactual, prediction and prevention assertions in Experiment 7	148
Table 4.3	Percentages of participants making each of the inferences from each assertion in Experiment 8	162

Table 4.4	Mean latencies (in milliseconds) to endorse each of the inferences from each assertion in Experiment 9	168
Table 4.5	Mean log transformed latencies (in milliseconds) to endorse both the modus ponens and denial of the antecedent inferences from each assertion in Experiment 9	169
Table 4.6	Percentages of participants making each of the inferences from each assertion in Experiment 9	173

Summary

The aim of this thesis was to compare two pervasive forms of human thinking: causal thinking about why an outcome happened (e.g., “I failed because I didn’t try”) and counterfactual thinking about how an outcome could have been different (e.g., “I wouldn’t have failed if I had tried”). We report nine experiments designed to test the hypothesis that causal thoughts are concerned with the prediction of outcomes, whereas counterfactual thoughts are concerned with their prevention.

In our first series of experiments, we compared the activation and content of causal and counterfactual thoughts. Experiments 1 and 2 showed that spontaneous causal and counterfactual thoughts are both activated by bad rather than good outcomes, causal but not counterfactual thoughts are evoked by unexpected rather than expected outcomes, counterfactual but not causal thoughts are activated by controllable rather than uncontrollable outcomes, and causal but not counterfactual thoughts are evoked by outcomes with unusual rather than normal antecedents. The contents of spontaneous causal and counterfactual thoughts also differ in systematic ways: causal thoughts are more likely than counterfactual thoughts to focus on general factors rather than specific antecedents, whereas counterfactual thoughts are more likely than causal thoughts to focus on internal rather than external factors, and inhibitory rather than facilitative antecedents. These findings generalise from spontaneous to directed thoughts, as Experiment 3 showed: directed causal thoughts focus on covarying and uncontrollable antecedents, whereas directed counterfactual thoughts focus on covarying and controllable antecedents.

Our second series of experiments compared the consequences of causal and counterfactual thoughts. In Experiment 4, counterfactual but not causal thoughts increased the perceived preventability and controllability of an outcome, but neither causal nor counterfactual thoughts affected the perceived predictability of the outcome. In Experiment 5, counterfactual but not causal thoughts decreased perceived predictability. In Experiment 6, causal thoughts

somewhat increased judgements about prediction compared to counterfactual thoughts. Both causal thoughts and upward counterfactual thoughts (about how an outcome could have been better) increased judgements about prevention, whereas downward counterfactual thoughts (about how an outcome could have been worse) had no effect. Downward counterfactual thoughts but not causal or upward counterfactual thoughts enhanced emotions.

Our final series of experiments compared how people understand and reason from causal and counterfactual assertions, which we compared to prediction and prevention assertions. Experiment 7 showed that people tend to interpret all four sorts of assertions as biconditionals: participants judged only the two situations in which both the antecedent and consequent occurred, or neither occurred, to have been possible given any one of the assertions. However, Experiments 8 and 9 suggested that people *initially* keep in mind only the former possibility for causal and prediction assertions, but both possibilities for counterfactual and prevention assertions. In Experiment 8, participants made more affirmative inferences (i.e., modus ponens) than negative inferences (i.e., denial of the antecedent) from affirmative causal and prediction assertions, but not from affirmative counterfactual and prevention assertions. In Experiment 9, participants made affirmative inferences faster than negative inferences from affirmative causal and prediction assertions, but not from affirmative counterfactual and prevention assertions.

The overall findings support the view that causal thoughts are concerned with prediction, whereas counterfactual thoughts are concerned with prevention. We suggest that people can generate causal assertions directly from their initial mental models of causal relations, whereas they must first flesh out their models in order to generate counterfactual assertions. As a result, counterfactual thinking may be more difficult, but also more instructive, than causal thinking. We propose that causal and counterfactual thinking are distinct activities which may depend on the same sorts of cognitive processes, and which may be united by the common goal of maintaining a sense of control.

Chapter 1 Introduction

The aim of this thesis is to compare two pervasive forms of human thinking: causal thinking about why an outcome happened, and counterfactual thinking about how an outcome could have been different.

Causal and counterfactual thinking

People engage in *causal thinking* when they explain past and present outcomes, such as when they recognise that the sudden movement of a football occurred because it was kicked (e.g., Michotte, 1946), when they attribute a person's quietness to their shy personality (e.g., Kelly, 1972a), when they judge that a space shuttle exploded due to faulty seals on the booster rockets (e.g., Hilton & Erb, 1996), when they decide that an individual committed a crime on impulse (e.g., Carroll & Payne, 1977), when they understand the relations between events in a story (e.g., Trabasso & Sperry, 1985) and when they realise that their rising body temperature is a result of jealousy (e.g., Schachter & Singer, 1962). People also engage in causal thinking when they predict future outcomes from antecedent conditions, such as when they infer that drinking more wine will cause them to have a hangover, or that using insect repellent will prevent them from getting bitten (e.g., Goldvarg & Johnson-Laird, 2000).

People engage in *counterfactual thinking* when they imagine how past and present outcomes could have been, or could be, different, such as when they dwell on how a child's death in a car accident could have been avoided *if only* he hadn't been let out (e.g., Davis, Lehman, Wortman, Silver & Thompson, 1995), when they wonder how the events of history would have turned out differently if Hitler had won World War Two (e.g., Tetlock & Belkin,

1996), when they consider how their present lives would be better if they had chosen a different career, or if they had never got married (e.g., Landman & Manis, 1992) and when they think about alternatives to more mundane everyday outcomes, such as being late for an appointment. Counterfactual thoughts may often be expressed as conditionals (i.e., “if...then...” assertions). The defining feature of a *counterfactual* conditional is that its antecedent and consequent are both false¹, and in English this is usually indicated by the use of the subjunctive mood. For example, “If I had driven fast, then I would have been on time” suggests that in reality, I drove slowly and I was late² (e.g., Carpenter, 1973; Fillenbaum, 1974). In contrast, a conditional in the indicative mood, such as “If I drove fast, then I was on time”, conveys no information about the truth or falsity of its constituent propositions. The aim of counterfactual thinking is typically to mentally *undo* a past or present outcome³, such as being late, by *mutating* antecedent events, such as driving slowly (e.g., Kahneman & Miller, 1986; Kahneman & Tversky, 1982).

In this thesis, we will test the hypothesis that causal and counterfactual thinking serve different functions: causal thinking is concerned with the prediction of outcomes, whereas counterfactual thinking is concerned with their prevention. We will report nine experiments which compare systematically the determinants of causal and counterfactual thinking: three

¹ The term “counterfactual” is also sometimes used to refer to conditionals in which only the antecedent is false, e.g., “Even if I had driven fast, then I would have been late”. However, such conditionals are usually termed *semifactuals*, rather than counterfactuals, because only one of their constituent propositions is counterfactual (e.g., Goodman, 1954/1983; McCloy, 2000; McCloy & Byrne, 1999). Our use of the term counterfactual does not include semifactuals.

² The same conditional in the subjunctive mood and the present tense, “If I were to drive fast, then I would be on time”, is not strictly counterfactual, since its constituent propositions could turn out to be true. However, people seem to understand both past and present tense subjunctive conditionals as counterfactuals (Byrne & Tasso, 1999), and so we will refer to both as counterfactuals.

³ People also think about how anticipated *future* outcomes could be different under different antecedent conditions, e.g., “If I were to drive fast tomorrow, then I would be on time”. However, such thoughts may be better termed *prefactuals* rather than counterfactuals, because they are “counter only to an imagined experience” (Gleicher, Boninger, Strathman, Armor, Hetts & Ahn, 1995, p. 294; see also Roese & Olson, 1993). Our use of the term counterfactual does not include prefactuals.

experiments which compare the activation and content of causal and counterfactual thoughts (Chapter 2), three experiments which compare the consequences of causal and counterfactual thoughts (Chapter 3), and three experiments which compare how people understand and reason from causal and counterfactual assertions (Chapter 4). This introductory chapter is divided into six sections: the first section describes the recent debate on the relation between causal and counterfactual thinking, the following four sections review past studies of the determinants of causal and counterfactual thinking, and the final section outlines the aims of our research.

The relation between causal and counterfactual thinking

The link between causal and counterfactual thinking has been recognised by philosophers for centuries. For example, Hume (1748/1999) suggested that “we may define a cause to be an object, followed by another, and where all objects, similar to the first, are followed by objects similar to the second. Or in other words, where, *if the first object had not been, the second never had existed*” (p. 146, italics added). Likewise, Mill (1843/1996) proposed that the identification of causes may be based on the contrast between a factual situation and a counterfactual alternative. For example, “if a person eats of a particular dish, and dies in consequence, that is, *would not have died if he had not eaten of it*, people would be apt to say that eating of that dish was the cause of his death” (p. 327, italics added). More recently, the relation between causal and counterfactual thinking has been explored by psychologists (e.g., Einhorn & Hogarth, 1986; Kahneman & Tversky, 1982; Lipe, 1991; Roese, 1997; Roese & Olson, 1995; 1997; Spellman & Mandel, 1999; Wells & Gavanski, 1989) and legal analysts (e.g., Hart & Honoré, 1959/1985), as well as philosophers (e.g., Goodman, 1954/1983; Lewis, 1973a; 1973b; Mackie, 1974; Stalnaker, 1968).

Recent psychological research suggests that causal and counterfactual thinking share similarities (e.g., Lipe, 1991; Wells & Gavanski, 1989) and differences (e.g., Mandel & Lehman, 1996; N'gbala & Branscombe, 1995), but the precise nature of the relation between causal and counterfactual thinking remains unclear (e.g., Roese & Olson, 1997). In the rest of this section, we will first consider the arguments for similarities in causal and counterfactual thinking, second, review the evidence for differences in causal and counterfactual thinking, and finally, describe two opposing views of the relation between causal and counterfactual thinking.

Similarities in causal and counterfactual thinking

Causal and counterfactual thinking clearly share some basic similarities. One similarity is that both activities seem to be universal parts of human mental life. Both causal and counterfactual thinking develop at an early age. Pre-school children are adept at explaining random events by citing possible invisible causes (e.g., Gelman & Kalish, 1993). Likewise, from the age of two, children express counterfactual thoughts about events that almost happened, their wishes for counterfactual situations, and previous misperceptions that counterfactual events had actually occurred (e.g., Bowerman, 1986). Moreover, neither causal nor counterfactual thinking depend on a particular language or culture. Weiner (1985a) cites several examples of causal thinking in very different times and places, including a seventeenth century Japanese warrior's contemplation of the causes of his success in battle (Musashi, 1645/1974, p. 35). Perhaps the earliest documentation of the use of counterfactual thinking is in the works of the early Greek philosophers, who often employed counterfactuals to explain ideas about the nature of reality (see Rescher, 1995). Even speakers of languages that do not have a distinct counterfactual construction, such as Chinese, seem to have no difficulty with counterfactual thinking (Au,

Recent psychological research suggests that causal and counterfactual thinking share similarities (e.g., Lipe, 1991; Wells & Gavanski, 1989) and differences (e.g., Mandel & Lehman, 1996; N'gbala & Branscombe, 1995), but the precise nature of the relation between causal and counterfactual thinking remains unclear (e.g., Roese & Olson, 1997). In the rest of this section, we will first consider the arguments for similarities in causal and counterfactual thinking, second, review the evidence for differences in causal and counterfactual thinking, and finally, describe two opposing views of the relation between causal and counterfactual thinking.

Similarities in causal and counterfactual thinking

Causal and counterfactual thinking clearly share some basic similarities. One similarity is that both activities seem to be universal parts of human mental life. Both causal and counterfactual thinking develop at an early age. Pre-school children are adept at explaining random events by citing possible invisible causes (e.g., Gelman & Kalish, 1993). Likewise, from the age of two, children express counterfactual thoughts about events that almost happened, their wishes for counterfactual situations, and previous misperceptions that counterfactual events had actually occurred (e.g., Bowerman, 1986). Moreover, neither causal nor counterfactual thinking depend on a particular language or culture. Weiner (1985a) cites several examples of causal thinking in very different times and places, including a seventeenth century Japanese warrior's contemplation of the causes of his success in battle (Musashi, 1645/1974, p. 35). Perhaps the earliest documentation of the use of counterfactual thinking is in the works of the early Greek philosophers, who often employed counterfactuals to explain ideas about the nature of reality (see Rescher, 1995). Even speakers of languages that do not have a distinct counterfactual construction, such as Chinese, seem to have no difficulty with counterfactual thinking (Au,

1984).

A second similarity is that both causal and counterfactual thinking seem to be important for successful interaction with the world. The usefulness of causal thinking has long been recognised by both philosophers and psychologists. For example, according to Hume (1748/1999, p. 121), "All reasonings concerning matters of fact seem to be founded on the relation of cause and effect". One way in which people may acquire causal knowledge is by explaining past and present outcomes (for a review, see Weiner, 1985b). The resulting causal information is necessary for many basic tasks, such as understanding the behaviour of others and representing sequences of actions (for a review, see Corrigan & Denton, 1996), and for more sophisticated processes, such as the use of analogies (e.g., Goswami, 1992; Keane, 1997). Moreover, such knowledge may allow people to predict and control future outcomes (e.g., Heider, 1958; Kelly, 1972a). More recently, experimental evidence has shown that counterfactual thoughts about how a past outcome could have been better help people to improve their subsequent performance (Nasco & Marsh, 1999; Roese, 1994). Such thoughts may be useful because they are converted into scripts for future action, which encourage behaviours that facilitate success (Roese, 1994). Some neuropsychologists have even suggested that the inflexible behaviour of patients with frontal lobe injuries may result in part from their impaired ability to engage in counterfactual thinking (Knight & Grabowecky, 1995).

The conversion of counterfactual thoughts about the past into useful prescriptions for future behaviour may depend on the relation between causal and counterfactual thinking (e.g., Roese & Olson, 1997). When people think about a counterfactual conditional, such as, "If I had driven fast, then I would have been on time", they may keep in mind not only the hypothesised counterfactual situation (in which I drove fast and I was on time), but also the presupposed

factual situation (in which I drove slowly and I was late). In contrast, when people think about an indicative conditional such as “If I drove fast, then I was on time”, they may keep in mind only the stated hypothetical situation (in which I drove fast and I was on time). A counterfactual conditional may require people to keep two explicit situations in mind, the counterfactual and the factual, whereas an indicative conditional may require them to keep just one explicit situation in mind, and to keep in mind that there may be alternatives to this situation, which they need not have thought out explicitly (Johnson-Laird & Byrne, 1991). The more explicit representation of counterfactual conditionals ensures that people make more of certain sorts of inferences from them than from indicative conditionals, even the more difficult *modus tollens* inference (from ‘I was late’ to the conclusion ‘I drove slowly’, e.g., Byrne & Tasso, 1999). They even tend to mistakenly recognise the factual situation as one that was described to them when in fact the counterfactual situation had been (Fillenbaum, 1974).

A possible result of understanding a counterfactual conditional by keeping in mind both the factual and the counterfactual situations is that the contrast between them may have consequences for causal inference. Mill’s (1843/1996) “method of difference” for the identification of causes states that if the only difference between two occurrences with different outcomes is the presence or absence of a specific antecedent, this antecedent can be judged to be a cause of the outcome. For example, when people think about the counterfactual conditional, “If Bob hadn’t left the window open, then the flies wouldn’t have got into the house”, they keep in mind that in reality, Bob *did* leave the window open and the flies *did* get into the house. The contrast between the two situations may in turn suggest that Bob’s leaving the window open *caused* the flies to get into the house (Roese & Olson, 1997, pp. 6-7). Several philosophers have suggested that the concept of causation involves the dependence of an outcome on preceding

events (e.g., Hume, 1739/1978; Lewis, 1973a; 1973b; Mackie, 1974; Mill, 1843/1996). For example, according to Lewis (1973a, p. 557), "we think of a cause as something that makes a difference...from what would have happened without it. Had it been absent, its effects...would have been absent as well". Lewis concludes that causation can be defined as counterfactual dependence: "if *c* and *e* are two actual events such that *e* would not have occurred without *c*, then *c* is a cause of *e*" (p. 563).

On this account, the psychological usefulness of counterfactual thinking may depend on the ability of counterfactual thoughts to convey causal information. The causal conclusions reached from counterfactuals may result in expectancies for the consequences of particular actions, for example, that leaving the window open will cause flies to get into the house. These expectancies should influence people's intentions for future action, for example, intending not to leave the window open, and the intentions may in turn determine their future behaviour. If the original causal conclusions were accurate, then performance should be improved by this process (Roese, 1997; Roese & Olson, 1997). This account suggests that causal and counterfactual thinking may be very closely related. Some theorists argue that causal judgements can be made using only counterfactual information (e.g., Lipe, 1991), or that counterfactual thinking serves as a heuristic for causal judgement (e.g., Kahneman & Tversky, 1982; Wells & Gavanski, 1989). For example, if a person answers "no" to the counterfactual question, "Would ABC Company have lost money if the economy had been good?", then they should infer that the poor economy caused ABC Company to lose money (Lipe, 1991, p. 458). It may even be questioned whether causal and counterfactual thinking are distinct processes. Roese (1997, p. 144) suggests that "some might argue that counterfactual and causal statements are not separate constructs, but semantically identical, differing only in wording. Thus, asserting that "If not for Watergate,

Nixon would not have resigned” expresses exactly the same idea as “Watergate caused Nixon’s resignation”. On the basis of this argument, the entire counterfactual literature might be viewed as an elaborate reformulation of traditional attribution theories.”

Evidence that causal and counterfactual thinking are similar comes from empirical studies which show that the availability of a counterfactual alternative to an outcome affects people's causal ratings of the events that preceded the outcome (Roese & Olson, 1996; Wells & Gavanski, 1989). For example, Wells and Gavanski gave participants a scenario about a couple who had to drive their own car to a restaurant because they were refused a lift by a taxi driver. On the way, they accidentally drove off a collapsed bridge and were killed. In one version of the scenario, the taxi made it across the bridge safely before it collapsed, and in the other version the taxi also drove off the bridge. Participants given the first version -- in which a different decision by the driver would counterfactually undo the deaths of the couple -- rated the driver’s decision as more causal of the outcome than those given the second version -- in which changing the driver’s decision would not undo the outcome. When participants were asked to produce counterfactual thoughts about how the outcome could have been avoided before they completed the causal ratings, their judgements of the causality of the driver's decision in the first version were even stronger. The results show that counterfactual thinking influences causal judgements.

Differences in causal and counterfactual thinking

An alternative view is that causal and counterfactual thoughts are quite different. Some philosophers have criticised counterfactual analyses of causation for their limited ability to account for how counterfactual and causal statements are actually used in conversation (e.g., Berofsky, 1973; Kim, 1973; Lombard, 1990; Swain, 1978). One objection is that theories of

causation cannot provide a complete account of the use of counterfactuals. For example, a counterfactual conditional such as, "if there had been a circle on the blackboard, then there would have been a triangle" suggests that in fact, there was no circle and there was no triangle (e.g., Byrne & Tasso, 1999). The contrast between the counterfactual and factual situations may suggest that the absence of a circle was necessary for the absence of a triangle. But it seems mistaken to suggest that people will infer that the absence of a circle *caused* the absence of a triangle. Clearly, not all counterfactual conditionals are concerned with causal relations. Counterfactuals can also express other sorts of relations, including for example, inferential relations, e.g., "if the number hadn't been divisible by 2, then it wouldn't have been even", and deontic relations, e.g., "if we hadn't promised, then we needn't have gone" (for further discussion, see Johnson-Laird & Byrne, 1991, pp. 70-71).

A possible response to this problem is to limit our discussion to counterfactuals whose content and context suggest that they are concerned with causal relations. However, even in such cases, counterfactuals do not provide a complete analysis of causation. A counterfactual conditional suggests that a particular antecedent was necessary for a particular outcome. For example, "If I hadn't forgotten my umbrella, then I wouldn't have got wet" suggests that forgetting my umbrella was necessary for me to get wet. Likewise, "If I hadn't been walking in the rain, then I wouldn't have got wet" suggests that walking in the rain was necessary for me to get wet. However, walking in the rain might be viewed as the cause of getting wet, whereas forgetting my umbrella would tend to be considered a mere condition that allowed or enabled getting wet to occur, rather than the cause. People do not consider all conditions to be causes (e.g., Cheng & Novick, 1990; 1991; Einhorn & Hogarth, 1986; Goldvarg & Johnson-Laird, 2000; Hart & Honoré, 1959/1985; Hilton, 1990; Hilton & Slugoski, 1986; Mackie, 1974). There

are competing views about how people distinguish causes from conditions, but these accounts concur that it cannot be based on counterfactual reasoning alone (e.g., Einhorn & Hogarth, 1986; Hart & Honoré, 1959/1985; Hilton, 1990; Hilton & Slugoski, 1986).

This view is consistent with the lack of evidence that people *spontaneously* generate counterfactual thoughts when considering the causes of events. For example, the availability of counterfactual alternatives to an event has less effect on participants' causal judgements when they make the causal judgements before generating counterfactual thoughts than when they are directed to generate counterfactual thoughts first (Wells & Gavanski, 1989). Moreover, Roese and Olson (1997) found that directing participants to verify a causal assertion such as "Heavy drinking caused Bill's grade" had no effect on the time they subsequently took to verify a related counterfactual assertion such as "If not for heavy drinking, Bill's grade would have been better", whereas response times to the causal assertion were facilitated by verifying the counterfactual assertion. They suggest that causal thinking does not facilitate counterfactual thinking because people do not typically bring to mind counterfactual information when they think about the causes of events. Furthermore, although participants judge counterfactual information to be useful for assessing a causal hypothesis, only one percent of their spontaneous questions to assess a causal hypothesis are counterfactual in nature (Lipe, 1991).

Further evidence that causal and counterfactual thoughts are different comes from experiments which show that people tend to focus on different events when they are asked to mentally undo an outcome compared to when they are asked about the cause of the outcome (Mandel & Lehman, 1996; N'gbala & Branscombe, 1995). For example, given the taxi scenario described on page 8 of this chapter, participants tended to generate counterfactual thoughts that focused on the driver's decision (e.g., "If only he had given them a lift,..."), but they tended to

assign greater causality to the collapsed bridge (N'gbala & Branscombe, 1995). Likewise, given a scenario about a car crash between a victim and a reckless driver, participants tended to generate counterfactual thoughts that focused on the victim's choice of route, whereas they tended to generate causal thoughts that focused on the reckless driver's behaviour (Mandel & Lehman, 1996).

Two views of the relation between causal and counterfactual thinking

Clearly, causal and counterfactual thoughts seem to share some important similarities and some important differences. Although counterfactuals may not always warrant an inference of a causal relation between an antecedent and a consequent, nonetheless counterfactual thoughts can influence causal thoughts, at least in some circumstances. One possible explanation is that causal and counterfactual thoughts are concerned with different sorts of causal relations. Causality is not a primitive unanalysable notion (Miller & Johnson-Laird, 1976; Johnson-Laird & Byrne, 1991). For example, a distinction can be made between *necessary* and *sufficient* causes. A necessary cause must be present for a particular outcome to occur, but the outcome need not always occur in its presence (e.g., oxygen is necessary to start a fire, but fires do not always start when oxygen is present). In contrast, a sufficient cause always brings about a particular outcome, but the outcome can also occur as a result of other causes (e.g., water is sufficient to stop a fire, but a fire can also be stopped by other means). Furthermore, the major causal relations include not only relations between the occurrence of an antecedent and the occurrence of an outcome, but also relations between the occurrence of an antecedent and the non-occurrence of an outcome. The former case is what we normally refer to as a relation of *causation*; in the latter case, the relation is one of *prevention*, in which the antecedent prevents the outcome (e.g.,

Johnson-Laird & Byrne, 1991; Mandel & Lehman, 1996).

One view about the relation between causal and counterfactual thinking is based on the distinction between necessary and sufficient causes. On this account, counterfactual thinking is a subcomponent of causal thinking: causal thoughts focus on antecedents which are both necessary and sufficient to produce an outcome, and counterfactual thoughts are used to test the necessity of possible causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993). A counterfactual conditional which mutates a factual antecedent suggests that this antecedent was necessary to bring about the factual outcome, but it does not imply that it was sufficient. For example, "If I had brought my umbrella, then I wouldn't have got wet" suggests that not bringing my umbrella was necessary for me to get wet, but it does not imply that it was sufficient. The absence of the antecedent would have led to a different outcome, and so the antecedent was a necessary cause, but it may or may not have been sufficient. Counterfactual thoughts have implications only for the necessity of possible causes, but they may be used as a heuristic for assigning causality (e.g., Kahneman & Tversky, 1982; Wells & Gavanski, 1989). Alternatively, counterfactual thinking may be only one part of the causal judgement process: people may simulate counterfactual alternatives only in order to decide whether an antecedent was necessary for an outcome, and they may use other strategies to establish whether it was sufficient (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993).

A second view is that causal thoughts are concerned with relations of causation, whereas counterfactual thoughts are concerned with relations of prevention. Mandel and Lehman (1996) concur that a counterfactual conditional which mutates a factual antecedent does not imply that this antecedent was sufficient to bring about the factual outcome, but they suggest that it does imply that the *absence* of the antecedent would have been sufficient to *prevent* the factual

outcome. For example, "If I had brought my umbrella, then I wouldn't have got wet" suggests that not bringing my umbrella would have been sufficient for me to not get wet. Although people may think about causal relations in terms of both necessity and sufficiency (e.g., Cummins, 1995; Jaspers, Hewstone & Fincham, 1983; Mackie, 1974; Mill, 1843/1996), they tend to consider sufficiency to be a more important attribute (e.g., Mandel & Lehman, 1998; McGill & Klein, 1993). Given that counterfactual thoughts have implications for sufficient preventors, but only for necessary causes, they may be more useful in thinking about how an outcome could have been prevented than how it was caused. Mandel and Lehman (1996) propose that causal and counterfactual thinking may serve *different functions*: causal thoughts allow the *prediction* of future outcomes, whereas counterfactual thoughts allow the *prevention* of unwanted outcomes. For example, participants given the car crash scenario described on page 11 of this chapter may have generated causal thoughts focusing on the reckless driver because reckless driving predicts car crashes, whereas they may have generated counterfactual thoughts focusing on the victim's choice of route because a different choice would have prevented the accident.

Conclusions

Despite the longstanding and continuing interest in causal and counterfactual thinking among psychologists, the precise nature of the relation between causal and counterfactual thinking remains unclear. A possible solution is to compare the determinants of causal and counterfactual thoughts, but only *three* previous studies have attempted to compare causal and counterfactual thinking (Mandel & Lehman, 1996; N'gbala & Brabscombe, 1995; Roese & Maniar, 1997). The principle aim of this thesis is to provide the first systematic comparison of causal and counterfactual thinking. However, causal and counterfactual thinking have both been

studied extensively in isolation from each other. In the rest of this chapter, we will review previous research on the determinants of causal and counterfactual thinking, which has focused on four main issues: the activation of causal and counterfactual thoughts, the content of causal and counterfactual thoughts, the consequences of causal and counterfactual thoughts, and how people understand and reason from causal and counterfactual assertions. All but three of the studies we will review have examined the determinants of either causal or counterfactual thoughts alone, rather than comparing the two. We will conclude the chapter by outlining the aims of our research.

The activation of causal and counterfactual thoughts

People often spontaneously engage in causal and counterfactual thinking, but they seem to be more likely to do so in some sorts of situations than others. No previous studies have examined the activation of *both* causal and counterfactual thoughts. However, we are aware of twelve studies which have examined the activation of causal thoughts alone, and four studies which have examined the activation of counterfactual thoughts alone. The studies have examined the effects of four variables on the activation of causal and counterfactual thoughts: first, whether an outcome is expected or unexpected, second, whether the outcome is good or bad, third, whether the antecedents of the outcome are normal or unusual, and finally, whether the outcome appears close to or distant from an alternative. We will consider each of these variables in turn.

Outcome expectancy

Causal thoughts seem to be activated by unexpected rather than expected outcomes. Several philosophers have suggested that causal reasoning is elicited by changes or deviations from the normal course of events (e.g., Hart & Honoré, 1959/1985; Mackie, 1974). Eleven

studies have examined the effects of outcome expectancy on the activation of causal thinking. Two out of three studies based on examining existing written reports found that more causal explanations were recorded for unexpected than expected outcomes. Bettman and Weitz (1983, cited in Weiner, 1985b) found that corporate reports contained most causal explanations when companies performed better or worse than anticipated. Likewise, Lau and Russel (1980) found that newspaper accounts of sports events contained more causal explanations for unexpected than expected outcomes, although Lau (1984) failed to replicate this finding.

Four studies of participants' thoughts about imaginary events showed that participants produce more spontaneous causal thoughts about unexpected than expected outcomes. Wong and Weiner (1981) found that participants who imagined an academic success or failure generated more "why" questions for unexpected than expected outcomes. Likewise, Clary and Tesser (1983) and Hastie (1984) found that participants produced more causal explanations while paraphrasing or completing a story when the protagonist's actions were inconsistent with his personality. Similarly, participants produced more causal explanations while paraphrasing a story in which the outcome was unexpected rather than expected (Kanazawa, 1992).

However, only one out of four studies which examined participants' thoughts about actual experiences found that unexpected outcomes evoked causal thinking. Pyszynski and Greenberg (1981) found that participants were more likely to seek information relevant to explaining an individual's behaviour when it was unexpected than when it was expected. In contrast, Schoenman, van Uchelen, Stonebrink and Cheek (1986) found that participants who recalled a past success or failure produced no more "why" questions about unexpected than expected outcomes. Likewise, Gilovich (1983) found that gamblers produced as many causal thoughts about expected as unexpected wins and losses. Moreover, Bohner, Bless, Schwarz and

Strack (1988) found no effect of expectancy on the frequency of participants' causal thoughts about their success or failure at a laboratory task. However, these inconsistencies may result from the greater difficulty of manipulating participants' expectancies about actual personal successes and failures, because in general people tend to expect success (e.g., McGill, 1989). A further possibility is that the need to produce excuses for failures overwhelmed participants' natural tendencies to seek the causes of unexpected outcomes (Kanazawa, 1992).

Based in part on the observation that unexpected outcomes seem to activate causal thinking, Roese and Olson (1995) speculated that unexpected outcomes may also activate counterfactual thinking. Sanna and Turley (1996) carried out three experiments in which participants were asked to describe and elaborate on an examination scenario, a real-life examination and their performance at a laboratory task, respectively. In all three studies, participants produced more counterfactuals following unexpected than expected outcomes. However, Klauer and Migulla (1995) and Roese and Hur (1997) both found no effect of outcome expectancy on the frequency of counterfactuals in participants' listings of the thoughts that might occur to the protagonist of a scenario. We suggest that these conflicting findings may result from the methodological differences in these studies. Sanna and Turley (1996) asked participants to describe events so that another person could understand them, an instruction which may have encouraged attempts to explain an unexpected outcome, whereas Roese and Hur (1997) and Klauer and Migulla (1995) simply asked participants to list the thoughts that might occur to the character in their scenarios.

In sum, unexpected outcomes seem to activate causal thinking, but the effect of outcome expectancy on the activation of counterfactual thinking remains unclear. Causal thinking may be activated by unexpected outcomes because causal thinking is motivated by the need to

understand the environment in order to predict and control future outcomes, and unexpected outcomes indicate that past attempts at prediction have failed (e.g., Heider, 1958; Kanazawa, 1992; Weiner, 1985b). Roese and Olson (1995) suggest that counterfactual thoughts may be activated by unexpected outcomes for the same reason. However, further research is clearly needed in order to clarify whether unexpected outcomes activate counterfactual thinking. In Chapter 2, we will report an experiment which provides the first *comparison* of the effects of outcome expectancy on the activation of causal and counterfactual thoughts (Experiment 1).

Outcome valence

Causal thoughts may be activated by bad rather than good outcomes. Ten studies have examined the effects of outcome valence on the activation of causal thinking. Two out of four studies based on examining existing written reports suggest that bad outcomes evoke causal thinking more than good outcomes. Foersterling and Groenvald (1983, cited in Weiner, 1985b) found that newspaper articles contained more causal explanations for political losses than wins, and Lau (1984) found the same pattern for sports losses and wins. In contrast, Lau and Russel (1984) found no effect of outcome valence on the frequency of explanations in newspaper sports pages and Bettman and Weitz (1983, cited in Weiner, 1985b) found no effect on the frequency of explanations in corporate reports.

Two out of three studies based on thoughts about imaginary scenarios found that participants engaged in more causal thinking about bad than good outcomes. Wong and Weiner (1981) found that participants generated more “why” questions about an imagined academic failure than an imagined success. Likewise, Gioia and Sims (1986) found that among participants role-playing an employer and employee, employers asked more “why” questions and employees

produced more causal explanations when the employee's performance was poor. In contrast, Kanazawa (1992) found no effect of outcome valence on the number of causal explanations in participants' paraphrases of stories about academic successes and failures.

Finally, three studies showed that participants produce more causal thoughts after actually experiencing bad rather than good outcomes. Schoeneman et al. (1986) found that participants who recalled an actual academic failure produced more "why" questions than participants who recalled an academic success. Moreover, Gilovich (1983) found that gamblers produced more causal explanations for actual losses than wins. Likewise, Bohnet et al. (1988) found that participants produced more causal thoughts when they were given negative rather than positive feedback about their performance at a task.

Likewise, four out of five studies suggest that counterfactual thinking is activated more by bad than good outcomes. Klauer and Migulla (1995) found that participants' listings of the thoughts of a character in a scenario contained more counterfactuals when the outcome was bad than when it was good. Moreover, in experiments based on scenarios, laboratory tasks and self-reports of recent experiences, both Sanna and Turley (1996) and Roese and Olson (1997) found that participants produced more counterfactual thoughts about bad than good outcomes. Roese and Hur (1997) replicated these findings and also showed that response times to a counterfactual probe were significantly shorter following bad than good outcomes. In contrast, Markman, Gavanski, Sherman and McMullen (1993) found that the frequencies of counterfactual thoughts produced by participants who thought aloud as they played a card game were unaffected by whether the outcome was framed as good or bad.

In sum, experimental evidence suggests that bad outcomes evoke both causal and counterfactual thinking. People may engage in causal and counterfactual thinking about bad

outcomes in order to learn how to avoid them in the future (e.g., Markman et al., 1993; Roese, 1994; Weiner, 1985b). However, Kanazawa (1992) argues that people should be as motivated to bring about good outcomes as to avoid bad outcomes (e.g., Thorndike, 1905, p. 203), and so understanding the causes of success is at least as adaptive as understanding the causes of failure. Kanazawa suggests that bad outcomes evoke causal thoughts only when there is an external demand for justifications for a bad outcome. In Chapter 2, we will report an experiment which provides the first *comparison* of the effects of outcome valence on the activation of causal and counterfactual thoughts (Experiment 1).

Antecedent normality

We are unaware of any previous research on the effect of antecedent normality on the activation of causal thinking. However, much previous research on counterfactual thinking has been based on the assumption that it is activated by outcomes with unusual or exceptional antecedents (e.g., Miller & McFarland, 1986; Kahneman & Tversky, 1982). There is strong evidence that the *content* of counterfactual thoughts is affected by antecedent normality, in that people tend to undo outcomes by mentally altering unusual antecedents rather than normal ones, perhaps because unusual antecedents are more cognitively “mutable” (e.g., Kahneman & Tversky, 1982; Wells, Taylor & Turtle, 1987). However, it is not clear whether antecedent mutability determines the spontaneous *activation* of counterfactual thoughts as well as their content.

Only two studies have directly examined the effects of antecedent normality on the activation of counterfactual thinking. Roese and Olson (1996) gave participants scenarios in which a man gets delayed in a traffic jam and misses his flight after taking either his normal

route or an unusual route to the airport. Participants produced more spontaneous counterfactual thoughts about the unusual route version than the normal route version. However, we suggest that the normality of the route may have been confounded with the likelihood that taking an alternative route would have led to a different outcome, because the normal route was described as the character's "favoured" route. Consistent with this speculation, Roese & Hur (1997) found no effect of antecedent normality on the activation of counterfactual thinking in an experiment which did not have this possible confound. In their experiment, the response times to a counterfactual probe and frequencies of spontaneous counterfactual thoughts of participants who performed an anagram task were unaffected by whether the decisions they had made about the task were framed as normal or unusual. Moreover, the frequency of counterfactual thinking in victims of traumatic events appears to be unrelated to the normality of preceding events (e.g., Davis et al., 1995).

In sum, no previous studies have examined the effect of antecedent normality on the activation of causal thinking, and two studies of its effect on the activation of counterfactual thinking have produced conflicting findings. In Chapter 2, we will report an experiment which provides the first *comparison* of the effects of antecedent normality on the activation of causal and counterfactual thoughts (Experiment 2).

Outcome closeness

We are unaware of any previous research on the effect of the perceived closeness of an outcome to an alternative on the activation of causal thinking. In contrast, two studies have shown that participants produce more spontaneous counterfactual thoughts about "near miss" than "far miss" outcomes. Meyers-Levy and Maheswaran (1992) found that participants given

scenarios about a fire in an uninsured house produced more counterfactuals when the owner had forgotten to renew the insurance policy three days before the fire occurred than when he had forgotten six months beforehand. Roese and Olson (1996) found that participants given the missed flight scenario described on pages 19 to 20 of this chapter produced more counterfactuals when the flight was missed by five minutes than when it was missed by an hour. However, close outcomes evoked “close counterfactuals”, which undo an outcome without mutating any antecedent events (e.g., “He could easily have caught his flight”), rather than counterfactual conditionals (e.g., “If he had driven faster, he could easily have caught his flight”). Kahneman and Varey (1990, p. 1101) state that a close counterfactual “does not invoke an alternative possible world, but states a fact about the history of this world – namely, that things were close to turning out differently than they did”.

In sum, no previous studies have examined the effect of outcome closeness on the activation of causal thoughts. Two studies have shown that close outcomes evoke counterfactual thoughts, but not necessarily counterfactual conditionals.

Conclusions

Studies of the activation of causal and counterfactual thinking suggest that both processes occur spontaneously in a range of situations, but they are evoked more by some sorts of outcomes than others. Unexpected outcomes seem to evoke causal thoughts, but the effect of outcome expectancy on the activation of counterfactual thinking remains unclear. Bad outcomes seem to evoke both causal and counterfactual thinking, but it is not known whether bad outcomes evoke causal thoughts when there is no external demand to make excuses for failure. The effect of antecedent normality on the activation of causal thinking has not been examined, and its effect

on the activation of counterfactual thinking remains unclear. Finally, close outcomes evoke “close counterfactual” thoughts, but the effect of outcome closeness on the activation of causal thinking has not been tested.

No previous studies have examined the activation of *both* causal and counterfactual thoughts. In Chapter 2, we will report two experiments which provide the first *comparison* of the effects of outcome expectancy, outcome valence, antecedent normality and outcome controllability on the activation of causal and counterfactual thoughts (Experiments 1 and 2).

The content of causal and counterfactual thoughts

The previous section considered factors that influence the essential first step in the processes of causal and counterfactual thinking, their activation. The second step involves constructing causal explanations or counterfactual alternatives. Any outcome can be explained or mentally undone by a potentially infinite number of antecedent events, but the contents of causal and counterfactual thoughts seem to be constrained by a number of factors. Only two previous studies have analysed the contents of *both* causal and counterfactual thoughts (Mandel & Lehman, 1996; N’gbala & Branscombe, 1995), but many studies have examined the content of causal or counterfactual thoughts alone. This section reviews research on the effects of eight variables on the content of causal and counterfactual thoughts: first, the order and proximity of events in a sequence, second, the similarity of an antecedent to an outcome, third, whether an antecedent is necessary and/or sufficient for the outcome, fourth, whether an antecedent covaries with the outcome, fifth, whether an antecedent is normal or unusual, sixth, whether an antecedent is an action or an inaction, seventh, whether an antecedent is controllable or uncontrollable, and finally, the individual’s perspective on the outcome.

Order and proximity

One constraint on the content of causal thoughts is the order and proximity of events in a sequence. People expect causes to precede their effects in time (e.g., Kun, 1978; Mandel & Lehman, 1998; Tversky & Kahneman, 1980), and so their causal thoughts tend to focus on events which precede an outcome rather than subsequent events (e.g., Bullock, Gelman & Baillargeon, 1982; Einhorn and Hogarth, 1986). Moreover, causal thoughts tend to focus on antecedents which are temporally and spatially close to an outcome rather than more distant events (e.g., Michotte, 1946; Shanks, 1989). Closeness seems to be a particularly important determinant of causal judgements in young children (e.g., Siegler & Liebert, 1974), perhaps because their lesser knowledge makes it more difficult for them to make causal links between distant events (e.g., Einhorn & Hogarth).

Order and proximity may also be important determinants of the content of counterfactual thoughts. For example, counterfactual thoughts tend to focus on the last antecedent in a sequence, rather than earlier antecedents. Miller and Gunesagaram (1990) presented participants with a scenario in which two characters lost the chance of winning \$1,000 by failing to obtain the same outcomes from tossing a coin. Most participants reported that the counterfactual alternative of the second person obtaining the same outcome as the first came to mind more easily than the first person obtaining the same outcome as the second. Moreover, the second event is more mutable than the first even when the two events are presented simultaneously (Miller & Gunesagaram, 1990), and even when the second event is mentioned first (Byrne, Segura, Culhane, Tasso, & Berrocal, 2000). A possible explanation is that “early events in a sequence are more likely to be presupposed or taken for granted” (Miller & Gunesagaram, 1990, p. 1113). Consistent with this view, Byrne et al. (2000) showed that the second event is more mutable only

when the first event provides the context in which later events are interpreted. In contrast, when the two target events are preceded by a different contextualising event, the two events are equally mutable.

Although the last antecedent of an *independent* sequence is more mutable than earlier antecedents, the first antecedent in a *causal* sequence (in which each antecedent depends on the preceding one) is more mutable than later antecedents. Wells et al. (1987) gave participants a scenario in which the order of four antecedents in a causal chain was manipulated, and asked them to list ways that the story could be changed so that the outcome would be different. Mutating any of the four antecedents would have been sufficient to undo the outcome. Participants were more likely to mutate the first antecedent than later ones, but there were no differences in the mutability of the three later antecedents. Wells et al. propose that the first event is more mutable because it is difficult to imagine later events in a causal sequence not occurring, given that the first event did occur.

In sum, causal thoughts tend to focus on antecedents which are close to an outcome, rather than more distant events. Similarly, counterfactual thoughts tend to focus on the last antecedent in an independent sequence of events. However, counterfactual thoughts tend to focus on the first antecedent in a causal sequence.

Similarity

Causal thoughts tend to focus on antecedents which are similar to an outcome. Nisbett and Ross (1980, p. 116) suggested that people think about causes as if they expect that “great events ought to have great causes, complex events ought to have complex causes and emotionally relevant events ought to have emotionally relevant causes”. Experimental evidence

supports the idea that people expect causes to be similar to their effects (e.g., Regan, Strauss & Fazio, 1974). Likewise, the content of counterfactual thoughts may depend in part on the perceived correspondence between antecedents and outcomes in factors such as their normality, valence and magnitude (e.g, Bouts, Spears & van der Plight, 1992; Gavanski & Wells, 1989; Sim & Morris, 1998). Counterfactual thoughts may tend to focus on antecedents which correspond to the factual outcome (e.g, Gavanski & Wells, 1989), or on antecedents which do not correspond to the counterfactual outcome (Sim & Morris, 1998).

In sum, both causal and counterfactual thoughts seem to focus on antecedents which are similar to an outcome.

Antecedent necessity and sufficiency

Several theorists have suggested that the content of causal thoughts depends on the perceived necessity and sufficiency of antecedents for an outcome (see this chapter, pp. 11-12). For example, Mill's (1843/1996) prescriptive "joint method of agreement and difference" for the identification of causes selects factors which were both necessary and sufficient for an outcome (see this chapter, p. 6). Moreover, Kelly (1973, p. 194) proposed a descriptive theory of causal thinking in which necessary and sufficient antecedents are selected as causes: "The effect is attributed to that condition which is present when the effect is present and absent when the effect is absent". Likewise, Jaspers et al.'s (1983) "natural logic model" of causal attribution identifies a factor which is both necessary and sufficient for an outcome as the cause. Judgements about the necessity and sufficiency of possible causes may be based on information about past occasions when the cause and effect did and did not occur. Mill (1843/1996, p. 312) states that "If two or more instances in which the phenomenon occurs have only one circumstance in

common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance; the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause of the phenomenon”.

However, Mill (1843/1996) noted that outcomes often result from a jointly sufficient set of necessary conditions, none of which alone is sufficient to produce the outcome. Mackie (1974) suggests that the fact that people often refer to one of these conditions as “the cause” suggests that factors other than perceived necessity and sufficiency also play a role in causal judgement. One possibility is that causation is based on probabilistic notions rather than concepts of necessity and sufficiency (e.g., Suppes, 1970). For example, smoking may be considered a cause of cancer because smoking significantly increases the probability that a person will get cancer, although smoking is neither necessary nor sufficient for cancer (e.g., Cheng & Nisbett, 1993). An alternative explanation is that causes are interpreted in relation to their context, and so some unnecessary and/or insufficient antecedents may be perceived as necessary and/or sufficient in relation to a set of assumed circumstances (e.g., Einhorn & Hogarth, 1986; Mackie, 1974; McGill, 1989). For example, smoking may be considered sufficient for cancer, assuming that a person smokes a minimum number of cigarettes, does not have a special resistance to cancer, and so on. Mandel and Lehman (1998) found that participants tend to define the term “cause” in terms of sufficiency (e.g., “when the cause is present the effect will occur”) rather than necessity (e.g., “if the cause is not present the effect won’t occur”), which suggests that perceived sufficiency may be a more important determinant of the content of causal thoughts than perceived necessity.

N’gbala and Branscombe (1995) argue that causal and counterfactual thoughts differ in their tendencies to focus on necessary and sufficient antecedents. They suggest that causal

thoughts focus on sufficient causes, whereas counterfactual thoughts focus on necessary causes, in line with Copi's (1982, p. 410) claim that people consider sufficient causes when they are concerned with "the production of something desirable", but focus on necessary causes when they are concerned with "the elimination of something undesirable". They report two experiments in which participants read scenarios about a car accident preceded by two focal antecedents, which were intended to be perceived as sufficient and necessary, respectively. Participants assigned greater causality to the sufficient cause, but they tended to mutate the necessary cause. For example, participants given the taxi scenario described on page 8 of this chapter judged the collapsed bridge the most important cause of the accident, but they tended to mutate the driver's decision, rather than the collapsed bridge. However, there was no evidence that participants had actually perceived these two antecedents to differ in their sufficiency and necessity. Moreover, there are several possible alternative explanations of the observed differences in the content of participants' causal and counterfactual thoughts, which we will discuss in more detail in Chapter 2.

In sum, it has been argued that causal thoughts tend to focus on antecedents which were sufficient for an outcome, whereas counterfactual thoughts tend to focus on necessary antecedents, but there is no direct evidence for differences in the roles of necessity and sufficiency in causal and counterfactual thinking.

Antecedent covariation

An alternative to the above view is that the content of causal thoughts depends on judgements of the covariation between antecedents and an outcome, that is the change in the probability of the outcome in the presence versus absence of the antecedents. For example,

according to Cheng and Novick (1991, p. 95), a cause need not be either necessary or sufficient for an outcome, rather, “a cause is a factor the presence of which (relative to its absence) noticeably increases the likelihood of the effect” (p. 95). Kelly (1967) proposed one of several covariation-based theories of causal reasoning in the social domain (e.g., Hewstone & Jaspers, 1987; Jones & Davis, 1965; Kelly, 1967; 1972a; 1972b; 1973). He suggested that people attribute behaviours to the individual, the stimulus, or the situation, by computing the covariation of each of these three variables with the behaviour, as in a three-way analysis of variance. Experimental evidence suggests that people use covariation information in causal attribution as Kelly predicted, although they exhibit some systematic deviations from his model (e.g., Kelly & Michela, 1980; McArthur, 1972; Ross, 1977; for a review see Kassin, 1979). Many general theories of causal thinking also propose that covariation with an outcome is important and perhaps necessary for an event to be considered a cause (e.g., Cheng, 1997; Cheng & Novick, 1991; 1992; Einhorn & Hogarth, 1986; Morris & Larrick, 1995; Shustack & Sternberg, 1981).

Similar to judgements of necessity and sufficiency, judgements of covariation may be based on past experience of occurrences and non-occurrences of an antecedent and outcome. However, there are competing views of how people detect covariation. According to the associationist view, people learn the associative strengths between events in a similar way to how animals learn the association between conditioned and unconditioned stimuli (e.g., Baker, Murphy & Vallée-Tourangeau, 1996). In contrast, the rule-based view proposes that people extract covariation information by applying a rule to compute the probabilities of events (e.g., Cheng, 1997; Cheng & Novick, 1991; 1992). For example, Cheng et al. argue that the causal relation between an antecedent and outcome is computed by comparing the probability of the outcome in the presence and absence of the antecedent, and they provide experimental evidence

which corroborates their model. Cheng (1993; 1997) suggests that people acquire general beliefs about the “power” of particular events to produce particular outcomes by computing covariations in order to determine the causes of past outcomes.

However, other theories propose that covariation information is largely acquired through social knowledge rather than actual experience of events (e.g., Fiedler & Semin, 1988). Kelly (e.g., 1967; 1972a; 1972b; 1973) acknowledged that although people may be able to make causal judgements from observation alone, in the manner of “naïve scientists”, knowledge and prior experience also play a role in determining the content of causal thoughts. For example, he suggested that “the interpretation of a given covariation analysis often depends on the perceiver’s prior causal attributions” (1972a, p. 3). Nisbett and Ross (1980) argued that people’s preconceptions about causes often override observed patterns of covariation. However, Alloy and Tabachnik (1984) showed that judgements of covariation depend on both prior expectations and current data. Kelly (1972b, p. 152) also proposed that people may use “abstract ideas about the operation and interaction of causal factors” called “causal schemata” to make causal inferences when they only have information about a single occurrence and therefore cannot compute covariations.

Mandel and Lehman (1996) suggest that causal and counterfactual thoughts differ in their tendencies to focus on antecedents which covary with an outcome (see this chapter, pp. 11-13). They suggest that only the content of causal thoughts is determined by a covariational criterion, whereas the content of counterfactual thoughts is determined by a controllability criterion. They report three experiments in which participants read scenarios about a car accident or plane crash preceded by two focal antecedents, which were intended to be perceived as covarying and controllable, respectively. Participants’ causal thoughts tended to focus on the covarying

antecedent, whereas their counterfactual thoughts tended to focus on the controllable antecedent. For example, participants given a scenario about a car crash between a victim and a reckless driver generated causal thoughts which focused on the reckless driver, whereas they generated counterfactual thoughts which focused on the victim's choice of route. However, these experiments suffer from the same limitation as N'gbala and Branscombe's (1995) studies (see this chapter, pp. 26-27), in that there was no evidence that participants had actually perceived the two antecedents to differ in the proposed ways. Moreover, the antecedents differed in many other ways which suggest alternative interpretations of the findings. We will return to this issue in Chapter 2, in which we report an experiment that compares the effects of *explicit* covariation information on the contents of causal and counterfactual thoughts (Experiment 3).

In sum, it has been argued that causal thoughts tend to focus on antecedents which covary with an outcome, whereas counterfactual thoughts may focus on non-covarying antecedents, but there is no direct evidence for a difference in the role of covariation information in causal and counterfactual thinking.

Antecedent normality

Cheng and Novick (1991) propose that people distinguish the cause of an outcome from the conditions which merely enabled an outcome by computing the covariation between potential causes and the outcome over different sets of events. A factor is a cause if it covaries with the outcome in the focal set of events implied by the context (e.g., lightning covaries with fires in forests), whereas it is an enabling condition if it only covaries with the outcome in another focal set (e.g., oxygen does not covary with fires in forests, but it covaries with fires in other contexts). In contrast, several philosophers have suggested that the distinction between causes and

conditions depends on normality. The legal analysts Hart and Honoré (1959/1985) proposed that abnormal conditions are more likely to be selected as causes than normal conditions. Based on the observation that people usually seek the causes of abnormal rather than normal outcomes, they suggest that a cause is “a difference from the normal course which accounts for the difference in the outcome” (p. 29). They stress that normality is defined relative to a particular context. On their account, oxygen would not be considered the cause of a fire in a forest, because oxygen is normally present in a forest, but it might be considered the cause of a fire in a laboratory where oxygen is normally excluded (p. 35). Similarly, Mackie (1974, p. 34) states that “what is normal, right and proper is not so readily called a cause as something abnormal or wrong”. He argues that anything that is normal in the context is part of the “causal field” and therefore “automatically ruled out as a candidate for the role of cause” (p. 35). For example, although birth is both necessary and sufficient for death, we would not say that a person’s birth caused their death, because their birth is part of the causal field.

Several psychologists have also suggested that causal thoughts tend to focus on abnormal events rather than normal events. For example, in line with Mackie (1974), Einhorn and Hogarth (1986, p. 5) propose that the people are unlikely to ascribe causality to events that “generally occur” in a particular context. Likewise, Turnbull and Slugoski (1988, p. 69) suggest that a “factor that is not taken for granted, which is unusual or abnormal in some way, is typically identified as the cause”. According to Kahneman and Miller (1986, p. 149), a cause need not be statistically unusual, but it cannot be a “default value”. For example, a person’s way of driving may be explained by their being young, old or female, although these attributes are not unusual, because the default car driver is a middle aged male. However, Cheng and Novick (1991) suggest that people may only tend to focus on abnormal events to explain abnormal outcomes,

whereas they often focus on normal events to explain normal outcomes. For example, the statement “the earth’s gravity causes objects near its surface to fall” seems sensible even though gravity would be considered a default rather than an abnormal condition (p. 93).

Counterfactual thoughts also tend to focus on unusual antecedents, rather than normal antecedents. Kahneman and Tversky (1982) first demonstrated this *exceptionality effect* using two versions of a car accident scenario, one in which the victim left work at his normal time, but drove home by an unusual route, and one in which he drove home by his normal route, but left work unusually early. Participants tended to undo the outcome by mutating the unusual event rather than the normal event. Kahneman and Tversky (1982, p. 205) suggest that it is cognitively easier to mutate unusual than normal antecedents: “the psychological distance from an exception to the norm that it violates is smaller than the distance from the norm to the same exception”. Wells et al. (1987) showed that an event is more mutable when it is described as exceptional than when it is described as normal, regardless of whether the other events in the scenario are normal or exceptional. They propose that unusual events may be more mutable because they are more focal, more likely to be encoded as occurrences (e.g., Chapman & Chapman, 1967) and more likely to be perceived as causes (e.g., Kahneman & Miller, 1986). In contrast, normal events may be relatively immutable because they are perceived as constrained by factors such as laws, social rules and habit.

The studies by Kahneman and Tversky (1982) and Wells et al. (1987) used only scenarios which had exceptional outcomes, which leaves open the possibility that people mutate exceptional antecedents only when their outcome is also exceptional. Gavanski and Wells (1989) hypothesised that people may undo exceptional outcomes by mutating exceptional antecedents, but undo normal outcomes by mutating normal antecedents, because they expect exceptional

outcomes to have exceptional causes and normal outcomes to have normal causes (e.g., Einhorn & Hogarth, 1986). In fact, they found that people undo exceptional outcomes by mutating exceptional antecedents towards greater normality, but they undo normal outcomes by mutating both normal and exceptional antecedents towards greater exceptionality. For example, participants given a scenario about a student who passed an exam (a normal outcome) often mutated exceptional antecedents, such as the student's unusual anxiety, towards greater exceptionality (e.g., "If she had been more anxious..."). Klauer, Jacobson and Migulla (1995) replicated Gavanski and Wells (1987) findings and suggested that a distinction should be made between mutation focus (whether normal or exceptional events are mutated) and mutation direction (towards greater normality or greater exceptionality). Kahneman and Tversky (1982) may have been correct that people tend to focus on exceptions rather than norms, but the direction of mutation may be towards greater exceptionality as often as towards greater normality.

Roese and Hur (1997) showed that the exceptionality effect generalises to spontaneous counterfactual thoughts. Participants performed an anagram task and antecedent normality was manipulated by framing the decisions they had made prior to the task as typical or atypical. Participants' spontaneous counterfactual thoughts were more likely to mutate these decisions when they were framed as atypical than when they were framed as typical. In contrast, Davis et al. (1995) found that most of the spontaneous counterfactual thoughts of victims of real-life traumatic events tended to focus on normal antecedents. However, this may be because many of the events did not have any unusual antecedents. In support of this suggestion, accident victims are more likely to believe that their accident was avoidable when it occurred in unusual circumstances (Bulman & Wortman, 1977; Davis, Lehman, Silver, Wortman & Ellard, 1996),

which suggests that when real events have exceptional antecedents, they are more mutable than normal antecedents. Moreover, although Davis et al.'s (1995) participants tended to focus on normal antecedents, in most cases the direction of mutation was not towards greater exceptionality – most participants replaced normal antecedents with equally normal antecedents.

In sum, both causal and counterfactual thoughts seem to focus on unusual antecedents rather than normal antecedents, although this difference may occur only for unusual rather than normal outcomes.

Antecedent action or inaction

We are unaware of any previous studies which have directly examined whether causal thoughts tend to focus on actions or inactions. However, studies of the *omission bias* (e.g., Ritov & Baron, 1990; Spranca, Minsk & Baron, 1991) suggest that actions may be perceived as more causal than inactions. The omission bias refers to the tendency to prefer inactions (e.g., letting someone die) rather than actions (e.g., actively bringing about someone's death). For example, participants often choose not to vaccinate a child, even when they know that the risk of the child dying from the disease is greater than the risk of the child dying from the vaccination. Many participants justify their choice by arguing that they would have a stronger sense of having caused the child's death if it occurred as a result of their action of vaccinating the child, than if it occurred as a result of their inaction of failing to vaccinate the child (e.g., Ritov & Baron, 1990; Spranca et al., 1991).

The action-inaction distinction has been directly examined in relation to the content of counterfactual thoughts. Kahneman and Tversky (1982) gave participants a scenario in which one character decides to switch his investment from one company to another, and another

character decides to leave his investment as it is. Both characters subsequently find that they would have been better off by \$1,200 if they had made the opposite choice. Most participants judged that the character who acted would feel more regret than the character who did not act, even though the consequences of their decisions were identical. The finding that people regret actions more than inactions has been frequently replicated, and a similar phenomenon has been uncovered in relation to good outcomes - people expect good outcomes resulting from actions to evoke more positive emotions than identical outcomes resulting from inactions (e.g., Gleicher, Kost, Baker, Strathman, Richman & Sherman, 1990; Landman, 1987). This *action effect* in emotional experience may occur because actions are more mutable than failures to act: "it is usually easier to imagine oneself abstaining from actions that one has carried out than carrying out actions that were not in fact performed" (Kahneman & Miller, 1986, p. 145). Consistent with this view, participants judge that an actor will generate "if only" thoughts focusing on his action more than a non-actor will generate "if only" thoughts focusing on his inaction (Byrne & McEleney, 2000).

The relative mutability of actions and inactions may depend on their context (e.g., Gilovich & Medvec, 1994; N'gbala & Branscombe, 1997; Roese & Olson, 1993; Zeelenberg, 1996). For example, Gilovich and Medvec (1994) found that actions evoke more regret only in the short-term, whereas inactions evoke more regret in the long-term. However, this temporal pattern seems to be confined to situations in which there is a difference between the short term and the long term in what is known about the consequences of actions and inactions. For example, participants given scenarios in which the short term and long term consequences for the action and the inaction are clearly specified judge that the actor will feel more regret than the non-actor in both the short term and the long term (Byrne & McEleney, 2000).

The content of people's counterfactual thoughts may depend on what they have made explicit in their mental representations of the factual situation (Byrne, 1996; 1997). We have suggested (Byrne & McEleney, 1997; 2000) that people's mental representations of actions include two explicit possibilities: the past pre-action state and the present post-action state, whereas their mental models of inactions include only one explicit possibility, because the past and present states are the same. This difference may account for the greater mutability of actions than inactions: the explicit model of the pre-action state provides a ready-made counterfactual alternative to the action, whereas people have to first make their mental representations more explicit in order to generate a counterfactual alternative to the inaction. For this reason, it may be cognitively easier to mutate actions than inactions.

In sum, no previous studies have directly examined the effect of actions and inactions on the content of causal thoughts, but several studies show that counterfactual thoughts tend to focus on actions rather than inactions.

Antecedent controllability

The legal analysts Hart and Honoré (1959/1985, p. 33) suggest that causal judgements are based on the "contrast...between a free and deliberate human action and all other conditions". They argue that the notion of voluntary human intervention in the normal course of events is central to conceptions of causality, not only in the law, but also in everyday life (see also von Wright, 1973). Likewise, Kelly (1972a, p. 22-23) argues that "controllable factors will have high salience as candidates for causal explanation", because "the purpose of causal analysis – the function it serves for the species and the individual – is effective control". He suggests that control concerns may form the basis of the tendency to attribute other people's behaviour to

changeable internal factors such as their attitudes and intentions, rather than relatively fixed external factors such as situational constraints (e.g., Jones & Nisbett, 1971). However, on this account, people should also attribute their *own* behaviour to internal rather than external factors, as their own attitudes and intentions should seem even more controllable than other people's, yet people tend to attribute their own behaviour to external rather than internal factors (e.g., Jones & Nisbett, 1971). Moreover, in the studies by N'gbala and Branscombe (1995) and Mandel and Lehman (1996), participants' causal explanations for the negative outcomes of scenarios tended to focus on the antecedents which were uncontrollable from the central character's point of view (see this chapter, pp. 10-11; pp. 29-30). We are unaware of any previous research which has directly tested the effect of antecedent controllability on the content of causal thoughts.

In contrast, research on the content of counterfactual thoughts has shown that antecedents that are under an individual's control are more mutable than actions that are either uncontrollable or constrained by prior circumstances. In one experiment, Girotto, Legrenzi and Rizzo (1991) gave participants a scenario in which an individual is delayed by four events, all four of which involved his own actions, but only one of which was under his control. Participants tended to undo the outcome by mutating the controllable action rather than the uncontrollable actions. One limitation of this study is that it compared the mutability of one specific controllable action (stopping at a bar) and several specific uncontrollable actions (e.g., stopping the car because of an asthma attack). McCloy and Byrne (2000; McCloy, 2000) suggest that the relative controllability of these events may have been confounded with their social acceptability. Their findings show that the mutability of controllable antecedents is mediated by their social acceptability. Nonetheless, other studies which have systematically varied the controllability of the same events support the view that controllable events are generally more mutable. For

example, in a study by Markman et al. (1995), participants' counterfactual thoughts tended to mutate whichever aspect of a computer-simulated game they believed they controlled.

Research on victims of traumatic events also supports the hypothesis that people tend to undo outcomes by mutating events under their own control. Bulman and Wortman (1977) found that accident victims tended to undo their accidents by changing actions which they had undertaken voluntarily, rather than less controllable events. Likewise, Davis et al. (1995) found that among participants who had lost a relative in a car accident, all those who spontaneously mentioned counterfactual thinking reported mutating either their own behaviour or that of the deceased. None reported mutating the circumstances of the accident or the other driver's behaviour, even though most of them blamed the other driver. Parents who had lost a child to Sudden Infant Death Syndrome were also much more likely to mutate their own behaviours than any other event, even though only a very small minority blamed themselves for the child's death. Davis et al. (1996) propose that people mutate their own controllable actions because their purpose for engaging in counterfactual thinking is to understand how they could have prevented an outcome.

In sum, no previous studies have directly examined the effect of antecedent controllability on the content of causal thoughts, but several studies show that counterfactual thoughts tend to focus on controllable rather than uncontrollable antecedents. In Chapter 2, we will report an experiment which provides the first *comparison* of the effects of antecedent controllability on the content of causal and counterfactual thoughts (Experiment 3).

Perspective

There is evidence that the content of people's causal thoughts depends on their

perspective on a situation. For example, as we mentioned earlier (see this chapter, pp. 36-37), people tend to attribute their own actions to external factors, but they attribute other people's actions to internal factors (e.g., Jones & Nisbett, 1971). This difference seems to arise at least in part because the external situation is most salient to the actor, whereas the individual who is acting is most salient to the observer (e.g., Storms, 1973). Several other studies have demonstrated that the content of causal thoughts depends on the prominence of stimuli in the observer's field of view (e.g., McArthur & Post, 1977; Taylor & Fiske, 1975; 1978). An individual's perspective on a situation may also depend on their prior experience and knowledge. Causal thinking may often be based on the use of knowledge structures containing detailed information about social and physical events, which may allow people to make causal inferences automatically as part of the general process of understanding a situation (e.g., Abelson & Black, 1986; Leddo & Abelson, 1986), and which may play a role in constraining the generation of multiple causal hypotheses, as people tend to accept explanations which fit with their prior theories (for a review see Fiske & Taylor, 1991).

Perspective also influences the content of counterfactual thoughts. For example, Kahneman and Tversky (1982) found that participants who were asked to generate counterfactual thoughts about a scenario describing a car crash between a victim and a reckless driver tended to focus on the reckless driver when they were directed to take his relatives' perspective, but not when they were directed to take the victim's relatives' perspective. Girotto et al. (1991) showed that participants who are asked to generate counterfactual thoughts about a scenario, but not told whose perspective to take, tend to focus on the actions of the central character, rather than external events, and they suggest that this is because readers tend to interpret events in a narrative from the central character's perspective (e.g., Bower & Morrow,

1990; Legrenzi, Girotto & Johnson-Laird, 1993).

In sum, both causal and counterfactual thoughts seem to focus on antecedents which are salient due to the individual's perspective on a situation.

Conclusions

We have considered the effects of eight variables on the content of causal and counterfactual thoughts. Previous research suggests that both causal and counterfactual thoughts focus on antecedents which are close to an outcome, although counterfactual thoughts tend to focus on the first antecedent in a causal sequence. Both causal and counterfactual thoughts focus on antecedents which are similar to an outcome. Causal thoughts seem to focus on sufficient antecedents, whereas counterfactual thoughts seem to focus on necessary antecedents. Causal thoughts seem to focus on covarying antecedents, whereas counterfactual thoughts seem to focus on non-covarying antecedents. Both causal and counterfactual thoughts focus on unusual rather than normal antecedents, at least for unusual outcomes. Causal thoughts seem to focus on uncontrollable antecedents, whereas counterfactual thoughts focus on controllable antecedents. The content of both causal and counterfactual thoughts depends on an individual's perspective on a situation.

Only two of the studies we have reviewed examined the content of *both* causal and counterfactual thoughts (Mandel & Lehman, 1996; N'gbala & Branscombe, 1995), and neither of these studies *systematically* compared the effects of specific variables on causal and counterfactual thinking, as we will discuss further in Chapter 2. Moreover, there have been no previous studies of the content of *spontaneous* causal or counterfactual thoughts. In Chapter 2, we will report an experiment which provides the first *systematic* comparison of the effects of

covariation and controllability on the content of directed causal and counterfactual thoughts (Experiment 3), and an experiment which provides the first comparison of the content of *spontaneous* causal and counterfactual thoughts (Experiment 1).

The consequences of causal and counterfactual thoughts

Both causal and counterfactual thinking have a range of important emotional, cognitive and behavioural consequences. Only one previous study has examined the consequences of *both* causal and counterfactual thoughts (Roese & Maniar, 1997), but numerous studies have examined the consequences of causal or counterfactual thoughts alone. In this section, we review research on four kinds of consequences of causal and counterfactual thoughts: first, their effects on emotions, second, their influence on social judgements, third, their role in the hindsight bias, and finally, their consequences for expectancies, intentions and behaviour.

Emotions

People’s emotional reactions to an outcome depend not only on whether the outcome is a good or bad, but also on the individual’s causal explanations for the outcome (e.g., Averill, 1982; McFarland & Ross, 1982; Weiner, 1985a; 1986; Weiner, Russell & Lerman, 1978). For example, McFarland and Ross (1982) gave participants either success or failure feedback about their performance on a social accuracy task and induced them to attribute the outcome to either their ability or to task difficulty. Participants’ scores on self-report measures of positive and negative affect depended more on the type of causal attribution than the type of outcome. The results suggest that affective reactions to success and failure are determined primarily by causal thinking.

Weiner (e.g., 1985a; 1986; Weiner et al., 1978) proposed that causal thoughts determine

not only the degree of pleasure or displeasure people experience, but also the specific type of emotion, such as anger, happiness, pity or pride. Weiner identified three key dimensions of causes: internal-external, stable-unstable and controllable-uncontrollable. For example, personal effort is an internal, unstable, and controllable cause of success or failure. He suggested that the dimensions of the causes of a success or failure determine which specific emotion is experienced. For example, attributing a failure to an internal, unstable and controllable cause results in feelings of guilt, whereas a failure attributed to an external, stable and uncontrollable cause evokes self-pity. Experimental evidence supports the basic tenets of Weiner's theory, although it has been applied mainly to achievement-related contexts (e.g., Averill, 1982; Weiner, 1986; Weiner et al., 1978).

Further evidence for emotional consequences of causal thinking comes from studies of depression. Beck (1967) first proposed that depression is often brought about, exacerbated and/or maintained by an individual's tendency to make internal attributions for negative events, which result in decreased self-esteem. Abrahamson, Seligman and Teasdale (1978) suggested that causal thoughts about bad outcomes which focus on stable rather than unstable causes may lead to depression by creating a chronic sense of helplessness about future outcomes. Moreover, they proposed that a further dimension of causes, global-specific, may determine whether this sense of helplessness generalises to different sorts of contexts. Sweeney, Anderson and Bailey's (1986) meta-analysis of 104 studies of the role of "attributional style" in depression shows that the tendency to make more internal, stable and global attributions for negative outcomes is associated with more severe symptoms of depression, and this effect is greater for clinically depressed individuals than for non-depressed participants.

Counterfactual thoughts also have consequences for positive and negative affect, but the

consequences depend on their *direction* as well as their content: *upward* counterfactuals about how an outcome could have been better lead to negative affect, whereas *downward* counterfactuals about how an outcome could have been worse lead to positive affect. Kahneman and Miller's (1986) *emotional amplification* hypothesis proposes that counterfactuals amplify affective reactions to events because they are used as standards against which reality is compared. Comparing a bad outcome with a better counterfactual alternative makes the real outcome seem worse in contrast, thereby amplifying negative affect, whereas comparing a good outcome with a worse counterfactual alternative makes the real outcome seem better in contrast, thereby amplifying positive affect. Consistent with this view, people feel worse about bad outcomes of actions than inactions, whereas they feel better about good outcomes of actions than inactions, seemingly because actions are more mutable than inactions (e.g., Byrne & McEleney, 1997; 2000; Gleicher et al., 1990; Landman, 1987). Moreover, for both bad and good outcomes, participants who spontaneously generate upward rather than downward counterfactuals express greater dissatisfaction with the outcome (Markman et al., 1993), and participants who are directed to generate upward counterfactuals report more negative affect than those who are directed to generate downward counterfactuals (Roese, 1994; McMullen, Markman & Gavanski, 1995).

Counterfactual thoughts not only moderate the intensity of general positive or negative affective reactions to outcomes, they can also shape the particular emotions people experience. Kahneman and Miller (1986) suggested that *counterfactual emotions*, such as regret and disappointment, cannot be experienced without prior counterfactual thinking. The *particular* counterfactual emotion experienced may in turn depend on the content of the counterfactual thoughts. Counterfactual thoughts may increase the perceived causal importance of the

antecedents which they focus on (e.g., Wells & Gavanski, 1989), and these causal implications may define which counterfactual emotions are experienced. For example, the specific emotion of regret not only requires the comparison of an outcome with a counterfactual alternative, as suggested by Kahneman and Miller (1986), it also requires that this counterfactual is based on the mutation of a personal decision. Zeelenberg (1996) points out that it is this “choice context dependency” that distinguishes regret from other counterfactual emotions, such as disappointment, and accounts for the observation that “a sense of personal responsibility is central to the experience of regret” (Gilovich & Medvec, 1994, p. 359). Like regret, the emotions of guilt and shame may depend on the construction of self-focused counterfactuals, but they may be distinguished from each other by which particular aspects of the self are mutated: guilt is associated with mutating one’s behaviour, whereas shame is associated with mutating one’s traits (Niedenthal, Tangney & Gavanski, 1994).

In sum, both causal and counterfactual thinking have consequences for emotional experience. The emotional consequences of causal thoughts depend on their content and the type of outcome they explain, whereas the emotional consequences of counterfactual thoughts depend on their direction and their content. No previous studies have examined the emotional consequences of *both* causal and counterfactual thinking. In Chapter 3, we will report an experiment which provides the first *comparison* of the emotional consequences of causal and counterfactual thoughts (Experiment 6).

Social judgements

Causal thinking also has consequences for social judgements, as shown by studies of legal decision making. For example, Carroll and colleagues (Carroll & Payne, 1976; 1977;

Carroll, Perkowitz, Lurigio & Weaver, 1987) have examined how the judgements of legal professionals and students are influenced by the content of their causal explanations for crimes. They found that participants' judgements were influenced by the three causal dimensions proposed by Weiner (e.g., 1986): internal-external, stable-unstable and controllable-uncontrollable. For example, participants expressed more anger at the perpetrator of a crime when the crime was attributed to internal rather than external causes, and controllable rather than uncontrollable causes, and they considered the individual a greater risk to society when the crime was attributed to stable rather than unstable causes. Furthermore, participants' decisions as to whether individuals should be granted parole were determined by their levels of anger and perceived risk.

Counterfactual thinking also influences social judgements (e.g., Branscombe, Owen, Garstka & Coleman, 1996; Macrae & Milne, 1992; Miller & McFarland, 1986; Wells & Gavanski, 1989). For example, Wells and Gavanski (1989) demonstrated that ratings of a character's responsibility for an outcome are elevated when the outcome can be undone by mutating their actions. Similarly, Branscombe et al. (1996) showed that the generation of counterfactual thoughts which mutate the actions of an individual involved in a crime increase the amount of blame assigned to that individual. For example, in one study based on a rape scenario, participants blamed the victim more and the assailant less when they imagined changes in the victim's behaviour undoing the outcome. In contrast, they blamed the victim less and assailant more when they imagined changes in the assailant's actions undoing the outcome. Victim blame was lowest when people imagined changes in the victim's behaviour which did not undo the outcome, perhaps because these semifactuals (see this chapter, p. 2) would make it appear less likely that the victim could have prevented the outcome.

In sum, both causal and counterfactual thoughts focusing on an individual's actions may influence judgements of that individual's role in a negative outcome. Several theorists have suggested that the effects of counterfactual thoughts on social judgements depend on their causal implications (e.g., Roese, 1997; Wells & Gavanski, 1989; Williams, Lees-Haley & Price, 1994), yet no previous studies have examined the effects of *both* causal and counterfactual thinking on social judgements. In Chapter 3, we will report an experiment which provides the first *comparison* of the consequence of causal and counterfactual thoughts for judgements of responsibility and blame (Experiment 6).

Hindsight bias

Causal thinking may be responsible for the *hindsight bias*, that is the tendency to judge the prior predictability of an outcome to be higher when it is known that the outcome has actually occurred than when the outcome is unknown, and to deny that outcome knowledge has influenced this judgement (e.g., Fischhoff, 1975; for reviews see Christensen-Szalanski & Willham, 1991; Hawkins & Hastie, 1990). For example, a student who thinks retrospectively that her failure in an examination was highly likely may have judged failure much less probable before receiving her result, and yet she may claim in all good faith that she "knew all along" that she was going to fail. The hindsight bias may occur because knowledge of an outcome allows people to generate causal explanations for the outcome, which increase its perceived predictability. For example, on finding out that she has failed, the student may think "I failed because I didn't study" and she may judge her failure more predictable as a result.

Consistent with this view, the hindsight bias is strongest following negative and unexpected outcomes (e.g., Schkade & Kilbourne, 1991), and these types of outcomes also seem

to evoke spontaneous causal thinking (e.g., Weiner, 1985b; see this chapter, pp. 14-19).

Moreover, following positive outcomes, the hindsight bias is stronger in participants who are directed to generate causal thoughts than in controls (Roese & Maniar, 1997). However, causal thinking seems to increase the perceived predictability of an outcome only when the outcome is attributed to foreseeable causes. In contrast, when an outcome is attributed to unforeseeable causes, the hindsight bias is reduced or even eliminated (Wasserman, Lempert & Hastie, 1991). These findings suggest that the hindsight bias may be a side effect of the role of causal thinking in enabling the prediction of future outcomes. By attributing a past outcome to foreseeable causes, people may make future outcomes more predictable. Their subsequent overestimates of the predictability of the past outcome may result from their inability to recall their pre-outcome state of mind (e.g., Hawkins & Hastie, 1990; Wasserman et al., 1991).

Roese and Olson (1996) suggest that counterfactual thoughts also increase the hindsight bias, and they report two experiments to support this view. However, their manipulations of counterfactual thinking were indirect, and we suggest that they were confounded with the actual predictability of the outcomes of their scenarios. For example, in one experiment, participants read scenarios about a student who does badly in an exam after forgetting to take some medication. Those who read that the medication usually made her feel better rated the outcome as more predictable than those who read that the medication rarely made her feel better. Roese and Olson propose that the former scenario evokes more counterfactual thoughts, but they provide no evidence for this claim. A possible alternative explanation for their findings is that the former scenario suggests that the student needed the medication and therefore was more likely to do badly when she did not take it. The findings do not provide any direct evidence about the effect of counterfactual thinking on the hindsight bias.

In contrast, Roese and Maniar (1997) examined the effects of direct manipulations of both causal and counterfactual thinking on the hindsight bias. They interviewed spectators either before or after their attendance at football games. Those interviewed before the games were asked to judge the likelihoods of various possible outcomes, whereas those interviewed after the games were asked to recall their pre-game perceptions of the outcome likelihoods. As expected, the post-game participants showed a hindsight bias, in that they recalled higher pre-game expectancies for the actual outcome than the pre-game participants had initially recorded. Moreover, post-game participants who were directed to generate either causal or counterfactual thoughts about the outcome before recalling their likelihood judgements showed an even greater hindsight bias. However, this study only examined the effects of causal and counterfactual thoughts on the hindsight bias for good outcomes (participants were students attending football games which were won by their college's team). Counterfactual thoughts about good and bad outcomes can have different consequences for emotions (see this chapter pp. 42-43), and so counterfactual thoughts about good and bad outcomes may also have different consequences for other phenomena such as the hindsight bias. Moreover, it may be more important to determine the consequences of causal and counterfactual thoughts about bad outcomes than good outcomes, because people tend to *spontaneously* generate causal and counterfactual thoughts about bad rather than good outcomes (see this chapter, pp. 17-19).

In sum, one previous study suggests that both causal and counterfactual thinking increase the hindsight bias for good outcomes (Roese & Maniar, 1997), but no previous studies have examined their effects on the hindsight bias for bad outcomes. In Chapter 3, we report an experiment which provides the first comparison of the effects of causal and counterfactual thoughts on the hindsight bias for a bad outcome (Experiment 5).

Expectancies, intentions and behaviour

The content of causal thoughts about past outcomes may affect people's expectancies for future outcomes, their intentions for future action and their actual behaviour. For example, "attributional retraining" techniques in clinical psychology teach participants to change their causal explanations for personal successes and failures in ways which increase confidence, enhance motivation and improve performance (e.g., Foersterling, 1985). Weiner (1986) suggests that the stability of the perceived cause of a past outcome determines changes in expectancies for future outcomes. Stable causes lead to expectations that conditions will remain the same and so the outcome will recur, whereas unstable causes lead to the belief that conditions may change and so a different outcome may occur. He reviews evidence from twelve studies which show that focusing on unstable rather than stable causes leads to greater expectancies of success after a prior failure, and lower expectancies of success after a prior success. For example, Eiser, Van der Plight, Raw and Sutton (1985) found that smokers who attributed their failure to give up smoking to unstable rather than stable causes expressed greater confidence about their ability to give up smoking in the future.

The consequences of causal thoughts for intentions and behaviour also depend on whether they focus on stable or unstable causes. For example, Eiser et al. (1985) found that the greater confidence in their ability to give up smoking of participants who attributed past failures to unstable causes was associated with greater intentions to give up smoking in the future and more frequent attempts to give up. Moreover, participants show greater persistence at a task when they are directed to change their attributions for failure from stable to unstable causes (e.g., Anderson, 1983; Dweck, 1975). The consequences of causal thoughts for intentions and behaviour also depend on whether they focus on internal or external causes. For example,

Kernis, Zucherman and McVay (1988) found that participants who attributed their success on a two-person task to internal rather than external causes performed better when they were subsequently tested on a related task.

Counterfactual thoughts also influence expectancies, intentions and behaviour, but their consequences may depend on their direction as well as their content. McMullen et al. (1995) showed that upward counterfactuals focusing on internal factors increase perceived control over a past outcome, whereas downward counterfactuals or upward counterfactuals focusing on external factors have no effect on perceived control. Moreover, upward counterfactuals may indirectly increase perceived control over future outcomes. Nasco and Marsh (1999) found that participants who generated more upward counterfactual thoughts about a recent test performance were more likely to change the circumstances preceding a subsequent test. Changed circumstances were associated with greater perceived control over the upcoming test, which in turn was associated with better subsequent grades. Similarly, Roese (1994) found that participants who generated upward counterfactuals about a poor test performance expressed greater intentions to carry out success-facilitating behaviours for future tests. In another experiment, participants who generated upward counterfactuals about their performance on an anagram task showed improved subsequent performance relative to those who generated downward counterfactuals, and the improvements appeared to be linked to the strategies identified in their counterfactual thoughts. Roese (1994) suggests that upward counterfactuals may be used as scripts for future action.

In sum, both causal and counterfactual thinking seem to influence expectancies, intentions and behaviour. It has been argued that the effects of counterfactual thinking on expectancies, intentions and behaviour depend on their causal implications (e.g., Roese, 1997),

yet no previous studies have examined the consequences of *both* causal and counterfactual thinking for these phenomena. In Chapter 3, we report two experiments which provide the first *comparison* of the consequences of causal and counterfactual thoughts for perceived control over a past outcome and expectancies and intentions for the future (Experiments 4 and 6).

Conclusions

Causal and counterfactual thoughts both have effects on a range of emotional, cognitive and behavioural phenomena. Previous research suggests that both causal and counterfactual thoughts may determine people's levels of positive versus negative affect, and the specific types of emotions they experience. Both causal and counterfactual thoughts have consequences for a range of social judgements. Both causal and counterfactual thoughts may increase the hindsight bias for a good outcome. Both causal and counterfactual thoughts may affect people's expectancies for future outcomes, their intentions and their actual behaviour.

However, only one of these previous studies has examined the consequences of *both* causal and counterfactual thoughts (Roese & Maniar, 1997). In Chapter 3, we will report three experiments which provide the first systematic comparison of the consequences of causal and counterfactual thoughts for emotions and judgements (Experiments 4, 5 and 6).

Understanding and reasoning from causal and counterfactual assertions

We have now examined the factors that influence the activation, content and consequences of self-generated causal and counterfactual thoughts. In this section, we will discuss two remaining related questions about causal and counterfactual thinking: first, how people understand causal and counterfactual assertions, and second, how people make inferences from causal and counterfactual assertions. No previous studies have examined understanding or

reasoning from *both* causal and counterfactual assertions, but previous research has examined understanding or reasoning from causal or counterfactual assertions alone.

Understanding causal and counterfactual assertions

People may understand an assertion by keeping in mind the possibilities with which it is consistent (e.g., Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). Most research on how people understand assertions has focused on their understanding of conditionals in the indicative mood (for a review see Evans, Newstead & Byrne, 1993). An indicative conditional such as “If John studies hard, then he gets good marks” may be interpreted as consistent with the following four possibilities:

- John studies hard and gets good marks.
- John studies hard and does not get good marks.
- John does not study hard and gets good marks.
- John does not study hard and does not get good marks.

or with any subset of these four possibilities, resulting in sixteen possible interpretations. Ten of the sixteen possible interpretations appear to be consistent with the everyday usage of indicative conditionals (Johnson-Laird & Byrne, 2000). Two common interpretations are the *biconditional* interpretation, in which the assertion is consistent with only the first and last possibilities shown above, and the *conditional* interpretation, in which it is also consistent with the third possibility. On the biconditional interpretation, the antecedent, studying hard, is necessary and sufficient for the consequent, getting good marks, whereas on the conditional interpretation, the antecedent is sufficient but not necessary for the consequent. People may construct models in their mind to represent these possibilities (e.g., Johnson-Laird & Byrne, 1991; 2000). For example, on the biconditional interpretation, the assertion “If John studies hard, then he gets good marks” would

be represented by the following set of models:

study hard	good marks
¬study hard	¬good marks

where “study hard” represents “John studies hard”, “good marks” represents “John gets good marks” and “¬” represents negation (e.g., “¬study hard” represents “John does not study hard”).

Separate models are represented on separate lines in the diagram, and each model represents a possibility which is consistent with the assertion (Johnson-Laird & Byrne, 1991). On the conditional interpretation, the assertion “If John studies hard, then he gets good marks” would be represented by the following set of models:

study hard	good marks
¬study hard	good marks
¬study hard	¬good marks

The meaning of causal assertions has received much less attention from psychologists than the meaning of indicative conditionals, although philosophers have debated the meaning of causation for centuries (e.g., Hume, 1748/1999; Mackie, 1974; Mill, 1843/1996). Goldvarg and Johnson-Laird (2000) have recently proposed a psychological theory of the meaning of causal assertions which is based on two basic principles. First, the *modal* principle of causation states that the meaning of a causal relation between two states of affairs concerns what is possible and what is impossible in their co-occurrences. For example, “Studying hard causes John to get good marks” means that it is impossible for John to study hard and not get good marks. Second, the principle of *temporal constraint* states that if one state of affairs has a causal influence on another, then the latter does not precede the former in time.

One possible alternative view is that causal assertions are interpreted as probabilistic

statements. For example, “A causes B” may mean that A increases the probability of B (e.g., Cheng, 1997; Reichenbach, 1956; Suppes, 1970). However, Goldvarg and Johnson-Laird argue that if causation were a probabilistic notion, then there would be no difference in meaning between “A causes B” and “A often causes B”, yet it seems less accurate to say “smoking causes lung cancer” than “smoking often causes lung cancer”. Another possibility is that the meaning of causation includes understanding of the mechanisms by which the cause produces its effect (e.g., Ducasse, 1968; Harré & Madden, 1975; Kant, 1781/1929; Mackie, 1974; Schultz, 1982; White, 1989). However, Goldvarg and Johnson-Laird point out that the concept of causation is meaningful even when the underlying mechanisms are not understood. For example, the following assertion does not seem to be self-contradictory: “Your psychokinesis will cause a bias in this random-number generator, but no mechanism or power exists to produce the effect”.

Goldvarg and Johnson-Laird (2000) suggest that a general causal assertion “A causes B”, means that any occurrence of A is accompanied by B. On this account, a causal assertion such as “Studying hard causes John to get good marks” would have the following set of models:

study hard	good marks
¬study hard	good marks
¬study hard	¬good marks

which correspond to the conditional interpretation, in which the antecedent is sufficient but not necessary for the consequent. Similarly, a general prevention assertion “A prevents B” means that no occurrence of A is accompanied by B, and so “Studying hard prevents John from not getting good marks” has the same set of models as “Studying hard causes John to get good marks”.

The models of causal assertions which convey information about the factual situation

may be tagged to show their epistemic status. For example, the models of a singular causal assertion where the outcome is known, such as “Studying hard caused John to get good marks”, include a model of the factual situation, and alternative models of once-possible counterfactual situations:

factual:	study hard	good marks
counterfactual:	¬study hard	good marks
	¬study hard	¬good marks

The preceding examples are weak causal relations, in which the antecedent is sufficient but not necessary for the consequent. There are also strong causal relations such as “A and only A causes B”, in which the antecedent is necessary and sufficient for the consequent. For example, “Studying hard and only studying hard causes John to get good marks” has the following models:

study hard	good marks
¬study hard	¬good marks

which correspond to the biconditional interpretation. Goldvarg and Johnson-Laird (2000) report several experiments which corroborate their theory. For example, in one experiment, participants were given causal assertions and asked to list all the possibilities that would make an assertion true and all the possibilities that would make it false. Consistent with the modal principle of causation, most participants judged that some possibilities were ruled out by causal assertions such as “A will cause B”, but not by control tautologies such as “A or not A will cause B or not B”.

How do people understand counterfactual assertions? Counterfactuals are often expressed as conditionals in the subjunctive mood, such as “If John had studied hard, then he would have

got good marks”. The meaning of such assertions has been debated extensively by philosophers (e.g., Chisholm, 1946; Goodman, 1954/1983; Lewis, 1986; Stalnaker, 1968), but much less by psychologists. Johnson-Laird and Byrne (1991; 2000) propose that a counterfactual conditional is represented by the same set of models as its corresponding indicative conditional. However, unlike an indicative conditional, which conveys no information about the truth or falsity of its constituent propositions, a counterfactual conditional suggests that its antecedent and consequent are false (e.g., Carpenter, 1973; Fillenbaum, 1974), and so the models of a counterfactual conditional are tagged to show their epistemic status. For example, on the conditional interpretation, the assertion “If John had studied hard, then he would have got good marks” has the following set of models:

factual:	¬study hard	¬good marks
counterfactual:	study hard	good marks
	¬study hard	good marks

Byrne and Tasso (1999) showed that when participants were asked to generate two shapes which best fit the description “If there had been a circle on the blackboard, then there would have been a triangle”, they always generated either a circle and a triangle, or no circle and no triangle, whereas when they were asked to generate two shapes which definitely go against the description, they often generated a circle and no triangle. The results support the view that people consider some possibilities to be consistent with a counterfactual conditional and other possibilities to be inconsistent with it.

In sum, people may understand both causal and counterfactual assertions by keeping in mind the possibilities with which they are consistent. However, no previous studies have compared people’s interpretations of causal assertions such as “Studying hard caused John to get

good marks” and counterfactual assertions about the same factual situation such as “If John had not studied hard, then he would not have got good marks”. In Chapter 4, we will report an experiment which provides the first *comparison* of the possibilities people believe are consistent with causal and counterfactual assertions (Experiment 7).

Reasoning from causal and counterfactual assertions

There are three main competing theories of how people make inferences from assertions (for a review see Evans et al., 1993). One view is that reasoning depends on the application of abstract formal rules of inference to language-like representations of premises (e.g., Braine & O’Brien, 1998; Rips, 1994). An alternative view is that reasoning depends on the use of domain-specific rules or schemas which are evoked by particular contexts (e.g., Cheng & Holyoak, 1985). A further view is that people reason by constructing and manipulating mental models which represent situations in which the premises are true. The mental model theory has been more successful than either formal rule theories or schema theories in explaining both correct and incorrect reasoning performance, and in providing a unified account of different kinds of thinking and reasoning (e.g., Evans et al., 1993; Johnson-Laird & Byrne, 1991; Oakhill & Garnham, 1996).

According to the mental model theory of reasoning, people understand premises by constructing models of the states of affairs consistent with the premises, as outlined in the previous subsection. However, working memory limitations make it difficult for people to keep many models in mind, and so they initially construct a set of *mental models* which represent explicitly as little information as necessary to capture their understanding of a situation. For example, an indicative conditional such as “If John studies hard, then he gets good marks” may

be understood as consistent with two possibilities (on the biconditional interpretation) or three possibilities (on the conditional interpretation), but it may initially be represented by the following set of mental models, which represents explicitly only the possibility stated in the assertion:

study hard	good marks
...	

where the three dots represent an implicit model indicating that there may be additional possibilities which have not been explicitly represented (e.g., Johnson-Laird & Byrne, 1991). The model theory proposes that reasoning from two or more assertions requires three stages. First, people construct mental models of each of the assertions, second, they combine the two sets of models and produce a parsimonious description of the combined models, and finally, they evaluate this conclusion by searching for alternative models that might falsify it (e.g., Johnson-Laird & Byrne, 1991). They may “flesh out” their mental models to be more explicit if necessary, but this requires effort, and so everyday reasoning is often based only on the initial set of mental models, which sometimes results in errors (e.g., Johnson-Laird & Byrne, 1991; Johnson-Laird, Byrne & Schaeken, 1992).

The four main conditional inferences are shown in Table 1.1 below. The *affirmative* inferences, modus ponens and affirmation of the consequent, are based on the affirmation of information in the assertion (either the antecedent, e.g., “John studies hard”, or the consequent, e.g., “John gets good marks”), whereas the *negative* inferences, denial of the antecedent and modus tollens, are based on the negation of information in the assertion (either the antecedent, e.g., “John does not study hard”, or the consequent, e.g., “John does not get good marks”). The *forward* inferences, modus ponens and denial of the antecedent, proceed from antecedent to

consequent (e.g., John studies hard. Therefore, John gets good marks), whereas the *backward* inferences, affirmation of the consequent and modus tollens, proceed from consequent to antecedent (e.g., John gets good marks. Therefore, John studies hard).

Table 1.1: The four main conditional inferences, illustrated for the conditional, “If John studies hard, then he gets good marks”

	Affirmative inferences:	Negative inferences:
	<i>Modus ponens</i>	<i>Denial of the antecedent</i>
Forward inferences:	If John studies hard, then he gets good marks. John studies hard. Therefore, John gets good marks.	If John studies hard, then he gets good marks. John does not study hard. Therefore, John does not get good marks.
	<i>Affirmation of the consequent</i>	<i>Modus tollens</i>
Backward inferences:	If John studies hard, then he gets good marks. John gets good marks. Therefore, John studies hard.	If John studies hard, then he gets good marks. John does not get good marks. Therefore, John does not study hard.

The denial of the antecedent and affirmation of the consequent inferences are valid on a biconditional interpretation, but they are invalid on a conditional interpretation, whereas the modus ponens and modus tollens inferences are valid regardless of whether a biconditional or conditional interpretation is taken. However, whereas people reliably make the modus ponens inference from indicative conditionals, they often fail to make the equally valid modus tollens inference. According to the mental model theory, in order to make the modus ponens inference, people construct a mental model of the first premise:

study hard

good marks

...

and a mental model of the second premise:

study hard

They then add the information from the second model to the first model and eliminate the implicit model, as follows:

study hard good marks

They describe the combined model as “John studies hard and gets good marks”, and since it is already stated in the second premise that John studies hard, they make the parsimonious conclusion, “John gets good marks”. The modus tollens inference is more difficult because it requires people to flesh out their initial models and keep several alternatives in mind. They again construct a mental model of the first premise:

study hard good marks
...
...
...

and a mental model of the second premise:

¬good marks

The two models cannot be combined because they have nothing in common, and so people often erroneously conclude that nothing follows. In order to make the correct conclusion, the initial set of mental models of the first premise must be fleshed out to be more explicit. It may be fleshed out to the biconditional set of possibilities:

study hard good marks
¬study hard ¬good marks

or the conditional set of possibilities:

study hard	good marks
¬study hard	good marks
¬study hard	¬good marks

In either case, the combination of the fleshed out set of models with the model of the second premise rules out all but one possibility:

¬study hard	¬good marks
-------------	-------------

leading to the correct conclusion that John does not study hard (Johnson-Laird & Byrne, 1991). People readily make inferences such as modus ponens which are supported by their initial mental models of indicative conditionals, but they make more errors and take longer to make inferences such as modus tollens which require them to flesh out their initial mental models to be more explicit (e.g., Johnson-Laird et al., 1992).

Goldvarg and Johnson-Laird (2000) provide experimental evidence that people rely on mental models to make deductive inferences from pairs of causal assertions. For example, most participants given the premises “A prevents B” and “B causes C” inferred that “A prevents C”. This conclusion is invalid according to the fully explicit models of the premises, but it follows from mental models of the premises which represent only the stated situation. People may construct the following mental model of the first premise:

A	¬B
...	

and the following mental model of the second premise:

B	C
...	

They then combine the first and second models, as follows:

A	$\neg B$	
	B	C
	...	

The only possibility containing A does not contain C, and so they conclude that A prevents C. In contrast, the combination of the fully explicit models of the two premises results in the following possibilities:

A	$\neg B$	C
A	$\neg B$	$\neg C$
$\neg A$	B	C
$\neg A$	$\neg B$	C
$\neg A$	$\neg B$	$\neg C$

which show that there is no causal relation between A and C. The specific patterns of correct and incorrect causal inferences demonstrated by Goldvarg and Johnson-Laird (2000) cannot easily be explained by the idea that people reason from causal assertions by using abstract formal rules of inference or specialised rules of inference which are attached to causal schemas (see this chapter, p. 57).

There is also evidence that people make inferences from counterfactual assertions using mental models. Unlike an indicative conditional, which conveys no information about the truth or falsity of its constituent propositions, a counterfactual conditional suggest that its antecedent and consequent are false (e.g., Carpenter, 1973; Fillenbaum, 1974). Consequently, people may construct a more explicit initial set of models of counterfactual conditionals than indicative conditionals. An indicative conditional such as “If John studied hard, then he got good marks” may initially be represented as only the hypothesised situation:

study hard good marks

...

In contrast, a counterfactual conditional such as “If John had studied hard, then he would have got good marks” may bring to mind both the hypothesised counterfactual situation and the presupposed factual situation:

factual:	–study hard	–good marks
counterfactual:	study hard	good marks

...

(Byrne & Tasso, 1994; 1999; Thompson & Byrne, 2000). If people construct different mental models of indicative and counterfactual conditionals, and reasoning is based on mental models, then people may make different sorts of inferences from indicative and counterfactual conditionals, even if they are interpreted as consistent with the same fully explicit sets of possibilities. In fact, people often make more negative (denial of the antecedent and modus tollens) inferences from counterfactual conditionals than from indicative conditionals, while they make about the same amount of affirmative (modus ponens and affirmation of the consequent) inferences (Byrne & Tasso, 1994; 1999; Thompson & Byrne, 2000). The negative inferences are supported by the mental models of counterfactual conditionals, whereas they require the mental models of indicative conditionals to be fleshed out to be more explicit. In contrast, the affirmative inferences can be made directly from the mental models of both assertions. Moreover, people are more likely to judge that the negated antecedent and consequent are implied by and consistent with a counterfactual conditional than an indicative conditional, and the differences in inferences from counterfactual and indicative conditionals depend on these different judgements (Thompson & Byrne, 2000). Differences in people’s inferences from

counterfactual and indicative conditionals would not be predicted by theories based on formal rules of inference or reasoning schemas (see this chapter, p. 57).

In sum, people may make inferences from both causal and counterfactual assertions by combining their mental models of the premises and generating a parsimonious description of the combined models. However, no previous studies have examined the sorts of inferences people make from *both* causal and counterfactual assertions. In Chapter 4, we will report two experiments which provide the first *comparison* of reasoning from causal and counterfactual assertions: one experiment on the sorts of inferences participants endorse from the assertions (Experiment 8), and a further experiment on their latencies to endorse the inferences (Experiment 9).

Conclusions

Previous research suggests that people understand both causal and counterfactual assertions by constructing models which represent possibilities consistent with the assertions. Because of working memory limitations, they may initially construct mental models which represent explicitly only some of the models from their fully explicit set, and they may make inferences by combining their mental models of two or more assertions.

However, no previous studies have compared people's understanding of causal and counterfactual assertions, or the sorts of inferences they make from them. In Chapter 4, we will report three experiments which provide the first comparison of how people understand and reason from causal and counterfactual assertions (Experiments 7, 8 and 9).

Summary and aims of the thesis

Our review of past research on causal and counterfactual thinking shows that both

activities are pervasive and universal aspects of human cognition, which can have important consequences for thoughts, feelings and actions. Causal and counterfactual thinking seem to share similarities (e.g., Wells & Gavanski, 1989) and differences (e.g., N'gbala & Branscombe, 1995), but the precise nature of the relation between causal and counterfactual thinking remains unclear (e.g., Roese, 1997; Spellman & Mandel, 1999). One possibility is that the primary purpose of counterfactual thinking is to determine the causes of outcomes: causal thoughts may identify antecedents which were necessary and sufficient for an outcome, and counterfactual thoughts may be used to test the necessity of possible causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see this chapter, pp. 11-12). A second possibility is that causal and counterfactual thinking serve different purposes: causal thoughts are concerned with how outcomes can be predicted, whereas counterfactual thoughts are concerned with how outcomes can be prevented (e.g., Mandel & Lehman, 1996; see this chapter, pp. 11-13).

The two views of the relation between causal and counterfactual thinking which we have outlined each imply specific predictions about the similarities and differences in the activation, content and consequences of causal and counterfactual thoughts, and in how people understand and reason from causal and counterfactual assertions. Previous research has investigated all of these aspects of causal and counterfactual thinking (see this chapter, pp. 14-64), but only three studies have attempted to *compare* causal and counterfactual thinking: two studies compared the content of causal and counterfactual thoughts (Mandel & Lehman, 1996; N'gbala & Branscombe, 1995), and one study compared the consequences of causal and counterfactual thinking (Roese & Maniar, 1997). There have been no previous comparisons of the activation of causal and counterfactual thoughts, or of how people understand and reason from causal and counterfactual assertions.

The primary aim of this thesis is to contribute to the ongoing debate on the relation between causal and counterfactual thinking by providing the first systematic comparison of causal and counterfactual thoughts. More specifically, we aim to directly test the predictions of the hypothesis that causal and counterfactual thinking serve different functions: causal thinking is concerned with the prediction of outcomes, whereas counterfactual thinking is concerned with their prevention. We report nine experiments which examine this view by comparing the determinants of causal and counterfactual thinking: in Chapter 2, we report three experiments which compare the activation and content of causal and counterfactual thoughts; in Chapter 3, we report three experiments which compare the consequences of causal and counterfactual thoughts; and in Chapter 4, we report three experiments which compare how people understand and reason from causal and counterfactual assertions. The results shed light on the cognitive and motivational determinants of causal and counterfactual thinking in addition to clarifying the relation between them, and we will discuss these implications in Chapter 5.

Chapter 2 The Activation and Content of Causal and Counterfactual Thoughts

In this chapter, we will examine the hypothesis that causal thinking is concerned with prediction and counterfactual thinking is concerned with prevention, by systematically comparing the activation and content of causal and counterfactual thoughts.

The view that causal thinking is concerned with prediction and counterfactual thinking is concerned with prevention suggests there should be specific similarities and differences in the activation of causal and counterfactual thinking. First, both causal and counterfactual thoughts should be activated by bad rather than good outcomes, because the aim of both prediction and prevention may be to avoid bad outcomes (Mandel & Lehman, 1996). Second, causal but not counterfactual thoughts should be activated by unexpected rather than expected outcomes, because unexpected outcomes indicate that prediction has failed. Third, counterfactual but not causal thoughts should be activated by controllable rather than uncontrollable outcomes, because controllable outcomes indicate that prevention was possible. Fourth, causal but not counterfactual thoughts should be activated by outcomes with unusual rather than normal antecedents, because unusual antecedents may be better predictors than normal antecedents (at least when the outcome is unusual). Past research shows that both causal and counterfactual thoughts are activated more by some sorts of outcomes than others (see Chapter 1, pp. 14-22), but no previous studies have *compared* the activation of causal and counterfactual thoughts.

The view that causal thinking is concerned with prediction and counterfactual thinking is concerned with prevention also suggests that there should be specific similarities and differences in the contents of causal and counterfactual thoughts. First, both causal and counterfactual thoughts should focus on antecedents which covary with an outcome, because the presence of a

covarying antecedent helps predict an outcome, whereas the absence of a covarying antecedent helps prevent an outcome. Second, causal but not counterfactual thoughts should focus on general factors which predict an outcome in a broad range of situations, rather than antecedents which are specific to a particular outcome. Third, counterfactual but not causal thoughts should focus on controllable internal factors which are potentially preventable, rather than external factors or uncontrollable internal factors which are not preventable. Finally, counterfactual but not causal thoughts should focus on inhibitory antecedents whose absence would prevent a bad outcome, rather than facilitative antecedents whose absence would not prevent a bad outcome. Past research shows that the contents of both causal and counterfactual thoughts depend on a number of factors (see Chapter 1, pp. 22-41), but only two previous studies have compared the contents of causal and counterfactual thoughts (Mandel & Lehman, 1996; N'gbala & Branscombe, 1995).

In the two previous studies which compared the content of causal and counterfactual thoughts, participants read scenarios which included two focal antecedents, which were intended to differ in either their necessity and sufficiency for the outcome (N'gbala & Branscombe, 1995) or their controllability and covariation with the outcome (Mandel & Lehman, 1996). After reading the scenarios, participants were directed to generate either causal or counterfactual thoughts. The contents of their thoughts were categorised according to whether they focused on the necessary or sufficient antecedent, or the controllable or covarying antecedent, respectively. Unfortunately, in neither case was there any evidence that participants had perceived the events to differ in these distinct ways, and in fact, the events differed in many ways, and so differences in participants' tendencies to focus on them in causal and counterfactual thinking could have arisen for many different reasons (see also Chapter 1, pp. 26-27; pp. 29-30).

Moreover, both sets of previous studies examined the content of *directed* rather than *spontaneous* causal and counterfactual thoughts, but differences in the content of *directed* causal and counterfactual thoughts may result merely from task demands. As Roese and Olson (1997, p. 37) suggest: "Asking someone how something could have been different might cue implicit expectations for personal, controllable actions, whereas asking someone about the cause of an event might implicitly suggest a wider causal search, one that embraces a variety of causal candidates". In contrast, differences in the content of *spontaneous* causal and counterfactual thoughts cannot be attributed to differences in the assumptions cued by different questions, as participants are not asked any direct questions.

The previous studies have also been united by the very specific nature of the scenarios that were used. First, all of the scenarios were about fictional characters, and so participants might have thought about the events differently from how they would think about events which they were personally involved in. Second, the scenarios were all about highly negative outcomes, car accidents and plane crashes, and so the findings may not generalise to how people think about neutral or positive outcomes. Third, accidents are a very specific type of negative outcome, because they reflect the failure of prevention goals, which are concerned with avoiding negative outcomes, rather than promotion goals, which are concerned with bringing about positive outcomes (Higgins, 1998). Roese, Hur and Pennington (1999) have shown that the content of counterfactual thoughts about negative outcomes depends on whether the outcome results from a prevention failure, such as losing money on an investment, or a promotion failure, such as failing to get a job. Finally, a specific feature of accidents is that they are dramatic and rare, and so the findings may not generalise to how people think about everyday situations.

In this chapter, we report three experiments on the activation and content of causal and counterfactual thinking which overcome all of the above limitations. The first two experiments examined the activation and content of *spontaneous* causal and counterfactual thoughts. In the first experiment, participants imagined themselves involved in a scenario about moving to a new town and attempting to make new friends. We compared the effects of bad and good outcomes, unexpected and expected outcomes, and controllable and uncontrollable outcomes on the frequencies of their spontaneous causal and counterfactual thoughts. We also compared the general (internal or external) versus specific (inhibitory or facilitative) content of their spontaneous causal and counterfactual thoughts. In the second experiment, we compared the effects of unusual and normal antecedents on the frequencies of participants' spontaneous causal and counterfactual thoughts about a similar scenario. The third experiment examined the content of directed causal and counterfactual thoughts, using a scenario about a car accident. We directly compared the effects of covariation and controllability information on the content of participants' causal and counterfactual thoughts.

Experiment 1

This experiment tested the idea that causal thinking is concerned with prediction and counterfactual thinking is concerned with prevention, by examining the activation and content of *spontaneous* causal and counterfactual thoughts.

Activation

The experiment examined three key variables concerning the activation of spontaneous causal and counterfactual thoughts: outcome valence, outcome expectancy, and outcome controllability. First, both causal and counterfactual thoughts should be evoked more by *bad*

outcomes than *good* outcomes, because the ultimate aim of both prediction and prevention may be to avoid future failures (Mandel & Lehman, 1996). In fact, studies of spontaneous causal thoughts (e.g., Bohnet et al., 1988; Weiner, 1985) and studies of spontaneous counterfactual thoughts (e.g., Klauer & Migulla, 1995; Roese & Olson, 1997; Sanna & Turley, 1996) suggest that they are both evoked by bad outcomes (see also Chapter 1, pp. 17-19). However outcome valence should affect counterfactual thinking more than causal thinking, because it is usually unnecessary to prevent good outcomes, whereas it may often be useful to predict good outcomes. No previous studies have compared the strength of the effect of outcome valence on spontaneous causal and counterfactual thoughts.

Second, causal thoughts should be evoked more by *unexpected* outcomes than *expected* outcomes, because unexpected outcomes indicate that prediction has failed. In contrast, counterfactual thoughts may be evoked more by expected outcomes, because expected outcomes may be easier to prevent than unexpected outcomes. In fact, studies of spontaneous causal thoughts show that they tend to be evoked by unexpected outcomes (e.g., Kanazawa, 1992; Weiner, 1985; see also Chapter 1, pp. 14-16). However, the effect of outcome expectancy on spontaneous counterfactual thoughts remains unclear (e.g., Klauer & Migulla, 1995; Roese & Olson, 1997; Sanna & Turley, 1996; see also Chapter 1, pp. 16-17), perhaps because unexpected outcomes may bring to mind their expected counterparts, so that counterfactual alternatives are more readily available cognitively for unexpected outcomes (e.g., Kahneman & Miller, 1986; Kahneman & Tversky, 1982). This cognitive availability may oppose the tendency for counterfactual thoughts to be evoked more by expected outcomes, so that some studies show no difference in the production of counterfactual thoughts between expected and unexpected outcomes.

Third, counterfactual thoughts should be evoked more by *controllable* outcomes than *uncontrollable* outcomes, because only controllable outcomes are potentially preventable. In contrast, outcome controllability should not affect the frequency of causal thoughts, because controllable and uncontrollable outcomes should be equally predictable. No previous studies have examined the effects of outcome controllability on the spontaneous generation of either causal or counterfactual thoughts.

Content

A second way in which the experiment tested the idea that causal thinking enables prediction and counterfactual thinking enables prevention is by examining a key variable in the *content* of spontaneous causal and counterfactual thoughts, that is their generality. We expected that causal thoughts should be more likely than counterfactual thoughts to focus on *general* factors that could explain similar outcomes in different situations (e.g., laziness, or bad luck, to explain an examination failure). Causal thoughts focusing on general factors would enable future prediction in a broader range of situations more than causal thoughts focusing on antecedents which were *specific* to that particular outcome (e.g., lack of study for that particular examination). In contrast, we expect that counterfactual thoughts should be more likely to focus on specific antecedents rather than general factors, as thoughts concerning specific concrete events would be more easily converted into plans for future preventative action (e.g., Taylor & Pham, 1996).

The distinction between general and specific factors can be broken down further. General factors can be either *internal* factors such as an individual's own actions or characteristics, or *external* factors such as other people's actions or situational factors (e.g., Heider, 1958; Kelly,

1971). Counterfactual thoughts should be more likely than causal thoughts to focus on internal rather than external factors. Internal factors are potentially controllable and therefore can potentially be changed to prevent an outcome, whereas external factors may be uncontrollable and therefore ruled out as possible preventors. However, internal and external factors would be equally useful predictors. Specific factors can be either *inhibitory* antecedents that increase the likelihood of a bad outcome or *facilitative* antecedents that increase the likelihood of a good outcome. Counterfactual thoughts should be more likely than causal thoughts to focus on *inhibitory* antecedents rather than *facilitative* antecedents. The absence of an inhibitory antecedent would be likely to prevent a bad outcome, whereas the absence of a facilitative antecedent would not. However, inhibitory and facilitative antecedents would be equally relevant to prediction.

To test these predictions, we compared the effects of the valence (bad or good), the expectancy (unexpected or expected) and the controllability (controllable or uncontrollable) of an outcome on the frequencies of spontaneous causal and counterfactual thoughts. We also compared the general (internal or external) versus specific (inhibitory or facilitative) content of spontaneous causal and counterfactual thoughts.

Method

Participants and Design

The 248 undergraduates from a variety of departments at the University of Dublin, who took part in the experiment voluntarily, included 190 women, 57 men, and one person who did not record gender. Their average age was 21 years, with a range from 17 to 55 years. They were randomly assigned to eight groups, in line with a two (bad or good outcome) by two (unexpected

or expected outcome) by two (controllable or uncontrollable outcome) between-participants factorial design.

Materials

We constructed eight versions of a scenario about moving to a new town and attempting to make new friends. The versions varied in whether the outcome was bad or good, unexpected or expected, and controllable or uncontrollable. All of the versions described four specific antecedent events that occurred after the move, two of which would inhibit a good outcome (not going to a party, and not joining a club), and two of which would facilitate a good outcome (attending a staff dinner, and meeting an old friend). We included these four events in order to provide participants in all conditions with a range of antecedents that they could use either to explain the outcome, or to construct counterfactual alternatives. The four antecedents were presented to participants in one of two possible orders (the order in the scenario below and the reverse order), to control for possible order effects.

The outcome was the individual’s social situation six weeks after the move. Outcome valence was manipulated by describing it as either bad (not having made friends and feeling lonely and isolated), or good (having made friends and feeling happy and at home). Outcome expectancy was manipulated by describing the individual’s expectations about settling in to the new town at the start of the scenario as either expecting it to be easy or hard, and by describing the person's reaction to the outcome at the end of the scenario as either very surprised or not surprised. Outcome controllability was manipulated by describing each of the four antecedent events as the result of either a controllable decision or some uncontrollable factor.

An example of the scenario - the bad, unexpected, controllable outcome version - is as

follows (see also Appendix 1a):

You're moving house to start a new job in a different city. The night before you leave, you write down your thoughts about the move in your diary: ...I've got mixed feelings about moving to a place where I know hardly anyone - my friends and social life are so important to me. But I'm sure it will be easy to settle in to the new town - I've never had any trouble making new friends...

A lot happens in your first two weeks in the new town. During your first week at work, a staff dinner is held. You decide to go because you want to get to know your colleagues. You enjoy the evening and meet a lot of people. That weekend, your next-door neighbours invite you to a party. Most of the people who live on your road will be there. However, that evening you decide to go to the cinema instead. The next week, you decide to ring an old friend who lives in the town and ask him to show you around. You arrange to go out with him the following evening and he introduces you to a lot of his friends. A few days later, a colleague tells you there's a membership vacancy at her sports club. You think joining would be a good way to meet people, but then you decide to spend the money on a new stereo instead.

Six weeks after the move, things have turned out nothing like you had expected. You haven't made any real friends in the new town and you feel very lonely and isolated. You are very upset, and very surprised.

Procedure

The materials and instructions were presented in booklet form with another unrelated experiment (see Appendix 1a). The first page contained the instructions: "The following story describes a real-life experience. Please read it slowly and carefully, and try to imagine yourself in the situation described. Try to experience the events as if they are really happening to you, and imagine what you would think." The second page contained one of the eight versions of the scenario. At the end of the page, participants were instructed, "Before turning the page, please take a moment to consider the events that have occurred and imagine the thoughts you might have about them." The third page contained the information, "That night, while thinking over the

events of the past six weeks, you realise that you haven't written in your diary since the night before you moved. In the space below, jot down what you might write in your diary about your current situation and the events leading up to it. You may turn back and re-read the story first if you wish. Please try to fill the page." Participants could use the rest of the page to write their thoughts.

Coding

The number and content of causal and counterfactual thoughts in each diary entry were recorded. A causal thought was defined as "any statement that attempts to explain why the outcome occurred", e.g., "I haven't made friends because I didn't go to that party". A counterfactual thought was defined as "any change to a scenario event that would change the outcome", e.g., "If only I had gone to that party, I would have made friends". Participants' causal and counterfactual thoughts were also classified by content based on whether they focused on general factors (either internal factors or external factors), or else any of the four specific antecedents mentioned in the scenario (either the inhibitory antecedents, or the facilitative antecedents). For example, both of the above examples focus on missing the neighbours' party, which is a specific inhibitory antecedent. All responses were coded by the author and by a second rater who was unaware of the hypotheses and experimental conditions. Initial interrater agreement was 79% for causal thoughts and 90% for counterfactual thoughts. Discrepancies were resolved by discussion. An example of a causal thought is, "I haven't made an effort to get to know people so obviously I can expect to feel isolated" and an example of a counterfactual thought is, "if I had chosen to go to the other functions I could have made more friends" (further examples are provided in Appendix 2).

Results and Discussion

Frequencies of causal and counterfactual thoughts

A total of 402 causal thoughts were produced, a mean of 1.62 per participant, ranging from 0 to 6 per person, as shown in Table 2.1. A total of 197 counterfactual thoughts were produced, a mean of 0.79 per participant, ranging from 0 to 4 per person.

Table 2.1 Frequencies (with means per participant in parentheses) of causal and counterfactual thoughts by condition in Experiment 1

Type of outcome		Thoughts	
		Causal	Counterfactual
Bad	Unexpected		
	Controllable (n = 32)	70 (2.19)	56 (1.75)
	Uncontrollable (n = 31)	58 (1.81)	26 (0.81)
	Expected		
	Controllable (n = 30)	51 (1.70)	49 (1.63)
	Uncontrollable (n = 30)	56 (1.93)	19 (0.65)
Good	Unexpected		
	Controllable (n = 32)	45 (1.45)	13 (0.42)
	Uncontrollable (n = 31)	48 (1.55)	9 (0.29)
	Expected		
	Controllable (n = 29)	36 (1.20)	17 (0.57)
	Uncontrollable (n = 33)	38 (1.15)	8 (0.24)
Total		402 (1.62)	197 (0.79)

Overall, participants produced more causal thoughts than counterfactual thoughts ($M = 1.62$ vs $M = 0.79$, Wilcoxon test, $z = 6.92$, $n = 248$, $p < 0.01$). As Table 2.1 shows, in six of the eight conditions, participants produced more causal than counterfactual thoughts, bad unexpected uncontrollable: $z = 2.31$, $n = 31$, $p = 0.02$; bad expected uncontrollable: $z = 3.39$, $n = 30$, $p <$

0.01; good unexpected controllable: $z = 3.13$, $n = 32$, $p < 0.01$; good unexpected uncontrollable: $z = 3.7$, $n = 31$, $p < 0.01$; good expected controllable: $z = 2.31$, $n = 29$, $p = 0.02$; and good expected uncontrollable: $z = 3.55$, $n = 33$, $p < 0.01$. The two exceptions were the bad, controllable outcome conditions, and in these conditions, participants produced as many counterfactual thoughts as causal thoughts, for both the unexpected version, $z = 1.49$, $n = 32$, $p = 0.14$; and the expected one, $z = 0.05$, $n = 30$, $p = 0.96$.

We carried out Mann-Whitney U tests to examine the effects of our manipulations on the frequencies of causal and counterfactual thoughts. As we predicted from the substantive hypothesis, participants produced more causal thoughts in the bad outcome conditions than in the good outcome conditions ($M = 1.91$ vs $M = 1.34$, $U = 5924$, $n = 248$, $p < 0.01$), and likewise, they produced more counterfactual thoughts for bad outcomes than for good outcomes ($M = 1.22$ vs $M = 0.38$, $U = 4384.5$, $n = 248$, $p < 0.01$). The difference between bad and good outcomes tended to be greater for counterfactual thoughts than for causal thoughts, although this difference did not reach significance ($U = 7068.5$, $n = 248$, $p = 0.26$). Also as predicted, more causal thoughts were produced for unexpected outcomes than for expected outcomes ($M = 1.79$ vs $M = 1.47$, $U = 6708.5$, $n = 248$, $p = 0.04$), whereas there was no difference in the frequencies of counterfactual thoughts for unexpected and expected outcomes ($M = 0.76$ vs $M = 0.86$, $U = 7500$, $n = 248$, $p = 0.36$). The difference between unexpected and expected outcomes tended to be greater for causal than for counterfactual thoughts, although this difference did not reach significance ($U = 7120$, $n = 248$, $p = 0.31$). Finally, as predicted, there were as many causal thoughts to the scenarios with controllable outcomes as uncontrollable ones ($M = 1.72$ vs $M = 1.6$, $U = 7559$, $n = 248$, $p = 0.41$), whereas there were more counterfactuals to the scenarios with controllable outcomes than uncontrollable ones ($M = 1.1$ vs $M = 0.5$, $U = 5662.5$, $n = 248$, $p <$

0.01). The difference between controllable and uncontrollable outcomes was significantly greater for counterfactual thoughts than for causal thoughts ($U = 6262.5$, $n = 248$, $p = 0.01$).

We used Wilson's (1956) extension of the median test⁴ to examine possible interactions between the three variables, valence, expectancy and controllability. Like the Mann Whitney tests, the Wilson tests showed that participants produced more causal thoughts for bad outcomes than for good outcomes, $\chi^2(1) = 4.12$, $p < 0.05$, and they produced more counterfactual thoughts for bad outcomes than for good outcomes, $\chi^2(1) = 37.32$, $p < 0.01$. Participants produced more causal thoughts for unexpected outcomes than for expected outcomes, although this difference did not reach significance, $\chi^2(1) = 1.96$, $p > 0.1$, whereas they produced as many counterfactuals for unexpected outcomes as expected outcomes, $\chi^2(1) = 0.02$, $p > 0.9$. Finally, participants produced as many causal thoughts for controllable outcomes as uncontrollable ones, $\chi^2(1) = 8 \times 10^{-7}$, $p > 0.99$, whereas they produced more counterfactual thoughts for controllable outcomes than uncontrollable ones, $\chi^2(1) = 9.3$, $p < 0.01$. The chi-square values for the interaction effects were non-significant for both causal thoughts, $\chi^2(1) = 3.16$, $p > 0.05$, and counterfactual thoughts, $\chi^2(1) = 2.18$, $p > 0.1$, which shows that valence, expectancy and controllability had independent effects on the frequencies of causal and counterfactual thoughts.

Content of causal and counterfactual thoughts

As Table 2.2 shows, participants' causal and counterfactual thoughts focused on specific antecedents that were explicitly mentioned in the scenario, and on general factors that might

⁴ Wilson's (1956) analysis is "a distribution-free (i.e., non-parametric) test of the hypotheses concerning main effects and interactions ordinarily tested by...analysis of variance" (p. 96). The median for the entire set of observations is determined, and contingency tables are set up showing the frequencies of observations above and below the median in each condition. Chi-square values are computed to assess the main effects and interaction effects (for further details see Wilson, 1956).

explain a similar outcome in a different set of circumstances. They focused on *specific inhibitory* antecedents, that is, missing the neighbours' party or not joining the sports club (32%), or on *specific facilitative* antecedents, that is, going to the staff dinner or meeting an old friend (23%). They also focused on *general internal* factors such as their own characteristics, decisions, and effort (20%), or *general external* factors, such as luck, the availability of opportunities, and other people's actions (17%). Examples of causal and counterfactual thoughts in each of the content categories are shown in Appendix 2.

Table 2.2 Percentages of causal and counterfactual thoughts in each content category in Experiment 1

Categories		Responses		
		Causal	Counterfactual	All responses
Specific:				
	inhibitory	14	68	32
	facilitative	33	3	23
	total	47	71	55
General:				
	internal	24	14	20
	external	26	0	17
	total	49	14	38
Other:	total	4	15	8

As predicted, participants' causal thoughts were more likely than their counterfactual thoughts to focus on general factors (49% vs 14%) rather than specific antecedents mentioned in the scenario (47% vs 71%), and this difference was reliable, Wilcoxon test, $z = 4.44$, $n = 248$, $p < 0.01$. Causal thoughts focused equally on specific antecedents and on general factors (47% vs 49%, $z = 0.26$, $n = 248$, $p = 0.79$). In contrast, counterfactual thoughts were more likely to focus on

specific antecedents than on general factors (71% versus 14%, $z = 6.72$, $n = 248$, $p < 0.01$). Also as predicted, counterfactual thoughts focused more on internal than external general factors (14% vs 0%, $z = 3.91$, $n = 248$, $p < 0.01$), whereas causal thoughts focused equally on internal and external general factors (24% vs 26%, $z = 0.55$, $n = 248$, $p = 0.58$), and this difference was reliable, $z = 2.02$, $n = 248$, $p = 0.04$. Also as predicted, counterfactual thoughts focused more on inhibitory than facilitative specific antecedents (68% vs 3%, $z = 7.87$, $n = 248$, $p < 0.01$), whereas causal thoughts focused more on facilitative than inhibitory specific antecedents (33% vs 14%, $z = 4.5$, $n = 248$, $p < 0.01$) and this difference was reliable, $z = 8.58$, $n = 248$, $p < 0.01$.

The results show that when people think spontaneously about a situation -- that is, when they remain unprompted to consider how things might have been different, or to explain how things came to be as they are -- their spontaneous causal and counterfactual thoughts are activated in situations that differ in some respects and are similar in other respects. Causal and counterfactual thinking are similar in that both are evoked significantly more by bad outcomes than by good outcomes. They seem to be different in the tendency for outcome valence to have a greater effect on the frequency of counterfactual thoughts than on the frequency of causal thoughts: participants produced three times as many counterfactual thoughts for bad outcomes as for good outcomes, but they produced only one and a half times as many causal thoughts for bad outcomes as for good outcomes. They are also different in that causal thoughts are evoked more often by unexpected than expected outcomes, whereas counterfactual thoughts are evoked equally by expected and unexpected outcomes. They are also different in that counterfactual thoughts are evoked more often by controllable than uncontrollable outcomes, whereas causal thoughts are evoked equally by controllable and uncontrollable outcomes. Perhaps as a result of these differences, people generate more causal thoughts than counterfactual thoughts overall.

However, following controllable bad outcomes, they generate as many counterfactual as causal thoughts. The content of spontaneous causal and counterfactual thoughts also differs in systematic ways. In particular, causal thoughts are more likely than counterfactual thoughts to focus on general factors rather than antecedents specific to a particular outcome, and counterfactual thoughts are more likely than causal thoughts to focus on internal rather than external factors, and inhibitory rather than facilitative antecedents.

The experiment suggests that the determinants of spontaneous causal and counterfactual thoughts exhibit systematic differences. The results are consistent with the suggestion that causal thinking is concerned with the prediction of future outcomes, whereas counterfactual thinking is concerned with the prevention of unwanted outcomes. Our next experiment rules out a competing interpretation of the results.

Experiment 2

The aim of the second experiment was to investigate a possible alternative to our explanation of the results of the first experiment. We interpreted the differences we found in the activation of spontaneous causal and counterfactual thoughts as evidence for differences in the functions of causal and counterfactual thinking. However, these differences could have arisen because causal and counterfactual thinking are subject to different *cognitive constraints*, rather than because they serve different purposes.

One cognitive factor that may constrain the *content* of counterfactual thoughts but not causal thoughts is the availability of imaginary alternatives to events (e.g., Kahneman & Tversky, 1982). In order to construct a counterfactual scenario, people have to mutate antecedent events, and it may be cognitively easier to mutate some types of events than others. For example,

studies of directed counterfactual thinking show that people tend to mutate unusual antecedents rather than normal ones (e.g., Kahneman & Tversky, 1982; Wells et al., 1987; see also Chapter 1, pp. 32-34), perhaps because unusual events automatically bring to mind their normal counterparts, whereas normal events are less likely to evoke counterfactual alternatives (e.g., Kahneman & Miller, 1986). Likewise, people tend to mutate controllable antecedents rather than uncontrollable ones (e.g., Branscombe, N'gbala, Kobryniewicz & Wann, 1997; Girotto et al., 1991; Markman et al., 1995; N'gbala & Branscombe, 1995; see also Chapter 1, pp. 37-38), perhaps because controllable events bring to mind counterfactual alternatives more readily than uncontrollable events. Participants in our first experiment might have produced more counterfactual thoughts about controllable outcomes not because the function of counterfactual thinking is to allow the prevention of unwanted outcomes, but because counterfactual alternatives are easier to generate when an outcome has controllable antecedents. In contrast, the frequency of causal thoughts might have been unaffected by controllability because causal thinking does not require the generation of alternatives to events.

To test this possibility, we carried out a second experiment which examined the effects of antecedent normality on the frequency of spontaneous causal and counterfactual thoughts. People may generate more spontaneous counterfactual thoughts about outcomes with controllable antecedents than uncontrollable ones, simply because counterfactual alternatives are more available for controllable antecedents. If so, they should also generate more counterfactual thoughts once alternatives to the antecedents are made available, regardless of whether the antecedents are controllable or uncontrollable. For example, they should generate more counterfactual thoughts about outcomes with unusual antecedents than normal ones, regardless of whether they are controllable or uncontrollable, because counterfactual alternatives are also

more available for unusual events. In contrast, if the frequency of counterfactual thinking is largely determined by its function to allow the prevention of unwanted outcomes, then the normality of antecedents should have no effect on the frequency of counterfactual thoughts. There does not seem to be any reason why people should be more or less concerned about prevention depending on whether the antecedents of an outcome are unusual or normal.

The effect of antecedent normality on the *content* of *directed* counterfactual thoughts has been reliably demonstrated (e.g., Gavanski & Wells, 1989; Kahneman & Tversky, 1982; Wells et al., 1987; see also Chapter 1, pp. 32-34). In contrast, the two previous studies which have examined its effects on the *activation* of *spontaneous* counterfactual thoughts have produced inconclusive results. Roese and Olson (1996) gave participants scenarios in which a man gets delayed in a traffic jam after taking either his normal route or an unusual route to the airport. Participants produced more spontaneous counterfactual thoughts about the unusual route version than the normal route version. However, the normal route was described as the character's "favoured" route. We suggest that this description may have suggested that traffic jams were more likely to occur on the unusual route than on the normal route, and so the normality of the route may have been confounded with the likelihood that taking an alternative route would have prevented the outcome. Participants may have generated more counterfactual thoughts about the unusual route version not because counterfactual thoughts are more available for unusual antecedents, but because the outcome seemed more preventable in the unusual route version. This speculation is supported by the fact that there was no effect of antecedent normality on the activation of counterfactual thinking in a different experiment which did not have this possible confound (Roese & Olson, 1997; see also Chapter 1, pp. 19-20). We are unaware of any previous studies which have examined the effects of antecedent normality on the activation of

spontaneous causal thinking.

In this experiment, we confined our attention to scenarios with bad and unexpected outcomes, since these scenarios evoked the highest frequency of causal and counterfactual thoughts in the previous experiment. We gave participants the same scenario as in the previous experiment, but instead of manipulating the controllability of the antecedents, we used scenarios in which the antecedents were all uncontrollable, and we manipulated their normality. We predicted that antecedent normality would not affect the frequency of spontaneous counterfactual thoughts, but that it would affect the frequency of causal thoughts. We used scenarios in which the bad outcome was described as unusual for the protagonist (the first paragraph states, “I’ve never had any trouble making new friends”). If the frequency of causal thinking is determined by its function to allow the prediction of future outcomes, then people may produce more causal thoughts when an unusual outcome has unusual antecedents than when it has normal antecedents. Normal antecedents are present whether or not an unusual outcome occurs, and so they cannot be reliable predictors of an unusual outcome. In contrast, unusual antecedents which are present when an unusual outcome occurs would be more likely to be good predictors of the outcome.

Method

Participants and Design

The 84 undergraduates from a variety of departments at the University of Dublin and University College Dublin, who took part in the experiment voluntarily, included 55 women and 29 men. Their average age was 20 years, with a range from 17 to 45 years. They were randomly assigned to two groups (unusual antecedents or normal antecedents).

Materials

We constructed two versions of a scenario based closely on the scenarios used in Experiment 1. Both versions had a bad, unexpected, uncontrollable outcome. They differed in whether each of the four antecedent events was described as unusual or normal, as follows (see also Appendix 1b):

You're moving house to start a new job in a different city. The night before you leave, you write down your thoughts about the move in your diary: ...I've got mixed feelings about moving to a place where I know hardly anyone - my friends and social life are so important to me. But I'm sure it will be easy to settle in to the new town - I've never had any trouble making new friends...

A lot happens in your first two weeks in the new town. During your first week at work, a staff dinner is held. You have to go because your boss has asked all the staff to be there, (as he always does / which he almost never does). You enjoy the evening and meet a lot of people. That weekend, your next-door neighbours invite you to a party. Most of the people who live on your road will be there. However, you suffer from very (frequent / occasional) migraines, and as very (often / rarely) occurs, that evening you have a really bad one, so you can't go. The next week, as (regularly / very unusually) happens, you happen to bump into an old friend who lives in the town and he insists on showing you around. He takes you out the following evening and he introduces you to a lot of his friends. A few days later, a colleague tells you there's a membership vacancy at her sports club. You think joining would be a good way to meet people. However, (as you would never have that kind of money, / although you would very rarely not have enough money, at that time) there's no way you can afford the membership fee.

Six weeks after the move, things have turned out nothing like you had expected. You haven't made any real friends in the new town and you feel very lonely and isolated. You are very upset, and very surprised.

Procedure

The instructions were identical to those used in Experiment 1. Participants were directed to read the scenario and to try to experience the events as if they were really happening. After reading the scenario, they were instructed to write a one-page diary about their imagined

experience (see Appendix 1b).

Coding

The numbers of causal and counterfactual thoughts in each diary entry were recorded using the same coding scheme as in Experiment 1. All responses were coded by the author and by a second rater who was unaware of the hypotheses or experimental conditions. Initial interrater agreement was 85% for causal thoughts and 93% for counterfactual thoughts. Discrepancies were resolved by discussion. An example of a causal thought is “Things have not turned out well at all, maybe it’s just bad luck”. An example of a counterfactual thought is “If I’d gone to my neighbours’ party and didn’t have the migraine I could have made some real friends”.

Results and Discussion

As Table 2.3 shows, a total of 110 causal thoughts were produced, a mean of 1.31 per participant, ranging from 0 to 5 per person. A total of 38 counterfactuals were produced, a mean of 0.45 per participant, ranging from 0 to 3 per person. As in Experiment 1, participants produced more causal than counterfactual thoughts overall ($M = 1.31$ vs $M = 0.45$, Wilcoxon test, $z = 4.87$, $n = 84$, $p < 0.01$), and they also did so in both conditions (unusual: $M = 1.62$ vs $M = 0.36$, $z = 4.38$, $n = 42$, $p < 0.01$; normal: $M = 1.00$ vs $M = 0.55$, $z = 2.26$, $n = 42$, $p = 0.02$). As Table 2.3 shows, participants produced more causal thoughts when the scenario had unusual antecedents than when it had normal antecedents, and this difference was reliable ($M = 1.62$ vs $M = 1.00$, $U = 671$, $n = 84$, $p = 0.03$). In contrast, participants who were given the scenario with unusual antecedents produced the same amount of counterfactuals as those given the version with normal antecedents. In fact, those in the normal condition produced somewhat more

counterfactuals than those in the unusual condition, although this difference was not significant ($M = 0.55$ vs $M = 0.36$, $U = 798$, $n = 84$, $p = 0.36$).

Table 2.3 Frequencies (with means per participant in parentheses) of causal and counterfactual thoughts by condition in Experiment 2

<i>Type of antecedents</i>	<i>Thoughts</i>	
	Causal	Counterfactual
Unusual (n = 42)	68 (1.62)	15 (0.36)
Normal (n = 42)	42 (1.00)	23 (0.55)
<i>Total</i>	<i>110 (1.31)</i>	<i>38 (0.45)</i>

The results suggest that the spontaneous activation of causal and counterfactual thoughts is primarily determined by their different functions, rather than their availability. Studies of directed counterfactual thinking suggest that counterfactual alternatives come to mind more easily for unusual events than normal events, yet we found that participants did not generate more spontaneous counterfactual thoughts about an outcome with unusual antecedents. This finding does not support the idea that participants in Experiment 1 generated more counterfactual thoughts about outcomes with controllable rather than uncontrollable antecedents because counterfactual alternatives come to mind more easily for controllable than uncontrollable events. In contrast, it should be equally easy to generate causal thoughts focusing on unusual and normal antecedents, as causal thinking does not require the construction of imaginary alternatives to events, yet participants produced more causal thoughts for outcomes with unusual antecedents. This finding is consistent with the idea that causal thinking is directed towards the prediction of

future outcomes, because unusual antecedents may generally be better predictors of an unusual outcome than normal antecedents. In contrast, participants were no more likely to generate counterfactual thoughts about outcomes with unusual antecedents because counterfactual thinking is directed towards the prevention of unwanted outcomes, and unusual antecedents are no better preventors than normal antecedents. The findings support our conclusion that participants in Experiment 1 produced more counterfactual thoughts about controllable outcomes, not because counterfactual alternatives were readily available to them, but because controllable outcomes can be prevented.

Our first two experiments show that there are systematic differences in the activation and content of spontaneous causal and counterfactual thoughts. However, our conclusions about the content of causal and counterfactual thoughts were based on coding participants' spontaneous thoughts into content categories, rather than examining the effects of specific manipulations on their contents. The next experiment examines the effects of two key variables on the content of participants' directed causal and counterfactual thoughts.

Experiment 3

The aim of the third experiment was to compare systematically the effects of covariation and controllability information on the content of causal and counterfactual thoughts. Mandel and Lehman (1996) suggest that people focus on covarying events in causal thinking and this helps them to predict future outcomes, whereas they focus on controllable events in counterfactual thinking and this helps them to prevent unwanted outcomes. However, in their experiments, information about covariation and controllability was neither explicitly provided, nor systematically manipulated. We aimed to remedy this problem in this experiment.

As we outlined earlier, participants in Mandel and Lehman's (1996) studies read scenarios that included two focal antecedents, one of which was assumed would be perceived as covarying with the outcome (e.g., reckless driving in a car accident scenario), and one of which was assumed would be perceived as a controllable event (e.g., the accident victim's choice of route). However, there was no evidence that participants had actually perceived the events to differ in these distinct ways. Furthermore, the events also differed in many other ways, and so differences in participants' tendencies to focus on them in causal or counterfactual thinking could have arisen for other reasons. For example, the controllable event always occurred first in the scenarios and the covarying event always occurred last, but it is well-known that people's counterfactual thoughts tend to focus on the first event in a causal chain (e.g., Wells et al., 1987), whereas they tend to judge as most causal events which occur closest to an outcome (e.g., Siegler & Liebert, 1974; see also Chapter 1, pp. 23-24).

In this experiment, we presented participants with a car accident scenario in which the central character, Mr. Jones, was driving carefully and another character, Mark Smith, was driving at high speed. The outcome was a collision of their two cars. The scenario was based on two focal antecedent events, the time Mr. Jones left work, and the route he took home (Kahneman & Tversky, 1982; see Chapter 1, p. 32). Unlike in Mandel and Lehman's studies, we provided participants with *explicit* covariation and controllability information, rather than relying on their possibly differing background knowledge. We varied which of the two antecedents, either the time or the route, was described explicitly as increasing the likelihood of an accident, and which of the two events was described explicitly as being more under Mr. Jones's control. We also controlled the order in which the two antecedents were presented to ensure that the finding of differences in counterfactual and causal focus is not an artefact due to order.

Participants were asked to make causal judgements about the antecedents and to generate counterfactual alternatives to the outcome, as in previous related studies (Mandel & Lehman, 1996; N'gbala & Branscombe, 1995), and we controlled the order in which participants completed the causal and counterfactual thinking tasks.

We expected that the antecedent event (either the time or the route) which we described as increasing the likelihood of an accident (i.e., covarying with the outcome) would be rated as more causal than the other *and* be mutated more readily than the other. The fact that the occurrence of an antecedent increased the likelihood of an outcome implies that its non-occurrence (relative to its occurrence) would have decreased the likelihood of the outcome, and so the occurrence of a covarying antecedent would help predict an outcome and its non-occurrence would help prevent it.

We expected that the antecedent event (either the time or the route) which we described as more under Mr. Jones's control would be mutated more readily than the other, as in past findings (e.g., Girotto et al., 1991; Markman et al., 1995; N'gbala & Branscombe, 1995), but that it would not be rated as more causal than the other. More controllable events should be better preventors than less controllable events, but more and less controllable events should be equally good predictors.

Of course, a key feature in the scenario is the reckless driving of Mark Smith, and it makes sense that it would be rated as more causal than either the time or the route -- participants' prior knowledge is bound to suggest that reckless driving covaries with accidents in general, and so reckless driving would be the best predictor of future accidents. However, we expected that Mark Smith's behaviour would not be most likely to be mutated, as his behaviour would be a poor preventor because it is totally uncontrollable from Mr. Jones's point of view.

Method

Materials and Design

We constructed four scenarios, based on Kahneman and Tversky's (1982) scenario about Mr. Jones's car accident. In each of the four scenarios, Mr. Jones left work at an unusual time and drove home by an unusual route. The scenarios differed in whether the time or the route was described as being more under Mr. Jones's control, and whether the time or the route was described as increasing the likelihood of having a car accident. An example of one of the scenarios, the time-controllable / route-uncontrollable⁵ and time-covarying / route-non-covarying scenario, is as follows (see also Appendix 1c):

Mr. Jones is 47 years old, the father of three and a successful banking executive. He was recently involved in a serious car accident.

On the day of the accident, Mr. Jones did not leave work at his usual time. He decided to leave earlier than usual as it was an exceptionally sunny day and he wanted to get out and enjoy it. Mr. Jones did not drive home by his usual route. He had to drive along the coast road to do some important errands for his wife, who was sick at home.

Mr. Jones knew that accidents were much more likely to happen at that time than at the time he usually drove home, because speeding was so much more common at that time. So, although he was aware that the coast road was safer than the route he usually took, Mr. Jones drove especially carefully.

The accident happened when Mr. Jones was about half-way home. He slowed down as he approached a bend in the road. At that moment a sports car, driven by a young man named Mark Smith, came around the corner at very high speed. Smith lost control, attempted to swerve to avoid the oncoming traffic, and crashed into the side of Mr. Jones's car. Mr. Jones was seriously injured and spent three weeks recovering in hospital. As commonly happens in such situations, he often thought about the accident for some time afterwards.

⁵ We intended participants to perceive both the time and the route as somewhat controllable, but to perceive one as more controllable than the other. However, for ease of expression, we will refer to the time and the route as being either "controllable" or "uncontrollable".

We assigned participants at random to the time-controllable, time-covarying scenario ($n = 39$), the time-controllable, route-covarying scenario ($n = 40$), the route-controllable, time-covarying scenario ($n = 38$), and the route-controllable, route-covarying scenario ($n = 39$). We constructed four versions for each of the four scenarios, in which we varied whether the information about the time or the route was presented first or second for the covariation manipulation and for the controllability manipulation. One quarter of the participants in each group were given each order.

The participants carried out two tasks. The causal task asked them to rate from 0 (not at all causal) to 8 (very causal) the extent to which each of the following factors caused Mr. Jones's accident: (a) the time Mr. Jones left work, (b) the route Mr. Jones took, (c) Mark Smith's driving, (d) Mr. Jones's driving, (e) Mrs. Jones's illness, (f) the weather (in a pre-test in which we asked ten participants, who were not taking part in this experiment, to read versions of the scenario and list the causes of the outcome, these six factors were the most frequently mentioned causes). The order in which the events were presented was varied, so that each event appeared in each position equally often. The counterfactual task asked them to provide their counterfactual thoughts, that is, to write down four ways that Mr. Jones's accident could have been avoided "if only....". Half the participants in each group were given the causal task first and the counterfactual task second, and the other half were presented with the tasks in the opposite order (see also Appendix 1c).

We also gave the participants two additional questions, as manipulation checks, again presented in counterbalanced order. To assess the effectiveness of the covariation manipulation, they rated from 1 (not at all) to 7 (very much), how much the likelihood of Mr. Jones having an

accident was increased by: (a) the time he left work, (b) the route he took home. To assess the controllability manipulation, they rated from 1 (no control) to 7 (total control), how much control Mr. Jones had over: (a) the time he left work, (b) the route he took home (see also Appendix 1c).

Participants and Procedure

The 156 undergraduates from a variety of departments at the University of Dublin, who took part in the experiment voluntarily, included 112 women, 42 men and two people who did not record their gender. Their mean age was 22 years, with a range from 17 to 52 years. The materials were presented with another unrelated experiment in a booklet that contained the instructions, the story, the counterfactual and causal tasks and the two manipulation check tasks, each on separate pages. The instructions were: "Please read the following story slowly and carefully and then answer the questions that follow. Please take your time, and work through the tasks in the order in which they are presented. Please do not look ahead, or go back over previous answers." The participants worked at their own pace and wrote their answers in the spaces provided on the question sheet (see also Appendix 1c).

Results and Discussion

Manipulation checks

Our manipulations were largely effective, as shown by the results of a two (time-controllable versus route-controllable) by two (time-covarying versus route-covarying) ANOVA. The time was rated as significantly more controllable in the time-controllable conditions than in the route-controllable conditions [$M = 6.16$ vs $M = 4.75$, $F(1, 152) = 31.36$, $p < 0.01$]. The route was rated as significantly more controllable in the route-controllable conditions than in the time-controllable conditions [$M = 6.64$ vs $M = 5.22$, $F(1, 152) = 31.13$, $p < 0.01$]. The time was

judged to have increased the likelihood of Mr. Jones having an accident significantly more in the time-covarying conditions than in the route-covarying conditions [$M = 4.13$ vs $M = 3.33$, $F(1, 152) = 8.00$, $p < 0.01$]. The route was judged to have increased the likelihood of an accident more in the route-covarying conditions than in the time-covarying conditions, although the difference did not reach significance [$M = 4.81$ vs $M = 4.60$, $F(1, 152) = 1.37$, $p = 0.24$].

Causal ratings

As we predicted, participants rated the covarying event as more causal than the non-covarying event (63% vs 37%, binomial $z = 2.61$, $n = 107$, $p < 0.01$), as Table 2.4a shows (and the covarying events are underlined in Table 2.4b). As expected, they did not rate the controllable event as more causal than the uncontrollable event, in fact, their ratings show the opposite trend (38% vs 62%, binomial $z = 2.42$, $n = 107$, $p < 0.01$), as Table 2.4a shows (and the controllable events are in bold in Table 2.4b).

As expected, participants rated Mark Smith's driving as more causal than either the time ($M = 6.07$ vs $M = 3.75$, $t(155) = 6.96$, $p < 0.01$) or the route ($M = 6.07$ vs $M = 4.46$, $t(155) = 5.76$, $p < 0.01$), as Table 2.5 shows. Overall they rated the route as more causal than the time ($M = 4.46$ vs $M = 3.75$, $t(155) = 3.48$, $p < 0.01$).

Counterfactual thoughts

As Table 2.5 shows, participants' first counterfactual thoughts were equally likely to focus on the time, or the route, or Mark Smith's actions (time vs route: 25% vs 30%, binomial $z = 0.86$, $n = 86$, $p = 0.39$; time vs Smith: 25% vs 26%, binomial $z = 0.11$, $n = 79$, $p = 0.91$; route vs Smith: 30% vs 26%, binomial $z = 0.75$, $p = 0.45$), and these three antecedents were also mentioned equally often overall (time vs route: 20% vs 20%, binomial $z = 0.27$, $n = 215$, $p =$

0.79; time vs Smith: 20% vs 21%, binomial $z = 1.4$, $n = 130$, $p = 0.16$; route vs Smith: 20% vs 21%, binomial $z = 0.27$, $n = 222$, $p = 0.79$).

Table 2.4a Percentages of participants who rated the covarying or non-covarying event and the controllable or uncontrollable event as more causal in Experiment 3

<i>Controllability</i>	<i>Covariation</i>		
	Covarying	Non-covarying	<i>Total</i>
Controllable	68	32	38
Uncontrollable	59	41	62
<i>Total</i>	63	37	100

Table 2.4b Percentages of participants who rated the time or the route as more causal for each scenario in Experiment 3

<i>Scenario</i>	<i>Event</i>	
	Time	Route
Time controllable		
Time covaries (n = 31)	<u>30</u>	70
Route covaries (n = 27)	16	<u>84</u>
<i>total (n = 58)</i>	22	78
Route controllable		
Time covaries (n = 21)	<u>62</u>	38
Route covaries (n = 28)	29	<u>71</u>
<i>total (n = 49)</i>	43	57
<i>Total (n = 107)</i>	32	68

Notes: Data from 49 of the 156 participants were excluded because they rated the time and the route as equally causal. Covarying events are underlined; controllable events are in bold.

Table 2.5 Mean causal ratings and percentages of counterfactual mutations of events in Experiment 3

<i>Event</i>	<i>Mean causal rating</i>	<i>%s of first counterfactuals</i>	<i>%s of all counterfactuals</i>
Time	3.75	25	20
Route	4.46	30	20
Smith	6.07	26	21
Jones	2.52	6	9
Illness	3.29	7	15
Weather	3.23	2	8
Other		4	6

Overall, 83% of participants mutated either the time or the route at some point in their counterfactual thoughts. As predicted, they tended to mutate the covarying event earlier than the non-covarying event (57% vs 43%, binomial $z = 1.58$, $n = 130$, $p < 0.06$), as Table 2.6a shows (and the covarying events are underlined in Table 2.6b), and they tended to mutate the controllable event earlier than the uncontrollable event (also 57% vs 43%, binomial $z = 1.58$, $n = 130$, $p < 0.06$), as Table 2.6a shows (and the controllable events are in bold in Table 2.6b).

Table 2.6a Percentages of participants whose counterfactuals mutated the covarying or non-covarying event and the controllable or uncontrollable event earlier in Experiment 3

<i>Controllability</i>	<i>Covariation</i>		
	Covarying	Non-covarying	<i>Total</i>
Controllable	55	45	57
Uncontrollable	59	41	43
<i>Total</i>	57	43	100

Table 2.6b Percentages of participants whose counterfactuals mutated the time or the route earlier for each scenario in Experiment 3

Scenario	Event	
	Time	Route
Time controllable		
Time covaries (n = 33)	<u>55</u>	45
Route covaries (n = 34)	47	<u>53</u>
total (n = 67)	51	49
Route controllable		
Time covaries (n = 32)	47	53
Route covaries (n = 31)	26	74
total (n = 63)	37	63
Total (n = 130)	44	56

Notes: Data from 26 of the 156 participants were excluded because they did not mutate either the time or the route in any of their counterfactual responses. Covarying events are underlined; controllable events are in bold.

Covarying events are perceived as more causal than non-covarying events, and this result corroborates previous suggestions (Mandel & Lehman, 1996). Our experiment provides the first demonstration that participants give higher causal ratings to the antecedent event, whether the time or the route, that is described explicitly as increasing the likelihood of an accident. It also replicates the finding that people judge the reckless driver the most important cause of the accident in these scenarios, perhaps as suggested by their background knowledge about events that covary with car accidents. Our experiment shows that counterfactual thoughts also focus on antecedents that are described as covarying with the outcome, rather than non-covarying

antecedents (although they are not most likely to focus on the reckless driver). Counterfactual thoughts focus on more controllable rather than less controllable antecedents, and this result corroborates earlier findings (Giroto et al., 1991; Markman et al., 1995; N'gbala & Branscombe, 1995). The experiment provides the first demonstration that less controllable antecedents may be perceived as more causal (if they covary with the outcome) than more controllable antecedents. Importantly, these experimental findings are provided for scenarios in which the covariation and controllability of antecedents were directly manipulated and the order in which the antecedents appeared was controlled.

The results of the third experiment provide further support for the view that causal thinking is concerned with the prediction of future outcomes, whereas counterfactual thinking is concerned with the prevention of unwanted outcomes. We propose that people may view covarying events as more causal than non-covarying events because the presence of a covarying antecedent helps predict an outcome, and they mutate covarying events rather than non-covarying events because the absence of a covarying antecedent helps prevent an outcome. They view events that general knowledge suggests covary with an outcome as most causal because they predict the outcome in a wider range of situations. They do not rate controllable events as more causal than uncontrollable events because controllable events need not be better predictors, but they mutate controllable events rather than uncontrollable events because controllable events are better preventors.

General discussion

The findings of our first three experiments show that there are systematic similarities and differences in the activation and content of causal and counterfactual thoughts. The results

corroborate the view that causal thinking is directed towards prediction, whereas counterfactual thinking is directed towards prevention.

The results of Experiment 1 showed that causal and counterfactual thoughts are both evoked more by bad outcomes than good outcomes -- the aim of both prediction and prevention may be to avoid bad outcomes. But bad outcomes tend to increase counterfactual thinking more than causal thinking -- it may be useful to predict both good and bad outcomes, whereas it is only necessary to prevent bad outcomes. Causal thoughts are evoked more by unexpected outcomes than expected outcomes -- unexpected outcomes indicate that prediction has failed -- whereas counterfactual thoughts are evoked equally by expected and unexpected outcomes. Counterfactual thoughts are evoked more by controllable outcomes than uncontrollable outcomes -- controllable outcomes indicate that prevention is possible -- whereas causal thoughts are evoked equally by controllable and uncontrollable outcomes.

Experiment 1 also showed that spontaneous causal and counterfactual thoughts differ in their content. Causal thoughts focus equally on general factors such as effort and specific antecedents mentioned in a scenario, whereas counterfactual thoughts focus mostly on specific antecedents. People may focus on general factors in causal but not counterfactual thinking because general factors enable future outcomes to be predicted in a broad range of situations. They may focus on specific antecedents in counterfactual thinking because these events are represented explicitly and hence are cognitively available (Legrenzi et al., 1993; Byrne, 1997), or because specific events can be easily converted into plans for preventing future outcomes (e.g., Taylor & Pham, 1996). Moreover, causal thoughts focus equally on internal and external factors, whereas counterfactual thoughts focus more on internal factors, perhaps because internal factors are potentially controllable and therefore preventable. Furthermore, causal thoughts focus more

on facilitative than inhibitory antecedents, whereas counterfactual thoughts focus more on inhibitory antecedents, perhaps because the absence of an inhibitory antecedent helps prevent a bad outcome.

The results of Experiment 2 ruled out the alternative possibility that people generate more spontaneous counterfactual thoughts about outcomes with controllable antecedents than uncontrollable ones because counterfactual alternatives are more cognitively available for controllable events (e.g., Branscombe et al., 1997; Markman et al., 1995). Participants generated as many spontaneous counterfactual thoughts about outcomes with unusual antecedents as normal ones, despite the well-established finding that counterfactual thoughts are more cognitively available for unusual events (e.g., Kahneman & Miller, 1986; Kahneman & Tversky, 1982). In contrast, participants engaged in more causal thinking about outcomes with unusual antecedents than normal ones, even though antecedent normality should not affect the cognitive ease of generating causal thoughts. The findings suggest that the frequency of spontaneous causal or counterfactual thoughts depends on their functions -- people may generate more causal thoughts about unusual outcomes when they have unusual antecedents than when they have normal antecedents because the function of causal thinking is to predict future outcomes, and unusual antecedents are generally better predictors of unusual outcomes than normal antecedents.

These findings generalise from spontaneous causal and counterfactual thoughts to directed thoughts, as Experiment 3 showed. Directed causal and counterfactual thoughts both focus on covarying rather than non-covarying events, when people are given explicit information about which events covary with the outcome. However, directed causal thoughts focus on uncontrollable rather than controllable antecedents, whereas directed counterfactual thoughts focus on controllable antecedents. Moreover, an uncontrollable event that prior knowledge

suggests covaries with the outcome is considered most causal, but it is not mutated more than the other events. Causal and counterfactual thoughts may both focus on covarying antecedents because the presence of a covarying antecedent helps predict an outcome, and the absence of a covarying antecedent helps prevent an outcome. Causal but not counterfactual thoughts may focus on antecedents which general knowledge suggests covary with an outcome because they predict the outcome in a wide range of situations. Counterfactual but not causal thoughts may focus on controllable antecedents because controllable antecedents are better preventors than uncontrollable antecedents.

A possible alternative explanation is that differences in the content of causal and counterfactual thoughts result from differences in their cognitive determinants. The findings of Experiment 2 suggest that the *activation* of causal and counterfactual thoughts depends on their functions, but past research suggests that the *content* of counterfactual thoughts depends on the availability of counterfactual alternatives. For example, people tend to mutate unusual rather than normal antecedents, perhaps because counterfactual alternatives come to mind more easily for unusual antecedents (e.g., Kahneman & Tversky, 1982). Likewise, counterfactual thoughts may focus on controllable rather than uncontrollable antecedents because counterfactual alternatives come to mind more easily for controllable antecedents (e.g., Girotto et al., 1991; Markman et al., 1995), whereas causal thoughts may focus on a broader range of factors because causal thoughts do not require the generation of alternatives to events. On this account, causal and counterfactual thinking may serve the same functions, even if they have different contents. We will examine this possibility in the next chapter, in which we compare the *consequences* of causal and counterfactual thoughts with the same and different contents.

Of course, the idea that causal thoughts are directed towards prediction and

counterfactual thoughts are directed towards prevention is not incompatible with the view that counterfactual thoughts are *also* used to test the necessity of potential causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). However, our findings are inconsistent with the idea that the principle function of counterfactual thinking is to determine the causes of events. On this account, causal and counterfactual thoughts should tend to be activated by the same sorts of outcomes and focus on the same sorts of antecedents (except that causal thoughts may focus on necessary and sufficient antecedents, whereas counterfactual thoughts may focus on necessary but insufficient antecedents). In contrast, our findings clearly demonstrate that there are differences in the activation and content of causal and counterfactual thoughts which are independent of necessity and sufficiency. The specific pattern of differences is consistent with the view that causal and counterfactual thinking serve different functions: causal thoughts are concerned with prediction, whereas counterfactual thoughts are concerned with prevention.

Chapter 3 The Consequences of Causal and Counterfactual Thoughts

The experiments reported in Chapter 2 showed that there are differences in the activation and content of causal and counterfactual thoughts which are consistent with the view that causal thoughts are concerned with prediction and counterfactual thoughts are concerned with prevention. However, differences in the content of causal and counterfactual thoughts do not provide direct evidence for differences in their functions, because the content of counterfactual thoughts may be determined by the cognitive availability of counterfactual alternatives to events as well as their function (see Chapter 2, p. 102). In contrast, evidence for differences in the consequences of causal and counterfactual thoughts would provide more direct evidence for differences in their functions. In this chapter, we will examine the hypothesis that causal thoughts are concerned with prediction and counterfactual thoughts are concerned with prevention, by systematically comparing the consequences of causal and counterfactual thoughts.

The idea that the function of counterfactual thinking is to allow the prevention of unwanted outcomes, whereas the function of causal thinking is to allow the prediction of future outcomes (Mandel & Lehman, 1996) suggests that causal and counterfactual thoughts should have different consequences for judgements related to prevention and prediction. To our knowledge, no previous studies have examined the consequences of either counterfactual or causal thinking for judgements about prevention, although there is evidence that counterfactual thoughts increase the perceived controllability of a past outcome (e.g., McMullen et al., 1995). The only previous study to compare the consequences of counterfactual and causal thinking examined their effects on judgements about prediction (Roese & Maniar, 1997), but this study only examined the consequences of thoughts about good outcomes, whereas most spontaneous

counterfactual and causal thoughts are evoked by bad outcomes as shown in Experiment 1 (see also Bohner et al., 1988; Roese & Olson, 1997; Sanna & Turley, 1996; Weiner, 1985; Chapter 1, pp. 17-19).

We suggest that counterfactual and causal thoughts may also have different consequences for emotions. Previous research suggests that *downward* counterfactuals about how an outcome could have been worse (e.g., “If I hadn’t been wearing my seatbelt, I could have been killed”) serve an emotional rather than preventative function (e.g., Markman et al., 1993). For example, participants directed to generate downward counterfactuals report more positive emotions, but lower perceived control, than those who generate *upward* counterfactuals about how an outcome could have been better (e.g., “If I hadn’t been speeding, I wouldn’t have had an accident”, e.g., McMullen et al., 1995; Roese, 1994). If the function of causal thinking is to allow people to predict future outcomes, rather than feel better about past outcomes, then downward counterfactuals and causal thoughts should have different consequences for emotions.

In this chapter, we will report three experiments which compare the consequences of counterfactual and causal thinking for judgements related to prevention and prediction, and for emotions. All of the experiments examined thoughts about bad outcomes, because most spontaneous counterfactual and causal thoughts are evoked by bad outcomes. In Experiment 4, participants imagined themselves in a scenario about an academic failure, and we compared the effects of counterfactual and causal thoughts on their judgements of the preventability, controllability and predictability of the outcome. In Experiment 5, participants read a scenario about a historical event, and we compared the effects of counterfactual and causal thoughts on their judgements of the predictability and likelihood of the outcome. In Experiment 6, participants recalled a recent negative life event and generated either upward counterfactuals,

downward counterfactuals or causal thoughts, focusing either on their own actions or on whatever events they wished. We compared the effects on six measures related to prevention, two measures related to prediction and six measures related to emotions.

Experiment 4

This experiment tested the idea that counterfactual thoughts are concerned with prevention, whereas causal thoughts are concerned with prediction, by comparing the effects of counterfactual and causal thoughts on participants' judgements of the preventability, controllability and predictability of an outcome. We are unaware of any previous research which has examined the effects of either counterfactual or causal thinking on perceived preventability. Moreover, no previous studies have *compared* the consequences of counterfactual and causal thoughts for perceived control, although there is evidence that counterfactual thoughts increase perceived control (e.g., McMullen et al., 1995). Only one previous study compared the consequences of counterfactual and causal thinking for judgements of predictability (Roese & Maniar, 1997), but this study only examined the effects of thoughts about good outcomes. This experiment compared the consequences of counterfactual and causal thoughts about a bad outcome.

Method

Participants and Design:

The 125 undergraduates from a variety of departments at the University of Dublin, who took part in the experiment voluntarily, included 97 women and 28 men. Their ages ranged from 17 to 51, with a mean age of 20 years (one person did not record their age). They were randomly assigned to the three conditions: counterfactual thinking (n = 43), causal thinking (n = 42) and

control (n = 40).

Materials and Procedure:

The experiment was run in two group sessions. Participants were presented with a three-page booklet containing the instructions and materials (see Appendix 1d). On the cover page, they were informed that the study was concerned with “how people think about everyday events”, and that there were no right or wrong answers. They were asked to complete the tasks in the booklet in the order in which they were presented.

The second page contained the instructions: “The following story describes a real-life experience. Please read it slowly and carefully, and try to imagine yourself in the situation described. Try to experience the events as if they are really happening to you, and imagine what you would think.” These instructions were followed by the scenario below (adapted from Kahneman & Tversky, 1982):

You are a student in your first year at University. You have to pass some end of year exams in order to continue into second year. Students are randomly assigned to be examined by either Dr. Smith in May or Dr. Jones in June. You are initially assigned to Dr. Smith in May, but you ask the course director if you can be reassigned to Dr. Jones in June, so that you’ll have more time to prepare for the exams. She agrees.

You sit your exams in June and you think they go fairly well. In July, both Dr. Smith and Dr. Jones post their students’ results on the notice board. You find your name on Dr. Jones’s list and learn that you have failed and will not be allowed into second year. You are very disappointed. Looking more closely at the lists, you notice that most of the other students examined by Dr. Jones also failed the year, whereas very few of those examined by Dr. Smith failed. You are very upset.

The scenario described a number of antecedents, including both personal actions and external factors, so that participants could easily generate either counterfactual or causal

thoughts. After reading the scenario, participants in the counterfactual thinking condition were asked to complete the following sentence stem in the way that seemed most natural to them: “You think that you would not have failed the year if only...”. Those in the causal thinking condition received the same instructions, but they completed the following sentence: “You think that you failed the year because...”. Participants in the control condition were instructed to write a one-sentence summary of the story. The third page contained the three dependent measures, presented in counterbalanced order. Participants were asked “To what extent could you have prevented failing the year?”, which they answered by circling a number on a scale from 1 (not at all preventable) to 9 (totally preventable). They were also asked “How much control did you have over whether or not you failed the year?”, which they rated from 1 (not at all controllable) to 9 (totally controllable), and “How predictable was it that you would fail the year?”, which they rated from 1 (not at all predictable) to 9 (totally predictable) (see also Appendix 1d).

Results and Discussion

Manipulation check

A manipulation check was carried out on the sentence completions of participants in the two experimental conditions, based on a dichotomous classification of whether or not they correctly recorded the type of thought corresponding to their condition (e.g., Roese, 1994). A completion was coded as a counterfactual thought if it stated how the outcome could have been different by changing some preceding event (e.g., “...if only *Dr. Jones was not such a difficult marker*”). A completion was coded as a causal thought if it stated why the outcome occurred without changing any preceding event (e.g., “...because *Dr. Jones may have marked the paper too harshly*”). The completions were coded by the author and by an independent rater who was

unaware of the purpose of the experiment. Initial interrater agreement was 99%. The one discrepancy was resolved by discussion. Ninety-eight per cent of participants in the experimental conditions provided a completion consistent with their condition. The remaining two participants were excluded from further analyses.

Content coding

Participants’ sentence completions were also categorised according to whether or not they mentioned *internal* factors such as their personal actions (e.g., “...if only *I had remained with Dr. Smith*”), and whether or not they mentioned *external* factors such as other people’s actions or situational factors (e.g., “...because *Dr. Jones may have marked the paper too harshly*”). The contents were coded by the author and by an independent rater who was unaware of the purpose of the experiment. Initial interrater agreement was 95%. The four discrepancies were resolved by discussion. 91% of participants who generated counterfactual thoughts mentioned internal factors, binomial $z = 5.34$, $n = 43$, $p < 0.01$, whereas 40 % of participants who generated causal thoughts mentioned internal factors, binomial $z = 1.26$, $n = 40$, $p = 0.21$, and this difference is significant, $\chi^2 (1) = 23.83$, $p < 0.01$. 90% of participants who generated causal thoughts mentioned external factors, binomial $z = 5.06$, $n = 40$, $p < 0.01$, whereas 12% of participants who generated counterfactual thoughts mentioned external factors, binomial $z = 5.03$, $n = 43$, $p < 0.01$, and this difference is significant, $\chi^2 (1) = 50.92$, $p < 0.01$ ⁶.

Preventability, controllability and predictability

As Table 3.1 shows, participants in all conditions tended to judge the outcome as

⁶ The totals sum to more than 100% because some participants mentioned both internal and external factors.

preventable and controllable, but unpredictable (mean ratings on scales of 1 to 9 of 6.6, 6.22 and 3.62, respectively).

Table 3.1: Mean ratings of preventability, controllability and predictability following different sentence completion tasks in Experiment 4

Sentence completion task	Rating		
	Preventability	Controllability	Predictability
Counterfactual	7.05 _a	7.05 _{ab}	3.65
Causal	6.64	6.17 _a	3.55
Control	6.10 _a	5.45 _b	3.65
Mean	6.60	6.22	3.62

Notes: Means within columns sharing a common subscript differ at $p \leq 0.03$. Ratings were on scales from 1 to 9. Higher ratings indicate higher preventability, controllability or predictability.

T-tests revealed that the outcome was judged more preventable by participants who generated counterfactual thoughts than by controls [$M = 7.05$ vs $M = 6.1$, $t(81) = 2.23$, $p = 0.03$], whereas it was judged equally preventable by participants who generated causal thoughts and controls [$M = 6.64$ vs $M = 6.1$, $t(80) = 1.28$, $p = 0.2$]. Participants who generated counterfactual thoughts rated the outcome as more preventable than those who generated causal thoughts, although this difference did not reach significance [$M = 7.05$ vs $M = 6.64$, $t(83) = 1.13$, $p = 0.26$].

The outcome was also judged significantly more controllable by participants who generated counterfactual thoughts than by controls [$M = 7.05$ vs $M = 5.45$, $t(81) = 3.61$, $p <$

0.01], whereas the controllability ratings of participants who generated causal thoughts did not differ from those of controls [$M = 6.17$ vs $M = 5.45$, $t(75) = 1.4$, $p = 0.16$]. Participants who generated counterfactual thoughts rated the outcome as significantly more controllable than those who generated causal thoughts [$M = 7.05$ vs $M = 6.17$, $t(83) = 2.33$, $p = 0.02$].

Counterfactual and causal thinking had no effects on the perceived predictability of the outcome. Participants who generated counterfactual thoughts and controls judged the outcome to be equally predictable [$M = 3.65$ vs $M = 3.65$, $t(75) = 0$, $p = 1$], as did those who generated causal thoughts and controls [$M = 3.55$ vs $M = 3.65$, $t(80) = 0.25$, $p = 0.8$]. There was no difference in the predictability ratings of participants who generated counterfactual and causal thoughts ($M = 3.65$ vs $M = 3.55$, $t(83) = 0.3$, $p = 0.77$).

The results provide the first demonstration that counterfactual thoughts increase perceived preventability, and they support previous findings that counterfactual thoughts increase perceived control (e.g., McMullen et al., 1995). In contrast, causal thoughts had no effects on either perceived preventability or perceived control. These findings are consistent with the view that counterfactual thoughts are more concerned with prevention than causal thoughts. On the other hand, the idea that causal thoughts are more concerned with prediction than counterfactual thoughts is not supported by the finding that neither counterfactual nor causal thoughts affected perceived predictability. However, a possible reason for this finding is that the outcome of the scenario was very unpredictable, as suggested by low predictability ratings in the control condition, averaging only 3.65 on a scale of a 1 to 9 (whereas the preventability and controllability ratings averaged 6.1 and 5.45, respectively). Our next experiment examines the effects of counterfactual and causal thinking on judgements of the predictability of an outcome with a more moderate baseline level of predictability.

The results support the findings of Experiments 1 and 3 that counterfactual and causal thoughts have different contents: counterfactual thoughts focused on internal factors, whereas causal thoughts focused on external factors (see also Davis et al., 1995; Mandel & Lehman, 1996; N'gbala & Branscombe, 1995). Moreover, it provides the first demonstration that there are also differences in the *consequences* of causal and counterfactual thoughts, and this novel finding provides more direct support for the idea that they serve different functions.

Experiment 5

The aim of this experiment was to further examine the consequences of counterfactual and causal thinking for judgements related to prediction by comparing their effects on the hindsight bias, that is the post-hoc exaggeration of an outcome's prior predictability (e.g., Fischhoff, 1975; see Chapter 1, pp. 47-49)⁷. We suggest that if counterfactual thinking is concerned with prevention, whereas causal thinking is concerned with prediction, then counterfactual and causal thoughts should have different consequences for the hindsight bias: counterfactual thoughts should increase the perceived preventability of an outcome, thereby decreasing the hindsight bias, whereas causal thoughts may have no effect on the hindsight bias. In line with this view, several theorists have suggested that counterfactual thinking is incompatible with the hindsight bias (e.g., Kahneman & Varey, 1990, p. 1103; Sherman, 1991, p. 182; Wells & Gavanski, 1989, p. 167). However, Roese and Maniar (1997) report an experiment which suggests that both counterfactual and causal thinking increase the hindsight bias (see Chapter 1, p. 48). Their findings suggest that counterfactual and causal thinking may be

⁷ This experiment was carried out in collaboration with Vittorio Girotto. We designed the experiment and ran several pilot studies during my three month research visit to the University of Provence in 1999. The data were subsequently collected by Vittorio Girotto at the University of Trieste and analysed by me in Dublin.

equally concerned with prediction (pace Mandel & Lehman, 1996).

However, Roese and Maniar's (1997) study only examined the consequences of counterfactual and causal thoughts about good outcomes (participants were students attending football games which were won by their college's team). There is evidence that counterfactual thoughts about good and bad outcomes can have different consequences for emotions (e.g., Byrne & McEleney, 1997; 2000; Gleicher et al., 1990; Landman, 1987; see Chapter 1, pp. 42-43), and so counterfactual thoughts about good and bad outcomes may also have different consequences for other phenomena such as the hindsight bias. The aim of this experiment was to determine the effects of counterfactual and causal thoughts on the hindsight bias for a bad outcome, because people tend to *spontaneously* generate counterfactual and causal thoughts about bad rather than good outcomes, as we showed in Experiment 1 (see also Chapter 1, pp. 17-19).

We gave participants a version of a scenario describing the destruction of a Belgian town during the First World War, which has previously been shown to evoke the hindsight bias (Roese & Olson, 1996, Experiment 2). We included two control groups in order to confirm that the scenario evoked the hindsight bias in our participants. The "outcome control" group read the complete scenario, and the "no-outcome control" group read the scenario with the last two sentences (which described the outcome) omitted. The two experimental groups both read the complete scenario. After reading the scenario, the two experimental groups generated either a causal thought or a counterfactual thought, respectively, and the two control groups wrote a one-line description of the scenario. All participants then responded to two sets of hindsight bias measures.

Method

Participants and Design

The 134 University of Trieste undergraduates, who took part in the experiment voluntarily, included 79 women and 55 men. Their ages ranged from 18 to 41, with a mean age of 21 years. They were randomly assigned to four conditions: counterfactual thinking (n = 35), causal thinking (n = 33), outcome control (n = 33) and no-outcome control (n = 33).

Materials and Procedure

The experiment was run in two group sessions. Participants were presented with a three-page booklet containing the instructions and materials, which were translated into Italian, the participants' first language, by Vittorio Girotto (see Appendix 1e). On the cover page, they were informed that the study was concerned with “how people think about everyday events”, and that there were no right or wrong answers. They were asked to take their time and complete the tasks in the booklet in the order in which they were presented.

On the second page, they were presented with the following scenario (adapted from Roese & Olson, 1996; see also Appendix 1e):

During the first month of the First World War (1914-1918), hundreds of thousands of lives were lost during the Battle of the Frontiers. At this time, the Germans fought the French and the British along the border between Germany and France. In these early weeks, the Germans quickly acquired a reputation for being ruthless and harsh, particularly as they burned and terrorised villages in Belgium and France.

During the battle, a small British force accidentally encountered a much larger German force. The British were forced to retreat into a village, putting all the inhabitants into extreme danger of attack. The village was small and peaceful. It was inhabited mainly by farmers, and among them many children. Unfortunately, the British commander, Lt. Dorian Moore, was slow, inexperienced, and indecisive. Moore ordered his troops to stay in the village, and to simply wait for a German attack. This spelled certain doom for the village.

One British soldier was Thomas Jensen. He was only 20 years old, and although somewhat shy and unsure of himself, he was bright and had already involved himself in several minor military decisions for the British. He saw a way of luring the Germans away from the village, thereby saving the village from destruction. The more he worked on his plan, the more he became convinced that he could save the village, as well as his fellow soldiers. The problem was: since Jensen was so young, it was not certain that Moore would listen to him.

Jensen faced a difficult decision. As he saw it, he had three alternatives: 1) He could simply forget his plan and follow the orders Moore had already given. 2) Since he was so well-liked, he could organise his fellow soldiers to disobey Moore's orders, and to follow his plan instead. 3) He could carefully explain his plan to another officer, and have that officer tell Moore. With luck, Moore might order the troops to follow his plan.

Of these three choices, choice 1 was the least attractive (since it meant certain doom for the village). Choice 2 was riskiest to Jensen himself (he could be tried and shot for mutiny if his plan failed) but it could save the village (assuming that the troops were willing to follow Jensen's plan). Choice 3 was safer for Jensen (his conduct could not be questioned) and it could also save the village (assuming Moore would change his orders and instruct the troops to follow Jensen's plan). After a sleepless night, Jensen decided on choice 3.

The outcome was that Moore rejected Jensen's plan and strictly reaffirmed his original orders to the troops. The village was destroyed.

In the no-outcome control condition, the last two sentences were omitted.

On the next page, participants in the counterfactual thinking condition were asked to complete the following description of the story: "The story is about the misfortune of a village during the First World War. The village might have been saved if..." Participants in the causal thinking and control conditions received the same instructions, except that they were given the sentence stems "The village was destroyed because...", and "The village was...", respectively (see also Appendix 1e).

All participants were then given the following hindsight instructions: "Some people taking part in this study were asked to read the story but were not told the outcome. Please try to

put yourself in their shoes and answer the following questions as if you do not know the outcome” (e.g., Fischhoff, 1975; Roese & Olson, 1996; Wasserman et al., 1991). These instructions were followed by two predictability measures. First, participants were asked to rate their level of agreement with each of the following statements on a scale from 1 (strongly disagree) to 9 (strongly agree): “It was predictable in advance that the village would be *saved*”; “It was predictable in advance that the village would be *destroyed*” (e.g., Roese & Olson, 1996). Second, they were asked to judge the likelihood of each of the following outcomes by writing a probability value of 0 to 100% next to each outcome: “The village is *saved*”; “The village is *destroyed*”. Participants were instructed that the total of their probability estimates should equal 100% (e.g., Fischhoff, 1975; Roese & Olson, 1996; Wasserman et al., 1991; see Appendix 1e).

Results and Discussion

Manipulation check

As in Experiment 4, a manipulation check was carried out on the sentence completions of participants in the two experimental conditions. A completion was coded as a counterfactual thought if it stated how the outcome could have been different by changing some preceding event (e.g., “if Jensen had convinced Moore”). A completion was coded as a causal thought if it stated why the outcome occurred without changing any preceding event (e.g., “because Moore refused Jensen’s plan”). The completions were coded by Vittorio Girotto and by an independent rater who was unaware of the purpose of the experiment. Initial interrater agreement was 94%. The four discrepancies were resolved by discussion. Ninety-nine per cent of participants provided a completion consistent with their condition. The remaining one participant was excluded from further analyses.

Content coding

Participants' sentence completions were also categorised according to whether or not they mentioned *internal* factors related to the central character, Jensen (e.g., "if *Jensen had changed his plan*"), and whether or not they mentioned *external* factors (e.g., "because *the British army stayed in the village*"). The contents were coded by Vittorio Girotto and by an independent rater who was unaware of the purpose of the experiment. Initial interrater agreement was 91%. The six discrepancies were resolved by discussion. 25% of participants in the counterfactual condition mentioned internal factors, binomial $z = 2.57$, $n = 34$, $p < 0.01$, whereas only 3% of participants in the causal condition mentioned internal factors, binomial $z = 5.23$, $n = 33$, $p < 0.01$, and this difference was significant, $\chi^2(1) = 7.26$, $p < 0.01$. 97% of participants in the causal condition mentioned external factors, $z = 5.23$, $n = 33$, $p < 0.01$, whereas only 74% of participants in the counterfactual condition mentioned external factors, binomial $z = 2.57$, $n = 33$, $p < 0.01$, and this difference was significant, $\chi^2(1) = 7.26$, $p < 0.01$.

Hindsight bias

We first assessed whether the scenario evoked the hindsight bias by carrying out within-participant and between-participant comparisons of the predictability ratings and likelihood judgements of outcome control and no-outcome control participants. The mean predictability ratings and likelihood judgements are shown in Table 3.2.

The overall findings provide tentative evidence for a hindsight bias. As expected, the village being destroyed was judged more likely than the village being saved by outcome controls, but not by no-outcome controls [$M = 59.06$ vs $M = 40.94$, $t(31) = 2.49$, $p = 0.02$; $M = 53$ vs $M = 47$, $t(32) = 0.82$, $p = 0.42$]. Also as expected, the village being destroyed was judged

more likely by outcome controls than no-outcome controls, and the village being saved was judged less likely by outcome controls than no-outcome controls, although neither of these differences reached significance [$M = 59.06$ vs $M = 53.00$, $t(63) = 1.18$, $p = 0.24$; $M = 40.94$ vs $M = 47.00$, $t(63) = 1.18$, $p = 0.24$]. Unexpectedly, the village being destroyed was rated as more predictable than the village being saved by both outcome control and no-outcome participants [$M = 6.21$ vs $M = 4.24$, $t(32) = 3.7$, $p < 0.01$; $M = 5.48$ vs $M = 3.88$, $t(32) = 2.86$, $p < 0.01$]. However, as expected, the village being destroyed was considered more predictable by outcome controls than no-outcome controls, although this effect did not reach significance [$M = 6.21$ vs $M = 5.48$, $t(64) = 1.51$, $p = 0.13$]. The village being saved was judged equally predictable by outcome controls and no-outcome controls [$M = 4.24$ vs $M = 3.88$, $t(64) = 0.74$, $p = 0.46$].

Table 3.2: Mean predictability ratings and likelihood judgements following different sentence completion tasks in Experiment 5

<i>Sentence completion task</i>	<i>Predictability</i>		<i>Likelihood</i>	
	Village destroyed	Village saved	Village destroyed	Village saved
Counterfactual	4.79 _{AB}	4.47	52.94 _a	47.06 _a
Causal	5.97 _A	4.06	61.56 _a	38.75 _a
Outcome control	6.21 _B	4.24	59.06	40.94
No-outcome control	5.48	3.88	53.00	47.00
<i>Mean</i>	<i>5.61</i>	<i>4.15</i>	<i>56.83</i>	<i>43.24</i>

Notes: Means within columns sharing a common subscript differ at $p < 0.01$ for upper case letters and $p \leq 0.09$ for lower case letters. Predictability ratings were on scales from 1 to 9. Higher ratings indicate higher predictability. Likelihood judgements were probabilities from 0 to 100%.

Effects of counterfactual and causal thinking

We next assessed the effects of counterfactual and causal thinking on the hindsight bias by comparing the predictability ratings and likelihood judgements of counterfactual thinking, causal thinking and outcome control participants.

As expected, counterfactual thinking tended to eliminate the hindsight bias. Participants who generated counterfactual thoughts rated the village being destroyed as just as predictable as the village being saved [$M = 4.79$ vs $M = 4.47$, $t(33) = 0.56$, $p = 0.58$], whereas outcome controls rated the village being destroyed as more predictable than the village being saved (see above). The village being destroyed was rated as less predictable by participants who generated counterfactual thoughts than outcome controls [$M = 4.79$ vs $M = 6.21$, $t(65) = 2.9$, $p < 0.01$], and the village being saved was rated as equally predictable by participants who generated counterfactual thoughts and outcome controls [$M = 4.47$ vs $M = 4.24$, $t(65) = 0.45$, $p = 0.65$]. Participants who generated counterfactual thoughts judged the village being destroyed equally likely as the village being saved [$M = 52.94$ vs $M = 47.06$, $t(33) = 0.86$, $p = 0.4$], whereas outcome controls judged the village being destroyed more likely than the village being saved (see above). The village being destroyed was judged somewhat less likely by participants who generated counterfactual thoughts than outcome controls, although this effect did not reach significance [$M = 52.94$ vs $M = 59.06$, $t(63) = 1.23$, $p = 0.22$]. The village being saved was judged somewhat more likely by participants who generated counterfactual thoughts than by outcome controls, although this effect also did not reach significance [47.06 vs $M = 40.94$, $t(63) = 1.23$, $p = 0.22$].

In contrast, causal thinking had no effect on the hindsight bias. Participants who generated causal thoughts rated the village being destroyed as more predictable than the village

being saved [$M = 5.97$ vs $M = 4.06$, $t(32) = 3.06$, $p < 0.01$], as did outcome controls (see above).

The village being destroyed was rated as equally predictable by participants who generated causal thoughts and outcome controls [$M = 5.97$ vs $M = 6.21$, $t(63) = 0.51$, $p = 0.61$], and the village being saved was rated as equally predictable by participants who generated causal thoughts and outcome controls [$M = 4.06$ vs $M = 4.24$, $t(64) = 0.35$, $p = 0.72$]. Participants who generated causal thoughts judged the village being destroyed more likely than the village being saved [$M = 61.56$ vs $M = 38.75$, $t(30) = 3.22$, $p < 0.01$], as did outcome controls (see above).

The village being destroyed was judged equally likely by participants who generated causal thoughts and outcome controls [$M = 61.56$ vs $M = 59.06$, $t(62) = 0.5$, $p = 0.62$], and the village being saved was judged equally likely by participants who generated causal thoughts and outcome controls [$M = 38.75$ vs $M = 40.94$, $t(62) = 0.44$, $p = 0.66$].

The village being destroyed was rated as less predictable by participants who generated counterfactual thoughts than by those who generated causal thoughts [$M = 4.79$ vs $M = 5.97$, $t(65) = 2.53$, $p = 0.01$], although they gave equal predictability ratings for the village being saved [$M = 4.47$ vs $M = 4.06$, $t(65) = 0.79$, $p = 0.43$]. The village being destroyed was judged somewhat less likely by participants who generated counterfactual thoughts than by those who generated causal thoughts, although this difference was marginally significant [$M = 52.94$ vs $M = 61.56$, $t(64) = 1.78$, $p = 0.08$]. The village being saved was judged somewhat more likely by participants who generated counterfactual thoughts than by those who generated causal thoughts, although this difference was also marginal [$M = 47.06$ vs $M = 38.75$, $t(64) = 1.72$, $p = 0.09$].

The results suggest that counterfactual thinking decreases the hindsight bias, whereas causal thinking has no effect on the hindsight bias. We used a scenario previously shown by Roese and Olson (1996) to evoke the hindsight bias and which also tended to evoke a hindsight

bias in our participants. Participants who generated counterfactual thoughts rated the outcome as significantly less predictable than controls, whereas those who generated causal thoughts showed as strong a hindsight bias as controls. Moreover, those who generated counterfactual thoughts rated the outcome as significantly less predictable, and somewhat less likely, than those who generated causal thoughts. The findings support the view that counterfactual thoughts are concerned with prevention, whereas causal thoughts are concerned with prediction. The hindsight bias seems to be a side effect of learning how an outcome can be predicted (e.g., Hawkins & Hastie, 1990; Wasserman et al., 1991), but if counterfactual thoughts show how an outcome can be prevented, they may decrease its perceived predictability, and therefore eliminate the hindsight bias.

Our results seem to contradict Roese and Maniar's (1997) finding that counterfactual and causal thinking both increase the hindsight bias. One possible reason for the difference in our results is that Roese and Maniar examined the effects of counterfactual and causal thinking on the hindsight bias for good outcomes, whereas we chose to focus on bad outcomes. A possible alternative explanation is that Roese and Maniar examined the effects of counterfactual and causal thinking on the perceived predictability of a real outcome, whereas our experiment was scenario-based. Our next experiment compared the consequences of counterfactual and causal thoughts for the perceived predictability of negative outcomes which participants had actually experienced.

Experiment 6

This experiment compared the consequences of counterfactual and causal thoughts about a real life negative outcome. One aim of the experiment was to further examine the idea that

counterfactual thoughts are directed towards prevention, by comparing the consequences of counterfactual and causal thoughts for a range of judgements related to prevention. We re-examined their consequences for perceived control, and we also compared their consequences for judgements of responsibility and blame. Moreover, we compared their consequences for the perceived preventability and controllability of a hypothetical future outcome and intentions for future preventative action. Previous research suggests that counterfactual thoughts may increase the amount of responsibility and blame assigned to an individual (e.g., Branscombe et al., 1996; see Chapter 1, p. 45) and that causal thoughts may influence related legal judgements (e.g., Carroll et al., 1987; see Chapter 1, pp. 44-45). However, no previous studies have compared the effects of counterfactual and causal thoughts on judgements of responsibility and blame. To our knowledge, no previous research has examined the consequences of either counterfactual or causal thoughts for judgements of the preventability or controllability of *future* outcomes. Past studies suggest that both counterfactual and causal thoughts influence people's intentions for future action (e.g., Eiser et al., 1985; Roese, 1994; see Chapter 1, pp. 49-51), but no previous research has compared the effects of counterfactual and causal thoughts on behavioural intentions.

A second aim of the experiment was to further examine the idea that causal thoughts are directed towards prediction by comparing the consequences of counterfactual and causal thoughts for the perceived likelihood of a hypothetical *future* outcome, as well as the perceived predictability of a past outcome. Past research suggests that causal thoughts may influence people's expectancies for future outcomes (e.g., Eiser et al., 1995; see Chapter 1, p. 49), but no previous studies have compared the effects of causal and counterfactual thinking for predictions about future outcomes.

A third aim was to investigate the idea that counterfactual and causal thoughts have different emotional functions. Previous research suggests that *upward* counterfactuals about how an outcome could have been better evoke negative emotions and *downward* counterfactuals about how an outcome could have been worse evoke positive emotions (e.g., Roese, 1994; see Chapter 1, pp. 42-43), whereas the emotional consequences of causal thoughts depend on their content and the type of outcome they explain (e.g., Weiner et al., 1978; see Chapter 1, pp. 41-42). No previous studies have compared the consequences of counterfactual and causal thoughts for emotions.

In this experiment, we examined the consequences of downward counterfactuals as well as upward counterfactuals and causal thoughts. Moreover, we directly examined the role of content in the consequences of counterfactual and causal thoughts by directing participants in the *internal* conditions to focus their thoughts on their own actions, and allowing participants in the *general* conditions to focus on whatever events they wished.

Method

Participants and design

The 213 undergraduates from a variety of departments at the University of Dublin, who took part voluntarily, included 146 women, 66 men and one person who did not record their gender. They were aged from 17 to 57, with a mean age of 19 (two participants did not record their age).

The design of the experiment was a 3 (thought type: upward counterfactual vs downward counterfactual vs causal) x 2 (thought focus: general vs internal) between-participants factorial design. Participants were randomly assigned to one of the seven conditions: general upward

counterfactual (n = 33), internal upward counterfactual (n = 29), general downward counterfactual (n = 33), internal downward counterfactual (n = 27), general causal (n = 32), internal causal (n = 28), and control (n = 31).

Materials and procedure

The experiment was run in two group sessions. Participants were presented with a five-page booklet containing the instructions and materials (see Appendix 1f). On the cover page, they were informed that the study was concerned with “how people react to negative life events”. They were asked to work through the tasks in the booklet in the order in which they were presented and to answer the questions as honestly and accurately as possible. On the second page, all participants received the following instructions (adapted from Markman & Weary, 1996):

Please take a moment to recall an event that has happened to you that had a negative outcome. The outcome of the event should have the following characteristics:

- It made you unhappy or upset you in some way.
- It directly involved you (not something that happened to someone you know).
- It happened recently (within the last year or so).
- A similar outcome could possibly happen to you again in the future.

Some examples of the types of event you might recall are doing badly in an exam or competition, the end of a relationship, a serious fight with a friend or relative, being involved in an accident or crime, a financial or material loss, etc.

They were then asked to write a brief description of the event and its outcome and to rate how upsetting the event was at the time, from 1 (not at all upsetting), to 7 (extremely upsetting). The next page contained the manipulations, which involved an open-ended thought listing task, rather

than the sentence completions used in Experiments 4 and 5. Participants in the general upward counterfactual thinking condition were given the following instructions (italics indicate parts that were different in the other conditions):

Following negative life events, people often think about *how the outcome could have turned out better*. Please think carefully about the event you have described, and in the space below, write down as many *ways that the outcome could have turned out better* as come to mind. Please take your time, write down everything that occurs to you, and feel free to elaborate on your opinions.

They could then use the rest of the page to write their response. The instructions given to the participants in the other conditions are shown in Appendix 1f.

The final two pages of the booklet contained the 14 dependent measures, which were presented in three sets. The emotion measures were presented first. Participants were asked to rate “your current feelings about the event” on scales from 1 to 7. The six bipolar ratings were regret-satisfaction, unhappiness-happiness, disappointment-relief, depression-elation, guilt/shame-pride and pessimism-optimism. The measures of past controllability, responsibility, blame and predictability were presented second. Participants were asked to rate “your current thoughts about the event” on scales from 1 to 7. They rated how much control they had over the outcome, how responsible they were for the outcome, how much they were to blame for the outcome, and how predictable the outcome was in advance. The measures of future likelihood, controllability, preventability and intentions were presented last, under the heading “your current thoughts about the future”. Participants were instructed to “imagine that in the future, you are in a situation where there is a *possibility* of the same *type* of negative outcome occurring again”. They then rated from 1 to 7 the likelihood that things would turn out better this time, how much control they would have over the outcome, the likelihood that they could prevent the same type

of negative outcome from occurring again and the likelihood that they would act differently from the way they acted last time (see also Appendix 1f).

Results and Discussion

Events recalled

The events participants recalled included interpersonal conflicts, e.g., an argument, infidelity, or relationship break-up (50%), achievement failures, e.g., a poor performance in an examination, competition or interview (21%), health problems, e.g., their own illness or injury, or the illness, injury or death of a loved one (15%), and a range of other events (15%). There were no differences between the conditions in the type of events participants recalled [$F(6, 206) = 0.77, p = 0.6$]. The mean rating for how upsetting the event was at the time was 5.2, on a scale from 1 (not at all upsetting) to 7 (extremely upsetting). There were no differences between the conditions in how upsetting the events were rated [$F(6, 206) = 0.8, p = 0.57$].

Manipulation check and content coding

We carried out a manipulation check on the thought listings of participants in the experimental conditions. An upward counterfactual was defined as a statement which suggests that the outcome could have been better by changing some preceding event (e.g., “*I could have studied a lot harder*”). A downward counterfactual was defined as a statement which suggests that the outcome could have been worse by changing some preceding event (e.g., “*Leaving home would have worsened the situation*”). A causal thought was defined as a statement which explains why the outcome occurred without changing any preceding event (e.g., “*I put my work before the relationship and so he may have felt neglected*”). Statements in these categories were further classified as *internal* (concerning the individual’s own actions) or *external* (concerning

other people's actions or situational factors). A manipulation was considered effective if a participant recorded more thoughts corresponding to their condition (e.g., internal *or* external causal thoughts for the general causal condition, etc.) than non-correspondent thoughts. The contents were coded by the author and by an independent rater who was unaware of the purpose of the experiment. Initial interrater agreement was 90%. Discrepancies were resolved by discussion. The manipulations were effective for 89% of participants. The remaining 24 participants were eliminated from all subsequent analyses.

Participants in the general conditions could generate either internal or external thoughts. In fact, those in the general upward counterfactual condition were as likely to produce internal as external upward counterfactuals (45% produced more internal, 42% produced more external, binomial $z = 0.19$, $n = 27$, $p = 0.85$). In contrast, those in the general downward counterfactual condition tended to produce more external than internal downward counterfactuals (although this effect was marginally significant, 31% produced more internal, 66% produced more external, binomial $z = 1.89$, $n = 28$, $p = 0.06$). Participants in the general causal condition tended to produce more internal than external causal thoughts (although this difference was also marginal, 56% produced more internal, 28% produced more external, binomial $z = 1.73$, $n = 27$, $p = 0.08$). There were significant differences between the three conditions in participants' tendencies to focus on internal or external factors, general upward vs general downward: $\chi^2(1) = 12.54$, $p < 0.01$; general upward vs general causal: $\chi^2(1) = 14.54$, $p < 0.01$; general downward vs general causal: $\chi^2(1) = 6.75$, $p < 0.01$.

Preventability, predictability, future likelihood and emotion

A principal components analysis carried out on the 14 dependent measures extracted four

factors. The measures of perceived predictability and the likelihood of a better outcome in the future had high loadings on distinct factors, and so these measures remained as separate scores, which we will call “predictability” and “future likelihood”. The three other “thoughts about the past” measures (past controllability, self-responsibility and self-blame) and the three other “thoughts about the future” measures (future preventability, controllability and intentions) all had their highest loadings on another factor, and so these six measures were averaged to form a single score, which we will call “preventability” ($\alpha = 0.85$). Finally, the six emotion measures (regret-satisfaction, unhappiness-happiness, disappointment-relief, depression-elation, guilt/shame-pride and pessimism-optimism) all had their highest loadings on another factor, and so these six measures were averaged to form a single “emotion” score ($\alpha = 0.69$).

Table 3.3 shows participants’ mean scores on the measures of preventability, predictability, future likelihood, and emotion. To assess the effects of the manipulations, we carried out four ANOVAs. Each one was a three (thought type: upward counterfactual vs downward counterfactual vs causal) by two (thought focus: general vs internal) ANOVA. First, the analysis of the *preventability* scores showed no effect of either thought type [$F(2, 152) = 1.56, p = 0.22$] or thought focus [$F(1, 152) = 1.86, p = 0.17$], but there was a significant interaction between thought type and thought focus [$F(2, 152) = 3.31, p = 0.04$]. T-tests showed that upward counterfactual thoughts and causal thoughts both increased perceived preventability in the internal conditions, but not in the general conditions. The scores of participants in the internal upward counterfactual condition ($M = 4.68$) and the internal causal condition ($M = 4.71$) were higher than those of participants in the control condition ($M = 3.74$), and these differences were reliable [$t(44) = 2.32, p = 0.02$; $t(46) = 2.41, p = 0.02$, respectively]. They were also both higher than the scores of participants in the internal downward counterfactual condition ($M =$

3.65) and these differences were reliable [$t(38) = 2.65, p = 0.01$; $t(40) = 2.78, p < 0.01$, respectively]. There were no other systematic differences between the groups (see Appendix 3).

Second, the analysis of the *predictability* scores revealed no significant effect of either thought type [$F(2, 152) = 2.03, p = 0.13$] or thought focus [$F(1, 152) = 0.24, p = 0.62$], and no interaction between thought type and thought focus [$F(2, 152) = 0.03, p = 0.97$]. However, t -tests showed that participants in the causal conditions rated the outcome as somewhat more predictable ($M = 4.33$) than those in the upward counterfactual conditions ($M = 3.65$) or the downward counterfactual conditions ($M = 3.61$), although these differences were both marginal [$t(102) = 1.81, p = 0.07$; $t(105) = 1.89, p = 0.06$, respectively]. There were no other differences between the groups (see Appendix 3).

Third, the *future likelihood* scores showed no effect of either thought type [$F(2, 152) = 2.03, p = 0.13$] or thought focus [$F(1, 152) = 0.24, p = 0.62$], but there was a reliable interaction between thought type and thought focus [$F(2, 152) = 3.91, p = 0.02$]. T -tests showed that internal causal thoughts somewhat increased the perceived likelihood of a better future outcome, whereas upward and downward counterfactuals had no effects. The future likelihood scores of participants in the internal causal condition ($M = 5.43$) were higher than those of controls ($M = 4.65$), although this difference was marginal [$t(50) = 1.73, p = 0.07$], whereas none of the other conditions differed from the control group (see Appendix 3).

Finally, the *emotion* scores showed a significant main effect of thought type [$F(2, 152) = 6.57, p < 0.01$], but no effect of thought focus [$F(1, 152) = 0.18, p = 0.67$], and no interaction between thought type and thought focus [$F(2, 152) = 0.48, p = 0.62$]. T -tests revealed that the main effect of thought type was due to the affect-enhancing effect of downward counterfactuals. Participants in the downward counterfactual conditions felt better ($M = 3.87$) than participants in

the control condition [$M = 3.12$, $t(75) = 4.00$, $p < 0.01$], the upward counterfactual conditions [$M = 3.38$, $t(102) = 2.72$, $p < 0.01$] and the causal conditions [$M = 3.33$, $t(99) = 3.32$, $p < 0.01$; see Appendix 3 for further details].

Table 3.3: Mean preventability, predictability, future likelihood and emotion scores following different thought listing tasks in Experiment 6

Thought listing task		Score			
Thought type	Thought focus	Preventability	Predictability	Future likelihood	Emotion
Upward Counterfactual	General	3.85 _{ab}	3.71	4.23 _A	3.44 _A
	Internal	4.68 _{aCD}	3.60	4.65	3.32 _B
	Mean	4.26	3.65 _a	4.44	3.38 _C
Downward counterfactual	General	4.17	3.66	4.93	3.77 _{DE}
	Internal	3.65 _{CE}	3.56	4.40 _B	3.97 _{ABFGH}
	Mean	3.91	3.61 _b	4.66	3.87 _{CIJ}
Causal	General	4.06	4.47	4.16 _C	3.28 _{EH}
	Internal	4.71 _{bEF}	4.19	5.43 _{ABCD}	3.38 _G
	Mean	4.38	4.33 _{ab}	4.79	3.33 _I
Control		3.74 _{DF}	3.90	4.65 _d	3.12 _{DFI}
Mean		4.12	3.87	4.64	3.45

Notes: Means within columns sharing a common subscript differ at $p \leq 0.03$ for upper case letters and $p \leq 0.07$ for lower case letters. Scores could range from 1 to 7. Higher scores indicate higher predictability, likelihood or preventability, or more positive emotions. “Preventability” scores are the average of ratings of past controllability, self-responsibility and self-blame, and future preventability, controllability and intentions. “Emotion” scores are the average of ratings of regret-satisfaction, unhappiness-happiness, disappointment-relief, depression-elation, guilt/shame-pride and pessimism-optimism.

The experiment shows that perceived control over a past outcome is highly related to perceptions of self-responsibility and self-blame, judgements of the preventability and controllability of future outcomes, and intentions for future preventative action, and so we combined these measures into single “preventability” scores. Internal upward counterfactuals increased these preventability scores. This result supports the finding of Experiment 4 that counterfactual thinking increases the perceived preventability and controllability of an outcome. It also supports past findings that upward counterfactual thinking influences judgements of responsibility and blame (e.g., Branscombe et al., 1996; Wells & Gavanski, 1989) and intentions for future action (e.g., Gleicher et al., 1995; Roese, 1994). Moreover, it provides the first demonstration that upward counterfactuals increase the perceived preventability and controllability of *future* outcomes. Internal causal thoughts also increased preventability scores. This finding suggests that causal thinking may have had no effects on perceived preventability and controllability in Experiment 4 because in that experiment, participants’ causal thoughts tended to focus on external rather than internal factors. In contrast, internal downward counterfactuals had no effects on preventability scores, which is consistent with McMullen et al.’s (1995) finding that upward, but not downward counterfactuals increase perceived control.

Upward counterfactual thinking did not increase judgements related to prevention in the general conditions, in which participants could focus their upward counterfactuals on whatever events they wished. This finding is inconsistent with the results of Experiment 4. A possible reason is that in this experiment, participants’ general upward counterfactuals focused equally on their own actions and on external factors, whereas participants in Experiment 4 focused their counterfactual thoughts overwhelmingly on personal actions, as in Experiments 1 and 3, consistent with previous studies of the content of counterfactual thoughts (e.g., Giroto et al.,

1991; Mandel & Lehman, 1991; N'gbala & Branscombe, 1995). One possible explanation for this difference is that many of the participants in this experiment could not easily generate counterfactuals focusing on their own actions because they recalled events over which they had no control (e.g., a relative's death from old age).

Causal thoughts resulted in somewhat greater perceived predictability than counterfactual thoughts, which supports the view that causal thoughts are concerned with the prediction of outcomes, whereas counterfactual thoughts are concerned with their prevention, as in Experiment 5. The results provide further support for our speculation that the lack of difference in the predictability ratings of participants who generated counterfactual and causal thoughts in Experiment 4 may have been due to the low baseline level of predictability. Moreover, this experiment provides further evidence for the link between causal thinking and prediction by showing that internal causal thoughts increase the perceived likelihood of a better future outcome, whereas upward and downward counterfactuals have no effects on future likelihood judgements. The findings suggest that the belief that an outcome was *caused* by one's actions may have a greater effect on predictions about future outcomes than the suggestion that the outcome might have been better or worse if one had acted differently.

The results suggest that counterfactual and causal thinking differ not only in their consequences for judgements related to prevention and prediction, but also in their consequences for emotions. Participants who generated downward counterfactuals felt significantly better than controls, whereas those who generated upward counterfactuals or causal thoughts felt no better or worse than controls. The results support previous findings that upward and downward counterfactuals have different consequences for emotions (e.g., Markman et al., 1993; Roese, 1994; McMullen et al., 1995), and they provide the first demonstration of differences in the

emotional consequences of counterfactual and causal thoughts. The results are consistent with the idea that downward counterfactuals serve an emotional function.

General discussion

The three experiments reported in this chapter provide the first demonstration of differences in the consequences of counterfactual and causal thinking. The results are consistent with the view that counterfactual and causal thinking serve different functions: upward counterfactuals are concerned with prevention, causal thoughts are concerned with prediction and downward counterfactuals are concerned with emotions.

The experiments provide the first comparison of the consequences of counterfactual and causal thoughts for judgements about prevention. In Experiment 4, counterfactual thoughts increased the perceived preventability of an outcome, whereas causal thoughts had no effect. Likewise, counterfactual thoughts increased the perceived controllability of the outcome, whereas causal thoughts had no effect. Moreover, Experiment 6 showed that judgements of the controllability of an outcome are closely related to judgements of responsibility and blame, the preventability and controllability of future outcomes, and intentions for future preventative action. The differences in the consequences of counterfactual and causal thoughts for judgements about prevention in Experiment 4 may have depended on the fact that participants' counterfactual thoughts tended to focus on personal actions, whereas their causal thoughts tended to focus on external events, as in Experiments 1 and 3 (see also Davis et al., 1995; Mandel & Lehman 1996; N'gbala & Branscombe, 1995). In Experiment 6, both counterfactual and causal thoughts increased preventability scores when participants were directed to focus on personal actions, but not when they could focus on whatever events they wished. The consequences of

counterfactual thoughts for judgements about prevention also depend on their direction: in Experiment 6, internal upward counterfactuals increased preventability scores, whereas internal downward counterfactuals had no effect. Overall, the findings suggest that upward counterfactuals tend to increase judgements related to prevention compared to either downward counterfactuals or causal thoughts.

The experiments also provide the first comparison of the effects of counterfactual and causal thinking on judgements about the prediction of a bad outcome. In Experiment 4, neither counterfactual nor causal thinking affected perceived predictability, but a possible reason is that the outcome of the scenario was genuinely unpredictable, as suggested by the low mean predictability rating in the control condition. This speculation is supported by the findings of Experiment 5, in which there were more moderate baseline ratings of predictability. The results show that counterfactual thoughts decreased predictability ratings, whereas causal thoughts had no effect on perceived predictability. In Experiment 6, neither counterfactual nor causal thoughts had a significant effect on perceived predictability compared to the control group, but as in Experiment 5, outcomes were rated somewhat more predictable by participants who generated causal thoughts than by those who generated counterfactual thoughts. Moreover, internal causal thoughts increased the perceived likelihood of a better future outcome compared to internal upward or downward counterfactuals. Overall, the findings suggest that causal thoughts result in greater perceived predictability than counterfactual thoughts.

The experiments also provide the first comparison of the influence of counterfactual and causal thoughts on emotions. In Experiment 6, downward counterfactuals enhanced emotions, whereas upward counterfactuals and causal thoughts had no effects on emotions. The findings

suggest that downward counterfactuals enhance emotions compared to either upward counterfactuals or causal thoughts.

Although some of the reported differences were only marginally significant and there were some inconsistencies in the results of the different experiments, perhaps as a result of differences in the experimental materials, the overall findings suggest that counterfactual and causal thinking tend to have different consequences. However, the results suggest that counterfactual and causal thoughts may tend to have different consequences at least in part because they tend to have different contents. In Experiments 4 and 5, counterfactual and causal thoughts had different consequences for prevention and prediction, but in both experiments, participants' counterfactual thoughts were more likely to focus on personal actions and less likely to focus on external factors than their causal thoughts. In Experiment 6, counterfactual and causal thoughts had the same consequences for prevention when participants in both conditions were directed to focus on personal actions. It remains unclear whether there are reliable differences in causal and counterfactual thoughts which have the same contents. The experiments we report in the next chapter examine this issue by comparing how people understand and reason from counterfactual and causal assertions with the same contents.

Overall, the results do not support the view that the principle function of counterfactual thinking is to determine the necessity of possible causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). On this account, counterfactual and causal thoughts might have different consequences for judgements about prevention and prediction due to differences in their contents. Causal thoughts may affect judgements about both prediction and prevention because they focus on antecedents which were both necessary and sufficient for an outcome (the presence of a necessary and sufficient antecedent predicts an

outcome, and its absence prevents the outcome). In contrast, counterfactual thoughts may affect judgements about prevention more than judgements about prediction because they may focus on antecedents which were necessary but insufficient for an outcome (the presence of a necessary but insufficient does not predict an outcome, but its absence prevents the outcome). However, in our experiments, counterfactual and causal thoughts had different consequences even though both tended to focus on antecedents which were necessary but insufficient for the outcome. For example, in Experiment 4, participants' counterfactual thoughts tended to focus on their decision to be assessed by Dr. Jones, whereas their causal thoughts tended to focus on Dr. Jones's harsh assessment. Although these two antecedents were equally necessary for the outcome, and neither would have been sufficient without the other, counterfactual thoughts increased the perceived preventability of the outcome, whereas causal thoughts had no effect. Our findings suggest that differences in the consequences of counterfactual and causal thoughts are not merely a result of possible differences in their tendencies to focus on necessary and sufficient antecedents.

Our overall findings suggest that counterfactual and causal thoughts tend to have different consequences for judgements related to prevention and prediction, and for emotions. The specific pattern of differences is consistent with the view that causal and counterfactual thinking serve different functions: causal thoughts enable prediction, upward counterfactual thoughts enable prevention, and downward counterfactual thoughts enhance emotions.

Chapter 4 Understanding and Reasoning from Causal and Counterfactual Assertions

The experiments reported in Chapters 2 and 3 showed that there are differences in the activation, content and consequences of causal and counterfactual thoughts, which are consistent with the view that causal thoughts are directed towards prediction, whereas counterfactual thoughts are directed towards prevention. However, causal and counterfactual thoughts may have different consequences merely because they tend to have different contents. One aim of the final series of experiments which we will report in this chapter was to compare causal and counterfactual thinking while controlling for their contents. Therefore, unlike the previous six experiments, which examined the determinants of self-generated causal and counterfactual thoughts, these experiments examined how people think about causal and counterfactual assertions which they are presented with as stimuli, such as “Exercise that was excessive caused the occurrence of angina” and “If exercise had not been excessive, then angina would not have occurred”.

A second aim of the experiments was to examine directly the links between causal and counterfactual thinking and thinking about prediction and prevention. Therefore, we examined not only how people think about causal and counterfactual assertions, but also how they think about prediction assertions, which we phrased as indicative conditionals such as “If exercise was excessive, then angina occurred” and prevention assertions such as “Exercise that was not excessive prevented the occurrence of angina”.

The experiments examined the hypothesis that causal thoughts are concerned with prediction and counterfactual thoughts are concerned with prevention, by comparing how people understand and reason from causal, counterfactual, prediction and prevention assertions. No

previous studies have compared how people understand or reason from these four sorts of assertions. People may understand an assertion by keeping in mind the possibilities with which it is consistent, and they may construct models in their minds to represent these possibilities (e.g., Johnson-Laird & Byrne, 1991; see Chapter 1, pp. 52-53). Because of the constraints of their limited working memories, they may construct an initial set of mental models that is less explicit than the fully fleshed out set of possibilities. They may make inferences from two or more assertions by combining their mental models of the assertions and generating a parsimonious conclusion (e.g., Johnson-Laird & Byrne, 1991; see Chapter 1, pp. 57-61). We suggest that if causal thoughts are concerned with prediction and counterfactual thoughts are concerned with prevention, then people may understand causal and counterfactual assertions with the same contents to be consistent with different sets of possibilities. Even if people understand causal and counterfactual assertions to be consistent with the same set of possibilities, they may construct a different initial set of mental models which they use to make inferences from them. In contrast, they should understand and reason from causal and prediction assertions in the same way if the function of causal thinking is prediction, and they should understand and reason from counterfactual and prevention assertions in the same way if the function of counterfactual thinking is prevention.

In this chapter, we will report three experiments which compared how people understand and reason from causal, counterfactual, prediction and prevention assertions. In Experiment 7, we examined the situations people understand to be consistent with the four sorts of assertions. Participants read assertions in the past tense and judged whether each of four situations had been possible or impossible for each assertion. In Experiments 8 and 9, we examined how people initially represent the four sorts of assertions, by comparing the kinds of inferences they make

from them. In Experiment 8, we compared the relative frequencies of modus ponens and denial of the antecedent inferences that participants made from the four sorts of assertions. In Experiment 9, we compared participants' relative latencies to endorse modus ponens and denial of the antecedent inferences from the four sorts of assertions.

Experiment 7

This experiment provides the first comparison of how people understand causal, counterfactual, prediction and prevention assertions. Any assertion which combines two events or states of affairs, such as exercise that is excessive and angina, may refer to four possibilities, such as:

excessive	angina
excessive	¬angina
¬excessive	angina
¬excessive	¬angina

or to any subset of these four possibilities (where “excessive” represents “Exercise that is excessive occurs” and “angina” represents “angina occurs” and “¬” represents negation; see Chapter 1, p. 52). The aim of this experiment was to determine which possibilities people believe each of the four assertions is consistent with and which possibilities they believe it is inconsistent with. For example, given the factual situation in which exercise that was excessive occurred and angina occurred, people might generate causal assertions such as “Exercise that was excessive caused the occurrence of angina” and counterfactual assertions such as “If exercise had not been excessive, then angina would not have occurred”. We suggest that if causal thinking is concerned with prediction, whereas counterfactual thinking is concerned with prevention, then people may understand the causal and counterfactual assertions to be consistent with different sets of

possibilities. In contrast, they should understand the causal assertion to be consistent with the same possibilities as a related prediction assertion such as “If exercise was excessive, then angina occurred”, whereas they should understand the counterfactual assertion to be consistent with the same possibilities as a related prevention assertion such as “Exercise that was not excessive prevented the occurrence of angina”.

Are causal and counterfactual assertions biconditionals or conditionals?

What might any of these four assertions be taken to mean? Logicians identify two meanings of indicative conditionals such as our prediction assertion “If exercise was excessive, then angina occurred”. On the material equivalence or *biconditional* meaning of “if” as “if and only if”, the assertion is true in the first and last situations shown above: the situation in which exercise that is excessive occurs and angina occurs, and the situation in which exercise that is excessive does not occur and angina does not occur. On this interpretation, the antecedent, exercise that is excessive, is necessary and sufficient for the consequent, angina. In contrast, on the material implication or *conditional* interpretation of “if”, the assertion is true in these two situations and also in the third situation shown above, in which exercise that is excessive does not occur and angina occurs. On this interpretation, the antecedent is sufficient but not necessary for the consequent. These two sorts of interpretations of “if” have been extensively examined for indicative conditionals (see Evans et al., 1993, for a review). A third interpretation of “if” which has received recent investigation is the enabling or *reversed conditional* interpretation in which the assertion is true in the two situations outlined for the biconditional, and it is also true in the second situation shown above, in which exercise that is excessive occurs and angina does not occur. In this case the antecedent is necessary but not sufficient for the consequent (e.g., Byrne,

1989; Byrne, Espino & Santamaria, 1999; Chan & Chua, 1994; Stevenson & Over, 1995). In fact, of the 16 interpretations that result from the four possibilities above and all of their subsets, ten of the interpretations appear to be consistent with the everyday usage of “if” (Johnson-Laird & Byrne, 2000). In assertions with realistic contents, the meaning of “if p then q” can be modulated by pragmatics and semantics, resulting in this wide variety of interpretations, shown in Table 4.1.

Table 4.1 Ten interpretations of the conditional “if p then q” (adapted from Johnson-Laird & Byrne, 2000)

Tautology		Biconditional		Conditional		Reversed conditional		Disabling	
p	q	p	q	p	q	p	q	p	q
q	¬q	¬p	¬q	¬p	q	p	¬q	p	¬q
¬p	q			¬p	¬q	¬p	¬q	¬p	q
¬p	¬q								
Strengthened antecedent		Relevance		Ponens		Tollens		DA & AC	
p	q	p	q	p	q	¬p	¬q	¬p	q
p	¬q	¬p	q						

Notes:
 The table shows the possibilities consistent with the assertion under each of the interpretations.
 DA & AC stands for ‘Denial of the antecedent and affirmation of the consequent’.

Which of these interpretations is the most common basic interpretation of causal, counterfactual, prediction and prevention assertions? One possibility is that all four assertions are interpreted to refer to the same possibilities. If people make biconditional interpretations of all

four assertions, then the causal assertion “Exercise that was excessive caused the occurrence of angina”, the counterfactual assertion “If exercise had not been excessive, then angina would not have occurred”, the prediction assertion “If exercise was excessive, then angina occurred” and the prevention assertion “Exercise that was not excessive prevented the occurrence of angina” should all be understood as consistent with the following two possibilities:

excessive	angina
¬excessive	¬angina

On such a biconditional interpretation, there is no situation which distinguishes causal and counterfactual assertions about the same factual situation. The biconditional interpretation conveys information useful for prediction: the presence of excessive exercise predicts angina, because there is only one possible situation in which excessive exercise occurs, and in that situation, angina occurs. It also conveys information useful for prevention: the absence of excessive exercise prevents angina, because there is only one possible situation in which excessive exercise does not occur, and in that situation, angina does not occur. If causal and counterfactual thinking do not differ in their concern with predictive and preventative causal relations, then all four assertions, causal, counterfactual, prediction and prevention, should be interpreted as consistent with just these two biconditional possibilities.

The alternative possibility is that causal and counterfactual assertions are interpreted to be consistent with different sets of possibilities. On any interpretation other than the biconditional interpretation⁸, there is at least one situation which distinguishes causal and counterfactual assertions about the same factual situation. For example, the conditional interpretation of the

⁸ We exclude the tautological interpretation, which is consistent with all four possibilities, and therefore conveys no information (Johnson-Laird & Byrne, 2000).

causal assertion above, which is of the form “p caused q”, is captured by the following three possibilities:

excessive	angina
¬excessive	¬angina
¬excessive	angina

In contrast, the conditional interpretation of the counterfactual assertion above, which is of the form “if not-p then not-q”, is captured by the following three possibilities:

excessive	angina
¬excessive	¬angina
excessive	¬angina

The third possibility in each set is different. On a conditional interpretation, causal and counterfactual assertions about the same factual situation are distinguished by their reference to one possibility which they do not hold in common.

Moreover, on a conditional interpretation, the causal assertion is aligned with the prediction assertion, whereas the counterfactual assertion is aligned with the prevention assertion. On a conditional interpretation, the prediction assertion above, which is of the form “if p then q”, will be understood with reference to the same three possibilities as the causal assertion, whereas the prevention assertion above, which is of the form “not-p prevented q”, will be understood with reference to the same three possibilities as the counterfactual assertion. On a conditional interpretation, the causal and prediction assertions both convey information useful for prediction (the presence of excessive exercise predicts angina, because there is only one possible situation in which excessive exercise occurs, and in that situation, angina occurs), but not for prevention (the absence of excessive exercise does not prevent angina, because there are

two possible situations in which excessive exercise does not occur, one in which angina does not occur and one in which angina occurs). In contrast, the counterfactual and prevention assertions both convey information useful for prevention (the absence of excessive exercise prevents angina, because there is only one possible situation in which excessive exercise does not occur, and in that situation, angina does not occur) but not prediction (the presence of excessive exercise does not predict angina, because there are two possible situations in which excessive exercise occurs, one in which angina occurs and one in which angina does not occur). If causal thinking is concerned with prediction and counterfactual thinking is concerned with prevention, then all four assertions, causal, counterfactual, prediction and prevention may be interpreted as conditionals.

The content of causal and counterfactual assertions

This experiment provides the first comparison of people’s interpretations of the meaning of causal, counterfactual, prediction and prevention assertions. We controlled for possible effects of the contents of the assertions by presenting all four assertions in the same contents, as in the “angina” example above. However, we were also concerned that participants would not be swayed by the particular contents used to interpret all four assertions as biconditionals or conditionals. Assertions with causal contents can be interpreted as expressing a biconditional relation in which the antecedent is both sufficient and necessary for the consequent, e.g., “If the butter is heated then it melts”, or as expressing a conditional relation in which the antecedent is sufficient but not necessary for the consequent, e.g., “If the car is out of gas then it stalls” (e.g., Cummins, 1995; Staudenmayer, 1975; Thompson, 1994). However, it may be possible to elicit a “basic” interpretation of assertions which describe causal relations by using “anomalous”

contents derived by arbitrarily recombining antecedent and consequent propositions from assertions describing real causal relations, e.g., “If she waters the plant then the light will go on” (Staudenmayer, 1975). Indicative conditionals with anomalous causal contents and their related causal assertions, e.g., “Her watering the plant causes the light to go on” may be interpreted as either biconditionals or conditionals (Goldvarg & Johnson-Laird, 1999; 2000; Staudenmayer, 1975).

We intended that participants would interpret the assertions as describing real causal relations, but that their interpretations would not be affected by their prior knowledge of the types of causal relations being described, e.g., whether the antecedent was sufficient or necessary for the consequent, and so we used “anomalous” causal contents. We used assertions in the past tense in the following five contents (adapted from Goldvarg & Johnson-Laird, 1999):

- Exercise that was excessive caused the occurrence of angina.
- Genes that were diverse caused the development of new diseases.
- A childhood that was happy caused the occurrence of amnesia.
- Use of solar energy caused the occurrence of global warming.
- Reactivity that was excessive caused the occurrence of damage to the reactor.

Method

Participants

The 25 postgraduates from a variety of departments at the University of Dublin, who took part voluntarily, included 14 women and 11 men. They were aged from 23 to 42 years, with a mean age of 29 years. None of the participants had any expertise in logic or the psychology of reasoning.

Design

The participants acted as their own controls. They were given four sorts of assertions: causal (e.g., Exercise that was excessive caused the occurrence of angina), counterfactual (e.g., If exercise had not been excessive, then angina would not have occurred), prediction (e.g., If exercise was excessive, then angina occurred), and prevention (e.g., Exercise that was not excessive prevented the occurrence of angina). They were given 20 problems, based on five instances of each type of assertion, corresponding to the five different contents. For each problem their task was to indicate whether each of the four logically possible situations (p and q, p and not-q, not-p and q, not-p and not-q) had been possible or impossible.

Materials and Procedure

The participants were tested in small groups. They were presented with booklets containing the instructions and materials (see Appendix 1g). They worked through the problems individually at their own pace. On the first page, they were given the following general instructions:

Imagine any two events, each of which may or may not occur, for example, drinking a certain liquid and becoming ill. Varying the occurrence versus non-occurrence of both events results in four possible scenarios:

- You drink the liquid and you become ill.
- You drink the liquid and you do not become ill.
- You do not drink the liquid and you become ill.
- You do not drink the liquid and you do not become ill.

However, if there is some *relationship* between the events, for example, if drinking or not drinking the liquid *causes* or *prevents* becoming ill, then it may be that only some of these scenarios are *really possible* and the others are *really impossible*. On the following pages, you will be given statements about relationships between pairs of past events. The statements themselves need not be true. In fact, some of them are true and some of them are false.

Regardless of whether you think that a statement is true or false, your task is to decide which scenarios were really possible and which were really impossible *if the statement is true*.

These instructions were followed by 20 problems of the following form:

Exercise that was excessive caused the occurrence of angina.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- | | |
|--|-------|
| Exercise that is excessive occurs and angina occurs. | |
| Exercise that is excessive occurs and angina does not occur. | |
| Exercise that is excessive does not occur and angina occurs. | |
| Exercise that is excessive does not occur and angina does not occur. | |

The problems were presented in a different random order to each participant, with three problems on each page of the booklet.

Results and Discussion

The biconditional interpretation was the most frequent interpretation, accounting for 56% of the interpretations of all assertions, as Table 4.2 shows. Participants made as many biconditional interpretations of causal assertions as counterfactual (53% vs 55%, Wilcoxon's $z = 0.34$, $n = 25$, $p = 0.74$), prediction (53% vs 58%, Wilcoxon's $z = 1.13$, $n = 25$, $p = 0.26$) or prevention assertions (53% vs 57%, Wilcoxon's $z = 0.68$, $n = 25$, $p = 0.49$). They made as many biconditional interpretations of counterfactual assertions as prediction (55% vs 58%, Wilcoxon's $z = 0.57$, $n = 25$; $p = 0.57$) or prevention assertions (55% vs 57%, Wilcoxon's $z = 0.23$, $n = 25$, $p = 0.82$). They made as many biconditional interpretations of prediction assertions as prevention assertions (58% vs 57%, Wilcoxon's $z = 0.78$, $n = 25$, $p = 0.78$).

Table 4.2: Percentages of ten interpretations of causal, counterfactual, prediction and prevention assertions in Experiment 7

Interpretation	Possibilities		Assertion				
			Causal	Counterfactual	Prediction	Prevention	Total
Biconditional	p not-p	q not-q	53	55	58	57	56
Conditional	p not-p not-p	q q not-q	25	19	31	23	25
Reversed conditional	p p not-p	q not-q not-q	7	7	5	6	6
Tautology	p p not-p not-p	q not-q q not-q	10	6	2	5	6
Ponens	p	q	3	0	0	4	2
Tollens	not-p	not-q	0	4	0	2	2
Strengthened antecedent	p p	q not-q	1	1	1	1	1
Relevance	p not-p	q q	1	0	0	1	0
Disabling	p p not-p	q not-q q	0	0	0	0	0
DA & AC	not-p	q	0	0	0	0	0
Other			1	8	2	2	3

Note: 25 participants each interpreted 5 assertions of each type

The conditional interpretation was the next most frequent interpretation, accounting for 25% of the interpretations of all assertions. Participants made as many conditional interpretations of causal assertions as counterfactual (25% vs 19%, Wilcoxon's $z = 0.83$, $n = 25$, $p = 0.41$), prediction (25% vs 31%, Wilcoxon's $z = 0.8$, $n = 25$, $p = 0.42$) or prevention assertions (25% vs 23%, Wilcoxon's $z = 0.31$, $n = 25$, $p = 0.76$). They made as many conditional interpretations of counterfactual assertions as prediction (19% vs 31%, Wilcoxon's $z = 1.41$, $n = 25$; $p = 0.16$) or prevention assertions (19% vs 23%, Wilcoxon's $z = 0.55$, $n = 25$, $p = 0.58$). They made as many conditional interpretations of prediction assertions as prevention assertions (23% vs 31%, Wilcoxon's $z = 1.26$, $n = 25$, $p = 0.21$).

The content of the assertions had no effect on the frequencies of biconditional interpretations, but it had some effects on the frequencies of conditional interpretations (see Appendix 4). There were individual differences in participants' tendencies to make biconditional or conditional interpretations of the assertions. 56% of participants interpreted more than half of the assertions as biconditionals, whereas 16% of participants interpreted more than half of the assertions as conditionals. None of the remaining eight interpretations was made for more than 10% of any of the assertions⁹.

The results suggest that causal, counterfactual, prediction and prevention assertions often have the same interpretation. Just over half of the assertions were interpreted as biconditionals,

⁹Three of the remaining interpretations were made more for some sorts of assertions than others. First, there were more tautological interpretations of causal than prediction assertions, although the difference was marginal (10% vs 2%, Wilcoxon's $z = 1.72$, $n = 25$, $p = 0.09$). Second, there were more "ponens" interpretations (referring only to the p and q possibility) of prevention assertions than either causal or prediction assertions, although these differences were both also marginal (both 2% vs 0%, Wilcoxon's $z = 1.83$, $n = 25$, $p = 0.07$). Third, there were more "tollens" interpretations (referring only to the not-p and not-q possibility) of counterfactual assertions than either causal or prediction assertions (both 4% vs 0%, Wilcoxon's $z = 2.02$, $n = 25$, $p = 0.04$). There were more "other" interpretations, which did not fall into any of the ten categories, of counterfactual assertions than causal, prediction, or prevention assertions (8% vs 1%, Wilcoxon's $z = 2.29$, $n = 25$, $p = 0.02$; 8% vs 2%, Wilcoxon's $z = 1.96$, $n = 25$, $p = 0.05$; 8% vs 2%, Wilcoxon's $z = 2.52$, $n = 25$, $p = 0.01$).

which is consistent with previous studies of the interpretation of causal assertions (Goldvarg & Johnson-Laird, 2000; Staudenmayer, 1975). We can conclude that many causal and counterfactual assertions are interpreted in ways which convey information about both prediction and prevention. However, the results also show that a sizable minority, almost a quarter, of causal and counterfactual assertions about the same factual situation have different interpretations, often a conditional interpretation. Moreover, there were as many conditional interpretations of prediction and prevention assertions as causal and counterfactual assertions. We can conclude that at least some causal assertions are interpreted in a way which conveys information about prediction but not prevention, whereas some counterfactual assertions are interpreted in a way which conveys information about prevention but not prediction.

Overall, the results suggest that people interpret a causal assertion such as “Exercise that was excessive caused the occurrence of angina” to refer to the same possibilities as a counterfactual assertion such as “If exercise had not been excessive, then angina would not have occurred”. Both assertions are taken to mean that it was possible for both exercise that was excessive and angina to occur, and for neither exercise that was excessive nor angina to occur. The two assertions are judged to mean that it impossible for exercise that was excessive to occur and angina to not occur, and that it was impossible for exercise that was excessive to not occur and angina to occur. Likewise, a prediction assertion such as “If exercise was excessive, then angina occurred”, and a prevention assertion such as “Exercise that was not excessive prevented the occurrence of angina” were also taken to refer to these same two possibilities and no other ones.

Of course, asking people to judge which situations were possible from a list of explicitly presented situations may force them to think more explicitly about each situation than they

would otherwise do, and hence it may remove subtle differences in which possibilities spontaneously come to mind for each of these sorts of assertions. Therefore, the aim of our next two experiments was to examine indirectly the initial possibilities that people may keep in mind of causal, counterfactual, prediction and prevention assertions.

Experiment 8

The aim of this experiment was to compare people’s *initial* mental models of causal, counterfactual, prediction and prevention assertions. The previous experiment showed that people often reach the same sorts of *fully explicit* interpretations of these assertions: they tend to judge all four assertions to mean that the two situations of a biconditional relation were possible. However, although people may judge an assertion to be consistent with multiple possibilities, they may need to mentally represent these possibilities economically because of their limited working memories (e.g., Johnson-Laird and Byrne, 1991; see Chapter 1, pp. 57-58). For example, an indicative conditional (one of our prediction assertions) such as “If fertiliser is put on the plants, then they grow quickly” may be interpreted as a biconditional, as we saw in the last experiment, and so it may be considered consistent with the following two possibilities:

fertiliser	grow
–fertiliser	–grow

However, it may initially be represented as a set of mental models which make explicit just some of the information, such as:

fertiliser	grow
...	

where the three dots represent an implicit model indicating that there may be additional possibilities which have not been explicitly represented. We suggest that if causal thinking is concerned with prediction, whereas counterfactual thinking is concerned with prevention, then people may construct different initial mental models of causal and counterfactual assertions, whereas they should construct the same initial mental models of causal and prediction assertions, and they should construct the same initial mental models of counterfactual and prevention assertions.

One way to examine people's initial representations of an assertion is to investigate the sorts of inferences they make from the assertion. There has been extensive research on the inferences people make from indicative conditionals such as "If fertiliser is put on the plants, then they grow quickly" (for a review see Evans et al., 1993). Given such an assertion, people may make the modus ponens inference from "Fertiliser is put on the plants" to "The plants grow quickly", the affirmation of the consequent inference from "The plants grow quickly" to "Fertiliser is put on the plants", the denial of the antecedent inference from "Fertiliser is not put on the plants" to "The plants do not grow quickly" and the modus tollens inference from "The plants do not grow quickly" to "Fertiliser is not put on the plants" (see Chapter 1, pp. 58-59). People's tendencies to make the affirmation of the consequent and denial of the antecedent inferences are quite variable, perhaps because these inferences are invalid on a conditional interpretation of the assertion, although they are valid on a biconditional interpretation. However, the modus ponens and modus tollens inferences are both valid regardless of whether a biconditional or conditional interpretation is taken, yet people reliably make the modus ponens inference more readily than the modus tollens inference (for a review see Evans et al., 1993).

This finding suggests that people’s inferences depend on their initial mental models rather than their fully explicit interpretation of an assertion. An indicative conditional such as “If fertiliser is put on the plants, then they grow quickly” may be initially represented by only the affirmative case and an implicit model:

fertiliser grow
...

For the modus ponens inference, the information in the second premise, “Fertiliser is put on the plants” may be represented as follows:

fertiliser

and the two models can be combined readily to result in the conclusion, “The plants grow quickly”. Similarly, for the affirmation of the consequent inference, the information in the second premise, “The plants grow quickly” may be represented as follows:

grow

and this model can be combined readily with the model of the first premise to result in the conclusion, “Fertiliser is put on the plants”. In contrast, for the modus tollens inference, the information in the second premise, “The plants do not grow quickly”:

¬grow

cannot be combined with the initial set of models of the first premise, and so people often erroneously conclude that nothing follows. Likewise, for the denial of the antecedent inference, the information in the second premise, “Fertiliser is not put on the plants”:

¬fertiliser

has nothing in common with the initial set of models of the first premise. People may make the modus tollens and denial of the antecedent inferences by fleshing out their initial models of the first premise to be more explicit. They may be fleshed out to the biconditional set of models:

fertiliser	grow
¬fertiliser	¬grow

and the information from either second premise can now be added to eliminate the first model, resulting in the conclusion, “Fertiliser is not put on the plants” for modus tollens, or “The plants do not grow quickly” for the denial of the antecedent. Alternatively, the initial set of models may be fleshed out to the conditional set of models:

fertiliser	grow
¬fertiliser	¬grow
¬fertiliser	grow

and for modus tollens, the information from the second premise eliminates all but the second model, again resulting in the conclusion, “Fertiliser is not put on the plants”. For the denial of the antecedent, the information from the second premise rules out all but the last two models, leading to the conclusion that nothing follows. In either case, inferences which require people to flesh out their initial models and keep multiple possibilities in mind are made less often and take more time than inferences that can be made directly from the initial models (e.g., Johnson-Laird et al., 1992; see also Chapter 1, pp. 57-61).

In this experiment, we will investigate people’s initial mental models of causal, counterfactual, prediction and prevention assertions by comparing the frequencies of their inferences from the four sorts of assertions. We will focus on the modus ponens and denial of the antecedent inferences because they may be most relevant to thinking about causal relations.

Modus ponens and denial of the antecedent are both *forward* inferences, in which an antecedent implies a consequent, whereas affirmation of the consequent and modus tollens are both *backward* inferences, in which a consequent implies an antecedent (see also Chapter 1, pp. 58-59). Although there is no difference in the logical validity of forward and backward inferences, people tend to think about causes having implications for their effects, rather than effects having implications for their causes, perhaps because effects cannot precede their causes in time (e.g., Tversky & Kahneman, 1980).

There is abundant evidence that the initial set of models for an indicative conditional usually represents explicitly only the affirmative case (e.g., Johnson-Laird & Byrne, 1991; for a review see Evans et al., 1993), and so we expect that participants will make more modus ponens than denial of the antecedent inferences from prediction assertions phrased as indicative conditionals. If prediction and causal assertions are aligned, then the initial representation of causal assertions such as “Putting fertiliser on the plants causes them to grow quickly” should also refer just to the affirmative case:

fertiliser	grow
...	

and so participants should also make more modus ponens than denial of the antecedent inferences from causal assertions. A single initial model of the affirmative case conveys information useful for prediction, but not prevention. It represents explicitly only one situation in which the plants grow quickly, and in that situation, fertiliser is put on the plants, and so it suggests that putting fertiliser on the plants predicts that they grow quickly. In contrast, it does not represent explicitly any situations in which the plants do not grow quickly, and so it conveys no information about what prevents the plants from growing quickly.

In contrast, a counterfactual assertion such as “if fertiliser were put on the plants then they would grow quickly” may be represented in a more explicit set of initial models from the outset (Byrne & Tasso, 1999; Johnson-Laird & Byrne, 1991; Moreno-Ríos, García-Madruga & Byrne, 2000; Thompson & Byrne, 2000). When people understand a counterfactual assertion, they may initially keep in mind not only the hypothesised affirmative case, in which fertiliser is put on the plants and they grow quickly, but also the presupposed factual case, in which fertiliser is not put on the plants and they do not grow quickly:

factual:	¬fertiliser	¬grow
counterfactual:	fertiliser	grow

and the models may be annotated to keep track of their epistemic status. As a result, from counterfactual assertions, participants make the inferences which require access to the negative case, such as the denial of the antecedent, as frequently as they make the inferences that require access to the affirmative case, such as modus ponens (e.g., Byrne & Tasso, 1999; Moreno-Ríos et al., 2000; Thompson & Byrne, 2000). Compared to prediction assertions, counterfactual assertions tend to evoke more negative inferences (e.g., Byrne & Tasso, 1999; Thompson & Byrne, 2000) and sometimes also less affirmative inferences (e.g., Byrne & Tasso, 1999; Moreno-Ríos et al., 2000). This latter finding suggest that keeping in mind both the affirmative and negative possibilities in the initial set of models of a counterfactual assertion may be a double-edged sword: it may enable the negative inferences such as the denial of the antecedent by removing the need to flesh out the initial models to make this inference, but it may also disable the affirmative inferences such as modus ponens due to the greater difficulty of keeping two possibilities in mind in the initial set of models (Byrne & Tasso, 1999).

We expect that participants will make as many modus ponens as denial of the antecedent inferences from counterfactual assertions, as in previous studies. If prevention and counterfactuals are aligned, then the initial representation of a prevention assertion such as “Putting fertiliser on the plants prevents them from not growing quickly” may also refer to both the affirmative and negative cases:

$\neg$ fertiliser	$\neg$ grow
fertiliser	grow
...	

and so participants should also make as many modus ponens as denial of the antecedent inferences from prevention assertions. Unlike the single initial model of an indicative conditional, these two initial models convey information useful for prevention: the contrast between the affirmative and negative cases shows how the affirmative case can be prevented.

In sum, the existing evidence suggests that people’s mental models of prediction assertions (indicative conditionals) represent explicitly only the affirmative case, and so people make more modus ponens than denial of the antecedent inferences from prediction assertions. In contrast, people’s mental models of counterfactual assertions represent explicitly both the affirmative and negative cases and so the difference between modus ponens and denial of the antecedent disappears. We expect that causal assertions are initially represented in the same way as prediction assertions, and so participants should also make more modus ponens than denial of the antecedent inferences from causal assertions, whereas prevention assertions are initially represented in the same way as counterfactual assertions, and so participants should also make as many modus ponens as denial of the antecedent inferences from prevention assertions. The alternative possibility is that causal and counterfactual thinking do not differ in their relations to

prediction and prevention, in which case people may construct the same mental models of causal, counterfactual, prediction and prevention assertions, and hence the relative frequencies of modus ponens and denial of the antecedent inferences from the four assertions should be the same.

Method

Participants

The 22 first year psychology students at the University of Dublin, who took part voluntarily, included 12 women and 9 men (and one person who did not record their gender). They were aged from 18 to 34 years, with a mean age of 21 years (1 person did not record their age). None of the participants had any expertise in logic or the psychology of reasoning.

Design

The participants acted as their own controls. We gave them 16 problems, half of them based on affirmative assertions and half based on negative assertions. The 8 affirmative problems were based on one modus ponens and one denial of the antecedent inference for each of four assertions in the present tense: a causal (e.g., Putting fertiliser on the plants causes them to grow quickly), a counterfactual (e.g., If fertiliser were put on the plants, then they would grow quickly), a prediction (e.g., If fertiliser is put on the plants, then they grow quickly) and a prevention assertion (e.g., Putting fertiliser on the plants prevents them from not growing quickly). The 8 negative problems had a similar structure except that the antecedent and consequent were negated.

Materials and Procedure

The participants were tested in small groups. They worked through the problems individually at their own pace. They were presented with booklets containing the instructions and materials (see Appendix 1h). On the first page, they were given the following general instructions:

On the following pages, you will be asked to make inferences from information about everyday events and situations. In each case, you will be presented with two factual statements, and you will be asked to say what, if anything, follows. Here is an example:

All Trinity students are intelligent.
Paul is a Trinity student.
.....
What, if anything, follows?

In this example, if you think that the two statements, “All Trinity students are intelligent” and “Paul is a Trinity student” necessarily imply that Paul is intelligent, then you should write down “Paul is intelligent” as your answer. If you think that the statements necessarily imply that Paul is not intelligent, then you should write “Paul is not intelligent”. If you think that the statements imply nothing definite about Paul’s intelligence, then you should write “Nothing follows”.

The problems that you will complete will have the same form as the example shown above, but in each case, you will be asked to draw conclusions from different statements. It is important to read each set of statements very carefully, and to think about your answers. There are 16 problems in total, and it is essential that you complete all of them. Please complete them in the order in which they are presented. Do not look ahead, or go back over previous answers.

These instructions were followed by 16 problems with the same form as the example above. The content of the assertions were realistic causal relations. To avoid participants’ interpretations being unduly influenced by their prior knowledge of the types of causal relations being described, we chose contents which described causal relations in which the antecedent is neither

completely necessary nor completely sufficient for the consequent (e.g, Cummins, 1995; Thompson, 1994). Each participant received one problem with each type of first premise in each of the following four contents (adapted from Cummins, 1995; see also Appendix 1h):

- Putting fertiliser on the plants causes them to grow quickly.
- Depressing the brake causes the car to slow down.
- Studying hard causes John to get good marks.
- Turning on the air conditioning causes Jenny to feel cool.

The contents were randomly assigned to the problems for each participant, and the problems were presented in a different random order to each participant, with four problems on each page of the booklet.

Results and Discussion

The results show a similar pattern of inferences from causal and prediction assertions, which differs in systematic ways from the pattern of inferences from counterfactual and prevention assertions, as Table 4.3 shows. For affirmative assertions, as we predicted, participants made more modus ponens than denial of the antecedent inferences from causal assertions (91% vs 55%, $n = 22$, binomial $p < 0.01$) and prediction assertions (100% vs 55%, $n = 22$, binomial $p < 0.01$), but they made as many modus ponens as denial of the antecedent inferences from counterfactual assertions (82% vs 57%, $n = 21$, binomial $p = 0.23$) and prevention assertions (73% vs 52%, $n = 21$, binomial $p = 0.18$). The difference between modus ponens and denial of the antecedent was greater for causal assertions (36% difference between 91% and 55%) than counterfactual assertions (25% difference between 82% and 57%), although this interaction did not reach significance (Wilcoxon's $z = 1.09$, $n = 21$, $p = 0.27$). Likewise, the difference between modus ponens and denial of the antecedent was greater for prediction

assertions (45% difference between 100% and 55%) than prevention assertions (21% difference between 73% and 52%), although this interaction did not reach significance (Wilcoxon's $z = 1.57$, $n = 21$, $p = 0.12$). These results support our suggestion that causal and prediction assertions are understood by constructing an initial set of models that makes available the affirmative possibility but keeps other information implicit; in contrast, counterfactual and prevention assertions are understood by constructing a more explicit set of initial models that makes available not only the affirmative possibility but also the negative possibility.

For negative assertions, participants made more modus ponens than denial of the antecedent inferences from all four assertions (causal: 100% vs 43%, $n = 21$, binomial $p < 0.01$; counterfactual: 91% vs 55%, $n = 22$, binomial $p < 0.01$; prediction: 91% vs 55%, $n = 22$, binomial $p = 0.02$; prevention: 90% vs 54%, $n = 22$, binomial $p = 0.04$). However, the difference between modus ponens and denial of the antecedent was somewhat greater for causal assertions (57% difference between 100% and 43%) than counterfactual assertions (36% difference between 91% and 55%), although this interaction did not reach significance (Wilcoxon's $z = 1.26$, $n = 21$, $p = 0.21$). The difference between modus ponens and denial of the antecedent was not greater for prediction assertions (36% difference between 91% and 55%) than prevention assertions (36% difference between 90% and 54%, Wilcoxon's $z = 0.28$, $n = 22$, $p = 0.78$). The greater difficulty of understanding negative assertions (see e.g., Evans et al., 1993) may make it more difficult for people to keep two models in mind, resulting in a difference in the frequencies of modus ponens and denial of the antecedent inferences from counterfactual and prevention assertions only when they are negative.

Table 4.3: Percentages of participants making each of the inferences from each assertion in Experiment 8

Assertion type	Assertion valence	Inference	
		Modus ponens	Denial of the antecedent
Causal	Affirmative	91	55
	Negative	100	43
	Mean	95	49
Counterfactual	Affirmative	82	57
	Negative	91	55
	Mean	86	56
Prediction	Affirmative	100	55
	Negative	91	55
	Mean	95	55
Prevention	Affirmative	73	52
	Negative	90	54
	Mean	81	53
All assertions	Affirmative	86	55
	Negative	93	52
	Mean	90	53

Participants tended to make more modus ponens inferences from causal and prediction assertions (95% in each case) than from counterfactual and prevention assertions (86% and 81%). They made more modus ponens inferences from causal assertions than counterfactual assertions (although this difference was marginal, 95% vs 86%, Wilcoxon’s $z = 1.83$, $n = 22$, $p = 0.07$) or prevention assertions (95% vs 81%, Wilcoxon’s $z = 2.37$, $n = 22$, $p = 0.02$), whereas they made as many modus ponens inferences from causal assertions as prediction assertions (95% vs 95%, Wilcoxon’s $z = 0$, $n = 22$, $p = 1$). They made as many modus ponens inferences from counterfactual assertions as from prevention (86% vs 81%, Wilcoxon’s $z = 0.89$, $n = 22$, p

= 0.37) or prediction assertions (86% vs 95%, Wilcoxon's $z = 1.35$, $n = 22$, $p = 0.18$). They made more modus ponens inferences from prediction assertions than prevention assertions (95% vs 81%, Wilcoxon's $z = 2.37$, $n = 22$, $p = 0.02$)¹⁰. People may make less modus ponens inferences from counterfactual and prevention assertions than from causal and prediction assertions due to the greater difficulty of keeping in mind two possibilities in their initial mental models of these assertions (e.g., Byrne & Tasso, 1999; Moreno-Ríos et al., 2000).

There were no differences in the frequencies of denial of the antecedent inferences made from any of the assertions. There were no differences for causal assertions compared to counterfactual (49% vs 56%, Wilcoxon's $z = 0.91$, $n = 20$, $p = 0.36$), prediction (49% vs 55%, Wilcoxon's $z = 0.63$, $n = 21$, $p = 0.53$) or prevention assertions (49% vs 53%, Wilcoxon's $z = 0.71$, $n = 21$, $p = 0.48$). There were no differences for counterfactual assertions compared to prediction (56% vs 55%, Wilcoxon's $z = 0$, $n = 21$, $p = 1$) or prevention assertions (56% vs 53%, Wilcoxon's $z = 0$, $n = 20$, $p = 1$). There were no differences between prediction and prevention assertions (55% vs 53%, Wilcoxon's $z = 0.28$, $n = 21$, $p = 0.78$)¹¹. Some previous

¹⁰ For affirmative assertions, participants made as many modus ponens inferences from causal assertions as counterfactual (91% vs 82%, $n = 22$, binomial $p = 0.5$), prediction (91% vs 100%, $n = 22$, binomial $p = 0.5$) or prevention assertions (91% vs 73%, $n = 22$, binomial $p = 0.22$). They made as many from counterfactual assertions as prediction (82% vs 100%, $n = 22$, binomial $p = 0.12$) or prevention assertions (82% vs 73%, $n = 22$, binomial $p = 0.69$), but they made more from prediction assertions than prevention assertions (100% vs 73%, $n = 22$, binomial $p = 0.03$). For negative assertions, they made as many modus ponens inference from all four assertions (causal vs counterfactual: 100% vs 91%, $n = 22$, binomial $p = 0.5$; causal vs prediction: 100% vs 91%, $n = 22$, binomial $p = 0.5$; causal vs prevention: 100% vs 90%, $n = 22$, binomial $p = 0.25$; counterfactual vs prediction: 91% vs 91%, $n = 22$, binomial $p = 1$; counterfactual vs prevention: 91% vs 90%, $n = 22$, binomial $p = 1$; prediction vs prevention: 91% vs 90%, $n = 22$, binomial $p = 1$).

¹¹ Participants made as many denial of the antecedent inference from all four assertions both for affirmative assertions (causal vs counterfactual: 55% vs 57%, $n = 21$, binomial $p = 1$; causal vs prediction: 55% vs 55%, $n = 22$, binomial $p = 1$; causal vs prevention: 55% vs 52%, $n = 21$, binomial $p = 1$; counterfactual vs prediction: 57% vs 55%, $n = 21$, binomial $p = 1$; counterfactual vs prevention: 57% vs 52%, $n = 20$, binomial $p = 1$; prediction vs prevention: 55% vs 52%, $n = 21$, binomial $p = 1$), and for negative assertions (causal vs counterfactual: 43% vs 55%, $n = 21$, binomial $p = 0.69$; causal vs prediction: 43% vs 55%, $n = 21$, binomial $p = 0.69$; causal vs prevention: 43% vs 54%, $n = 21$, binomial $p = 0.55$; counterfactual vs prediction: 55% vs 55%, $n = 22$, binomial $p = 1$; counterfactual vs prevention: 55% vs 54%, $n = 22$, binomial $p = 1$; prediction vs prevention: 55% vs 54%, $n = 22$, binomial $p = 1$).

studies have found higher rates of denial of the antecedent inferences from counterfactual than indicative conditionals (Byrne & Tasso, 1999; Thompson & Byrne, 2000). However, the demonstration that participants made more affirmative than negative inferences from affirmative indicative conditionals, but not from affirmative counterfactual conditionals replicates previous findings (Moreno-Ríos et al., 2000).

The results show a similar pattern of inferences from causal and prediction assertions, which differs in systematic ways from the pattern of inferences from counterfactual and prevention assertions. For affirmative assertions, participants made more modus ponens than denial of the antecedent inferences from causal and prediction assertions, but they made just as many modus ponens as denial of the antecedent inferences from counterfactual and prevention assertions. These results support our proposal that causal and prediction assertions are represented initially by a single explicit mental model of the affirmative possibility, which may convey information useful for prediction but not prevention. In contrast, counterfactual and prevention assertions are represented from the outset by a richer set of mental models corresponding to both the affirmative and the negative possibilities, which may convey information useful for prevention. The next experiment provides a further test of this hypothesis by comparing the times people take to make modus ponens and denial of the antecedent inferences from causal, counterfactual, prediction and prevention assertions.

Experiment 9

The aim of this experiment was to compare the length of time it takes people to make the modus ponens and denial of the antecedent inferences from causal, counterfactual, prediction and prevention assertions. Inferences that require people to flesh out their models and keep multiple

possibilities in mind are not only made less often than inferences that can be made directly from the initial mental models, they also take longer when they are made (e.g., Johnson-Laird & Byrne, 1991). Hence, we expect that people will make modus ponens inferences more quickly than denial of the antecedent inferences from causal and prediction assertions. Causal and prediction assertions may be represented by an initial set of models that makes explicit only the affirmative case, and so the modus ponens inference may be made readily from the initial models, whereas the denial of the antecedent inference may require the models to be fleshed out, which takes time. In contrast, we expect that there will be no differences in the latencies of modus ponens and denial of the antecedent inferences from counterfactual and prevention assertions. Counterfactual and prevention assertions may be represented by an initial set of models that makes explicit both the affirmative and the negative cases. Hence, the modus ponens inference may be made less quickly than from causal and prediction assertions due to the greater difficulty of keeping in mind two possibilities in the initial set of models, whereas the denial of the antecedent inference may be made more quickly than from causal and prediction assertions because it can be made without fleshing out the initial models.

Method

Participants and design

Thirty-one members of the participant panel of the Psychology Department, University of Dublin, participated in the experiment. The participant panel is composed of members of the general population and they were paid for their participation at a rate of IR£4 per hour, plus £2 expenses. The participants were 20 women and 11 men aged from 18 to 65 years, with a mean age of 42 years. The design of the experiment was exactly the same as that of Experiment 8. In

this experiment the latency to make each inference was recorded as well as the nature of the inference made.

Materials and Procedure

Participants were tested individually. The experiment was run on a portable PC. Participants were first given verbal instructions by the experimenter (see Appendix 1i). They were informed that they would work through a number of problems on the computer at their own pace. They were told that for each problem, the computer screen would display a sentence, followed by another sentence, followed by a question. They were told that when they had finished reading a sentence and wanted to proceed to the next screen, they should press the key marked “NEXT”, and when they wanted to answer a question, they should press one of the keys marked “Y” for “yes”, “N” for “no” or “CT” for “can’t tell”. The space bar was marked “NEXT”, and the “q”, “y” and “[“ keys were used for the answers. To control for possible effects of the positions of the answer keys, different keys corresponded to different answers for different participants. Participants were instructed that they could take a break and ask the experimenter questions between problems, but they shouldn’t stop in the middle of a problem. They were not informed that their responses were being timed until they had completed the experiment. When it was clear that a participant had understood the procedure, the experimenter demonstrated it with one example problem, and then the participant completed eight different example problems (see Appendix 1i). The example problems were based on non-conditional and non-causal premises such as relational, quantificational and disjunctive assertions. Participants were encouraged to ask questions during and after completing the example problems.

As in Experiment 8, participants completed 16 experimental inference problems. The problems had the following form:

Putting fertiliser on the plants causes them to grow quickly.

Fertiliser is put on the plants.

Do the plants grow quickly?

Each problem was assigned a different content in this experiment, and the 16 contents are shown in Appendix 1i. They include the four non-necessary and non-sufficient causal contents used in Experiment 8 (adapted from Cummins, 1995) and an additional 12 non-necessary and non-sufficient causal contents (adapted from Thompson, 1994). The 16 contents were rotated over the 16 problems, and so there were 16 different versions of the materials. The different versions were assigned at random to the participants and the 16 problems were presented in a different random order to each participant. Participants worked through the problems at their own pace under the supervision of the experimenter. The first premise, second premise and inference question were presented separately on the computer screen. Participants pressed a key when they had read a premise. The keypress to the first premise replaced it by the second premise, and the keypress to the second premise replaced it with the inference question. The computer recorded the time they took for each key press and their answers to the inference question.

Results and Discussion

Latencies

The mean latencies to endorse the modus ponens and denial of the antecedent inferences are shown in Table 4.4. The times taken to read each premise are also shown.

Table 4.4: Mean latencies (in milliseconds) to endorse each of the inferences from each assertion in Experiment 9

Assertion type	Assertion valence	Inference					
		Modus ponens			Denial of the antecedent		
		1 st pre.	2 nd pre.	concl.	1 st pre.	2 nd pre.	concl.
Causal	Affirmative	624	222	338	586	272	362
	Negative	1199	364	297	955	301	264
	Mean	911	293	316	770	286	315
Counterfactual	Affirmative	671	276	241	928	309	404
	Negative	928	309	298	960	265	242
	Mean	799	292	269	944	287	329
Prediction	Affirmative	504	252	260	541	320	449
	Negative	848	278	330	828	309	274
	Mean	676	265	277	684	314	368
Prevention	Affirmative	1145	254	402	1344	445	361
	Negative	763	319	413	902	236	231
	Mean	954	286	409	1123	340	296
All assertions	Affirmative	736	251	295	850	336	398
	Negative	934	317	317	911	278	254
	Mean	835	284	314	880	307	331

Notes:

“1st pre.” refers to the time taken to read the first premise; “2nd pre.” refers to the time taken to read the second premise; “concl.” refers to read the time taken to make a conclusion. Latencies are for endorsements of an inference. Latencies which exceeded 3000 ms were excluded.

Data plots showed that the latencies approximated an exponential distribution, and so they were log transformed for the statistical analyses. We carried out within-participant comparisons of the latencies of pairs of inferences by selecting participants who made both inferences, and comparing their log-transformed conclusion times for the two inferences.

Comparing modus ponens and denial of the antecedent

We first compared the latencies of modus ponens and denial of the antecedent inferences from each assertion. Table 4.5 reports the mean log-transformed conclusion times of participants who made both inferences from an assertion.

Table 4.5: Mean log transformed latencies (in milliseconds) to endorse both the modus ponens and denial of the antecedent inferences from each assertion in Experiment 9

Assertion type	Assertion valence	Inference	
		Modus ponens	Denial of the antecedent
Causal	Affirmative (n = 16)	3.33	3.46
	Negative (n = 16)	3.42	3.36
Counterfactual	Affirmative (n = 19)	3.35	3.47
	Negative (n = 15)	3.41	3.35
Prediction	Affirmative (n = 19)	3.34	3.48
	Negative (n = 19)	3.39	3.36
Prevention	Affirmative (n = 7)	3.59	3.56
	Negative (n = 14)	3.39	3.34

Note: Participants who endorsed *both* the modus ponens and the denial of the antecedent inference from an assertion were selected. Latencies refer to the time taken to make a conclusion; they do not include the time taken to read the premises. Latencies which exceeded 3000 ms were excluded.

As in Experiment 8, the results show a similar pattern from causal and prediction assertions, which differs in systematic ways from the pattern for counterfactual and prevention assertions. For the affirmative assertions, participants made the modus ponens inference more quickly than the denial of the antecedent inference from the causal assertion [$M = 3.33$ vs $M = 3.46$, $t(15) = 2.33$, $p = 0.03$] and the prediction assertion [although this difference was only marginally significant, $M = 3.34$ vs $M = 3.48$, $t(18) = 1.83$, $p = 0.08$], whereas they made the modus ponens inference at the same speed as the denial of the antecedent inference from the counterfactual assertion [$M = 3.35$ vs $M = 3.47$, $t(18) = 1.47$, $p = 0.16$] and the prevention assertion [$M = 3.59$ vs $M = 3.56$, $t(6) = 0.22$, $p = 0.83$]^{12, 13}. People may make modus ponens inferences faster than denial of the inferences from causal and prediction assertions because their mental models of these assertions initially represent explicitly only the affirmative possibility, and so they must flesh out their models to be more explicit in order to make the denial of the antecedent inference. They may make modus ponens and denial of the antecedent inferences at the same speed from counterfactual and prevention assertions because their mental models of these assertions initially represent explicitly both the affirmative and negative cases, and so they can make both inferences without having to flesh out their models.

For negative assertions, participants made the modus ponens inference at the same speed as the denial of the antecedent inference from all four assertions [causal: $M = 3.42$ vs $M = 3.36$, $t(15) = 1.07$, $p = 0.3$; counterfactual: $M = 3.41$ vs $M = 3.35$, $t(14) = 0.94$, $p = 0.36$; prediction: M

¹² This last result should be interpreted with caution as only 7 participants made both inferences from the affirmative prevention assertion.

¹³ Unfortunately, we could not analyse the interaction between the effects of the type of inference and the type of assertion on the latencies, due to the low frequencies of participants (ranging from 2 to 12) who made all four inferences involved in each comparison.

= 3.39 vs M = 3.36, $t(18) = 0.7$, $p = 0.49$; prevention: M = 3.39 vs M = 3.34, $t(13) = 0.71$, $p = 0.49$]. As in Experiment 8, the difference in findings for affirmative and negative assertions may reflect the greater difficulty of processing negative assertions. This additional source of difficulty may conceal the greater difficulty of denial of the antecedent than modus ponens inferences from causal and prediction assertions.

Comparing assertions

We next examined the effects of the type of assertion on the latencies of modus ponens and denial of the antecedent inferences. Again, for each comparison, we selected participants who made both inferences and compared their log-transformed conclusion times for the two inferences. For affirmative assertions, participants made the modus ponens inference faster from the prediction assertion than the causal assertion [M = 3.29 vs M = 3.49, $t(18) = 2.77$, $p = 0.01$], but there were no other differences between the assertions [causal vs counterfactual: M = 3.38 vs M = 3.3, $t(17) = 1.55$, $p = 0.14$; causal vs prevention: M = 3.32 vs M = 3.34, $t(5) = 0.39$, $p = 0.71$; counterfactual vs prediction: M = 3.33 vs M = 3.35, $t(17) = 0.35$, $p = 0.73$; counterfactual vs prevention: M = 3.37 vs M = 3.39, $t(8) = 0.51$, $p = 0.63$; prediction vs prevention: M = 3.41 vs M = 3.54, $t(8) = 1.07$, $p = 0.32$]¹⁴. For negative assertions, there were no differences between the assertions for the latencies of modus ponens inferences [causal vs counterfactual: M = 3.4 vs M = 3.42, $t(21) = 0.4$, $p = 0.69$; causal vs prediction: M = 3.42 vs M = 3.47, $t(23) = 1.15$, $p = 0.26$; causal vs prevention: M = 3.45 vs M = 3.44, $t(17) = 0.12$, $p = 0.9$; counterfactual vs prediction: M = 3.49 vs M = 3.43, $t(22) = 0.23$, $p = 0.82$; counterfactual vs prevention: M = 3.45

¹⁴ It should be noted that 3 of these comparisons involved less than 10 participants (causal vs prevention, counterfactual vs prevention and prediction vs prevention).

vs $M = 3.4$, $t(16) = 0.67$, $p = 0.51$; prediction vs prevention: $M = 3.45$ vs $M = 3.47$, $t(18) = 0.24$, $p = 0.81$].

There were no differences between the assertions for the latencies of the denial of the antecedent inferences for the affirmative assertions [causal vs counterfactual: $M = 3.44$ vs $M = 3.47$, $t(17) = 0.42$, $p = 0.68$; causal vs prediction: $M = 3.47$ vs $M = 3.43$, $t(16) = 0.42$, $p = 0.68$; causal vs prevention: $M = 3.57$ vs $M = 3.46$, $t(11) = 0.96$, $p = 0.36$; counterfactual vs prediction: $M = 3.48$ vs $M = 3.41$, $t(19) = 0.94$, $p = 0.36$; counterfactual vs prevention: $M = 3.46$ vs $M = 3.47$, $t(14) = 0.2$, $p = 0.84$; prediction vs prevention: $M = 3.53$ vs $M = 3.48$, $t(16) = 0.57$, $p = 0.58$], or for the negative assertions [causal vs counterfactual: $M = 3.37$ vs $M = 3.36$, $t(13) = 0.44$, $p = 0.71$; causal vs prediction: $M = 3.35$ vs $M = 3.83$, $t(15) = 0.6$, $p = 0.55$; causal vs prevention: $M = 3.33$ vs $M = 3.33$, $t(11) = 0.08$, $p = 0.93$; counterfactual vs prediction: $M = 3.34$ vs $M = 3.39$, $t(14) = 0.86$, $p = 0.41$; counterfactual vs prevention: $M = 3.35$ vs $M = 3.36$, $t(11) = 0.15$, $p = 0.88$; prediction vs prevention: $M = 3.33$ vs $M = 3.35$, $t(14) = 0.45$, $p = 0.66$].

The results show a similar pattern of inference latencies from causal and prediction assertions, which differs in systematic ways from the pattern of inference latencies from counterfactual and prevention assertions. For affirmative assertions, participants made modus ponens inferences more quickly than denial of the antecedent inferences from causal and prediction assertions, but they made modus ponens inferences at just the same speed as denial of the antecedent inferences from counterfactual and prevention assertions. This pattern for the latencies exactly replicates the pattern for endorsements found in the previous experiment. It provides further corroboration for our suggestion that causal and prediction assertions are represented in the same way as each other, by an initial model corresponding to the affirmative case, whereas counterfactual and prevention assertions are represented in the same way as each

other, by a more explicit set of models corresponding to both the affirmative and the negative cases.

Inferences endorsed

Table 4.6 shows the percentages of participants who endorsed each of the inferences from each assertion.

Table 4.6: Percentages of participants making each of the inferences from each assertion in Experiment 9

Assertion type	Assertion valence	Inference	
		Modus ponens	Denial of the antecedent
Causal	Affirmative	71	68
	Negative	84	61
	Mean	77	65
Counterfactual	Affirmative	77	74
	Negative	77	55
	Mean	77	65
Prediction	Affirmative	74	81
	Negative	87	71
	Mean	82	76
Prevention	Affirmative	35	58
	Negative	65	58
	Mean	50	58
All assertions	Affirmative	65	70
	Negative	78	61
	Mean	71	66

As in Experiment 8, for the affirmative assertions, participants made as many modus ponens as denial of the antecedent inferences from counterfactual assertions (77% vs 74%, $n =$

31, binomial $p = 1$) and prevention assertions (35% vs 58%, $n = 31$, binomial $p = 0.12$). However, they also made as many modus ponens as denial of the antecedent inferences from causal assertions (71% vs 68%, $n = 31$, binomial $p = 1$) and prediction assertions (74% vs 81%, $n = 31$, binomial $p = 0.75$). The lack of differences in the frequencies of modus ponens and denial of the antecedent inferences from causal and prediction assertions removes the differences observed in the previous experiment, although of course, those differences were observed in the latencies of this experiment. However, when the assertions were negative, participants made more modus ponens than denial of the antecedent inferences from causal assertions and prediction assertions (although only the first difference is marginally significant, 84% vs 61%, $n = 31$, binomial $p = 0.09$; 87% vs 71%, $n = 31$, binomial $p = 0.23$, respectively). They made as many modus ponens as denial of the antecedent inferences from prevention assertions (65% vs 58%, $n = 31$, binomial $p = 0.75$), but they made somewhat more modus ponens than denial of the antecedent inferences from counterfactual assertions (77% vs 55%, $n = 31$, binomial $p = 0.06$).

Participants made as many modus ponens inferences from causal assertions as counterfactual (77% vs 77%, Wilcoxon's $z = 0$, $n = 31$, $p = 1$) and prediction assertions (77% vs 82%, Wilcoxon's $z = 0.63$, $n = 31$, $p = 0.53$), but they made more modus ponens inferences from causal assertions than prevention assertions (77% vs 50%, Wilcoxon's $z = 3.04$, $n = 31$, $p < 0.01$). They made as many modus ponens inferences from counterfactual assertions as prediction assertions (77% vs 82%, Wilcoxon's $z = 0.53$, $n = 31$, $p = 0.59$), but they made more modus ponens inferences from counterfactual assertions than prevention assertions (77% vs 50%, Wilcoxon's $z = 3.25$, $n = 31$, $p < 0.01$). They made more modus ponens inferences from prediction assertions than prevention assertions (82% vs 50%, Wilcoxon's $z = 3.51$, $n = 31$, $p <$

0.01)^{15, 16}.

Participants made as many denial of the antecedent inferences from causal assertions as counterfactual (65% vs 65%, Wilcoxon's $z = 0$, $n = 31$, $p = 1$), prediction (65% vs 76%, Wilcoxon's $z = 1.38$, $n = 31$, $p = 0.17$) or prevention assertions (65% vs 58%, Wilcoxon's $z = 0.8$, $n = 31$, $p = 0.42$). They made as many denial of the antecedent inferences from counterfactual assertions as prediction (65% vs 76%, Wilcoxon's $z = 1.58$, $n = 31$, $p = 0.11$) or prevention assertions (65% vs 58%, Wilcoxon's $z = 0.88$, $n = 31$, $p = 0.38$). They made more denial of the antecedent inferences from prediction assertions than prevention assertions (76% vs 58%, Wilcoxon's $z = 2.4$, $n = 31$, $p = 0.02$)¹⁷.

For affirmative assertions, participants made as many modus ponens as denial of the antecedent inferences from all four assertions; for negative assertions they made somewhat more modus ponens than denial of the antecedent inferences from causal and counterfactual assertions,

¹⁵ For affirmative assertions, participants made as many modus ponens inferences from causal as counterfactual (71% vs 77%, binomial $p = 0.75$) and prediction assertions (71% vs 74%, binomial $p = 1$), but they made more from causal than prevention assertions (71% vs 35%, binomial $p = 0.03$). They made as many from counterfactual as prediction assertion (77% vs 74%, binomial $p = 1$), but they made more from counterfactual than prevention assertions (77% vs 35%, binomial $p < 0.01$). They made more from prediction than prevention assertions (74% vs 35%, binomial $p = 0.01$). For negative assertions, they made as many from causal as counterfactual (84% vs 77%, binomial $p = 0.69$), prediction (84% vs 87%, binomial $p = 1$) or prevention assertions (84% vs 65%, binomial $p = 0.11$). They made as many from counterfactual as prediction (77% vs 87%, binomial $p = 0.37$) or prevention assertions (77% vs 65%, binomial $p = 0.34$). They made more from prediction than prevention assertions (87% vs 65%, binomial $p = 0.04$).

¹⁶ The particularly low rates of modus ponens inferences from the affirmative prevention assertions may result from the fact that they have negative consequents, e.g., "Putting fertiliser on the plants prevents them from not growing quickly". Since the verb prevents is implicitly negative (e.g., Goldvarg & Johnson-Laird, 2000), the negation of the consequent results in a double negative, which may be particularly difficult to understand (e.g., Evans et al., 1993).

¹⁷ For affirmative assertions, participants made as many denial of the antecedent inferences from causal as counterfactual (68% vs 74%, binomial $p = 0.73$), prediction (68% vs 81%, binomial $p = 0.39$) or prevention assertions (68% vs 58%, binomial $p = 0.61$). They made as many from counterfactual as prediction (74% vs 81%, binomial $p = 0.73$) or prevention assertions (74% vs 58%, binomial $p = 0.23$). They made more from prediction than prevention assertions (81% vs 58%, binomial $p = 0.04$). For negative assertions, they made as many from all four assertions (causal vs counterfactual: 61% vs 55%, binomial $p = 0.69$; causal vs prediction: 61% vs 71%, $p = 0.51$; causal vs prevention: 61% vs 58%, binomial $p = 1$; counterfactual vs prediction: 55% vs 71%, binomial $p = 0.18$; counterfactual vs prevention: 55% vs 58%, binomial $p = 1$; prediction vs prevention: 71% vs 58%, binomial $p = 0.34$).

and they made as many modus ponens as denial of the antecedent inferences from prediction and prevention assertions. Possible reasons for the differences in the relative frequencies of modus ponens and denial of the antecedent between the previous experiment and this one include the fact that the samples were drawn from university undergraduates in the former and the general population in the latter, the materials were presented in booklets in the former and on a computer in the latter, and all premises were available for inspection for each problem in the former, whereas the premises disappeared before the conclusion was presented in the latter. Nonetheless, the latencies recorded in this experiment provide a replication of the endorsements observed in the previous experiment and provide additional evidence for the suggested representational similarities and differences for counterfactual, causal, prediction and prevention assertions.

General discussion

The results of our final three experiments show that there are systematic similarities and differences in how people understand and reason from causal and prediction assertions on the one hand, and counterfactual and prevention assertions on the other. The findings provide further evidence for the view that causal thinking is directed towards prediction, whereas counterfactual thinking is directed towards prevention.

The results of Experiment 7 suggest that people often interpret a causal assertion such as “Exercise that was excessive caused the occurrence of angina” to refer to the same “biconditional” possibilities as a counterfactual assertion about the same factual situation such as “If exercise had not been excessive, then angina would not have occurred”. Both assertions are often taken to mean that it was possible for both exercise that was excessive and angina to occur, or for neither exercise that was excessive nor angina to occur, whereas it was impossible for

exercise that was excessive to occur and angina to not occur, or for exercise that was excessive to not occur and angina to occur. Likewise, a prediction assertion such as “If exercise was excessive, then angina occurred” and a prevention assertion such as “Exercise that was not excessive prevented the occurrence of angina” are also often taken to refer to these same two possibilities and no other ones.

Participants made this “biconditional” interpretation, which conveys information useful for both prediction and prevention, for just over half of the assertions. However, a sizable minority, almost a quarter, of the assertions were given a “conditional” interpretation, in which causal and counterfactual assertions about the same factual situation are distinguished by their reference to one possibility which they do not hold in common. On a conditional interpretation, a causal assertion may convey information useful for prediction but not prevention, whereas a counterfactual assertion may convey information useful for prevention but not prediction. We can conclude that at least some causal assertions are interpreted in a way which has implications only for prediction but not prevention, whereas some counterfactual assertions are interpreted in a way which has implications only for prevention. The function of causal thinking may be to determine how outcomes can be predicted, whereas the function of counterfactual thinking may be to determine how outcomes can be prevented.

An alternative view is that counterfactual thinking is a subcomponent of causal thinking: causal thoughts may identify antecedents which were necessary and sufficient for an outcome, and counterfactual thoughts may be used to test the necessity of possible causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). On this account, a causal assertion such as “Exercise that was excessive caused the occurrence of angina” should be interpreted as a biconditional (excessive exercise was necessary and sufficient for angina),

whereas a counterfactual assertion about the same factual situation such as “If exercise had not been excessive, then angina would not have occurred” should be interpreted as either a biconditional (excessive exercise was necessary and sufficient for angina) or a conditional (excessive exercise was necessary but not sufficient for angina). However, participants in Experiment 7 interpreted only 53% of causal assertions as biconditionals. They interpreted 25% as conditionals (excessive exercise was sufficient but not necessary for angina), but only 7% as reversed conditionals (excessive exercise was necessary but not sufficient for angina). These results suggest that people judge that causes are sufficient for their effects, but they need not be necessary (see also Mandel & Lehman, 1998; Chapter 1, p. 26). Moreover, participants made slightly less conditional interpretations of counterfactual assertions than causal assertions (19% vs 25%), which is inconsistent with the view that counterfactual thoughts focus on necessary causes more than causal thoughts (pace N’gbala & Branscombe, 1995, see Chapter 1, pp. 26-27). The results do not support the view that counterfactual thoughts test the necessity of potential necessary and sufficient causes.

Of course, asking people to judge which situations were possible from a list of explicitly presented situations does not reveal which possibilities *spontaneously* come to mind for different sorts of assertions. Therefore, the aim of our next two experiments was to examine indirectly the possibilities that initially come to mind for causal and counterfactual assertions, and for prediction and prevention assertions, by examining the sorts of inferences people make from the assertions. The results of Experiment 8 showed a similar pattern of inferences from causal and prediction assertions, which differed in systematic ways from the pattern of inferences from counterfactual and prevention assertions. For affirmative assertions, participants made more modus ponens than denial of the antecedent inferences from causal and prediction assertions, but

they made just as many modus ponens as denial of the antecedent inferences from counterfactual and prevention assertions. This difference did not occur for negative assertions, but we suggest that it may have been masked by the greater difficulty of understanding negative assertions (see e.g., Evans et al., 1993).

The results of Experiment 9 replicated this pattern for the time taken to make the inferences. For affirmative assertions, participants made modus ponens inferences more quickly than denial of the antecedent inferences from causal and prediction assertions, but they made modus ponens inferences at just the same speed as denial of the antecedent inferences from counterfactual and prevention assertions. Again, this difference did not occur for negative assertions, perhaps due to the greater resources required to process them. This pattern for the latencies exactly replicates the pattern for endorsements found in the previous experiment. It provides further corroboration for our suggestion that causal and prediction assertions are represented in the same way as each other, by an initial model corresponding to the affirmative case, which may convey information useful for prediction but not prevention, whereas counterfactual and prevention assertions are represented in the same way as each other, by a more explicit set of models corresponding to both the affirmative and the negative cases, which may convey information useful for prevention. The finding that people tend to initially represent causal and counterfactual assertions differently is also consistent with the view that counterfactual assertions convey information about the necessity of potential necessary and sufficient causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). However, this account would not predict the similarities we have demonstrated between causal and prediction assertions on the one hand, and counterfactual and prevention assertions on the other.

We have suggested that people understand and reason from causal and counterfactual assertions by constructing and manipulating mental models. Alternatively, reasoning may depend on the application of formal rules of inference to language-like representations (e.g., Braine & O'Brien, 1998; Rips, 1994) or the use of domain-specific reasoning schemas which are evoked by particular contexts (e.g., Cheng & Holyoak, 1985). However, neither of these theories would predict the specific similarities and differences we have found in people's inferences from causal, counterfactual, prediction and prevention assertions. In contrast, our findings are consistent with the core tenets of the mental model theory: people understand an assertion by constructing internal models which represent possibilities consistent with the assertion, people initially represent explicitly as few models as necessary to capture the meaning of the assertion, and inferences which require less models to be kept in mind and which can be made from the initial representation of an assertion are easier -- and therefore made more often and faster -- than inferences which require multiple models to be kept in mind and the initial models to be fleshed out (Johnson-Laird & Byrne, 1991). These principles not only account for our novel findings, they also account for many of the well-established phenomena of deductive reasoning (e.g., Johnson-Laird & Byrne, 1991; Johnson-Laird et al., 1992).

The three experiments reported in this chapter provide the first comparison of how people understand and reason from causal and counterfactual assertions. The results suggest that there are similarities and differences in how people interpret and represent causal and counterfactual assertions about the same factual situation. The differences are consistent with the view that causal and counterfactual thoughts tend to be concerned with different sorts of causal relations: causal thoughts are concerned with relations in which an antecedent predicts an outcome,

whereas counterfactual thoughts are concerned with relations in which an antecedent prevents an outcome.

Chapter 5 Discussion

At the end of Chapter 1, we stated the aims of this thesis. Our primary aim was to provide a systematic comparison of causal and counterfactual thinking. More specifically, we aimed to test the hypothesis that causal and counterfactual thinking serve different functions: causal thinking is concerned with the prediction of outcomes, whereas counterfactual thinking is concerned with their prevention (e.g., Mandel & Lehman, 1996; see Chapter 1, pp. 11-13). An alternative view is that counterfactual thinking is a subcomponent of causal thinking which tests the necessity of potential necessary and sufficient causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). We have now reported nine experiments which compared the determinants of causal and counterfactual thinking: in Chapter 2, we presented three experiments that compared the activation and content of causal and counterfactual thoughts, in Chapter 3, we described three further experiments that compared the consequences of causal and counterfactual thoughts, and in Chapter 4, we reported three final experiments that compared how people understand and reason from causal and counterfactual assertions. In this chapter, we will first summarise the findings of each set of experiments in turn, and show how they relate to our aims, second, discuss the implications of the overall findings for research and theory, and finally, draw some conclusions.

Summary of the findings

In this section, we summarise the main results of each of the three sets of experiments in turn and indicate the primary conclusions that can be drawn from them.

The activation and content of causal and counterfactual thoughts

In Chapter 2, we reported three experiments designed to examine the functions of causal and counterfactual thinking by comparing their activation and content. We reported two experiments which provided the first comparison of the activation and content of spontaneous causal and counterfactual thoughts (Experiments 1 and 2), and one experiment which provided the first direct test of the effects of two key variables (covariation and controllability) on the content of directed causal and counterfactual thoughts (Experiment 3).

In Experiment 1, we compared the effects of the valence (bad or good), the expectancy (unexpected or expected) and the controllability (controllable or uncontrollable) of an outcome on the frequencies of participants' spontaneous causal and counterfactual thoughts. The view that counterfactual thinking is a subcomponent of causal thinking suggests that causal and counterfactual thoughts serve the same ultimate aim and therefore should be evoked by the same sorts of outcomes. However, our findings demonstrate that there are systematic differences in the activation of causal and counterfactual thinking. The specific pattern of differences supports the idea that causal and counterfactual thinking serve different functions: causal thinking is concerned with prediction, whereas counterfactual thinking is concerned with prevention. First, causal and counterfactual thoughts were both activated by bad rather than good outcomes -- the aim of both prediction and prevention may be to avoid bad outcomes. However, the difference between bad and good outcomes tended to be greater for counterfactual thoughts than for causal thoughts -- it may be useful to predict both good and bad outcomes, whereas it is only necessary to prevent bad outcomes. Second, causal but not counterfactual thoughts were activated by unexpected rather than expected outcomes -- unexpected outcomes indicate that prediction has failed. Finally, counterfactual but not causal thoughts were activated by controllable rather than

uncontrollable outcomes -- controllable outcomes indicate that prevention is possible.

Experiment 2 ruled out the possibility that people generate more counterfactual thoughts about controllable than uncontrollable outcomes, not because counterfactual thoughts are concerned with prevention, but because it is cognitively easier to generate counterfactual alternatives to controllable antecedents. We tested this possibility by examining the effects of the normality of antecedents (unusual or normal) on the frequencies of participants' spontaneous causal and counterfactual thoughts. Despite the well-established finding that unusual antecedents are more mutable than normal antecedents (e.g., Kahneman & Miller, 1986; Kahneman & Tversky, 1982), participants generated as many *spontaneous* counterfactual thoughts about outcomes with unusual antecedents as normal ones. In contrast, participants engaged in more causal thinking about outcomes with unusual antecedents than normal ones, even though antecedent normality should not affect the cognitive ease of generating causal thoughts, because causal thinking does not require the mutation of antecedents. The results suggest that the frequency of spontaneous causal or counterfactual thoughts depends on their functions rather than their cognitive availability: participants may have generated more causal thoughts about unusual outcomes when they had unusual antecedents because unusual antecedents are better predictors of unusual outcomes than normal antecedents. In contrast, they may have generated as many counterfactual thoughts about outcomes with unusual antecedents as normal ones because unusual antecedents are no better preventors than normal antecedents.

This interpretation is also consistent with the content of the causal and counterfactual thoughts generated by participants in Experiments 1 and 3. In Experiment 1, we compared the general (internal or external) versus specific (inhibitory or facilitative) content of spontaneous causal and counterfactual thoughts. Causal thoughts focused on both general factors such as

effort and luck, and specific antecedents mentioned in the scenario, whereas counterfactual thoughts focused mostly on specific antecedents -- general factors may predict future outcomes in a broad range of situations, whereas specific antecedents may be more easily converted into plans for preventing future failures. Moreover, causal thoughts focused on both internal factors such as effort, and external factors such as luck, whereas counterfactual thoughts focused mostly on internal factors -- internal and external factors are equally relevant to prediction, but only internal factors are potentially controllable and therefore relevant to prevention. Furthermore, causal thoughts focused on antecedents which would facilitate a good outcome, whereas counterfactual thoughts focused on antecedents which would inhibit a good outcome -- a facilitative antecedent predicts a good outcome, whereas the mutation of an inhibitory antecedent prevents a bad outcome.

We further examined the content of causal and counterfactual thoughts in Experiment 3, by comparing systematically the effects of two key variables, covariation and controllability, on the content of participants' directed causal and counterfactual thoughts. We varied which of two antecedents was described as covarying with an outcome, and which of the two was described as controllable. Participants' causal and counterfactual thoughts both focused on the covarying rather than the non-covarying antecedent. Their causal thoughts also focused on the uncontrollable rather than the controllable antecedent, whereas their counterfactual thoughts focused on the controllable antecedent. Participants rated as most causal an uncontrollable antecedent which general knowledge would suggest covaries with the outcome, whereas they did not counterfactually mutate this antecedent any more than the others. Causal and counterfactual thoughts may both focus on covarying antecedents because the presence of a covarying antecedent helps predict an outcome, and the absence of a covarying antecedent helps prevent an

outcome. Causal but not counterfactual thoughts may focus on antecedents which general knowledge suggests covary with an outcome because they predict the outcome in a wide range of situations. Counterfactual but not causal thoughts may focus on controllable antecedents because controllable antecedents are better preventors than uncontrollable antecedents. A possible alternative explanation is that it is easier to generate counterfactual alternatives to controllable events, whereas it is not easier to generate causal thoughts focusing on controllable events. The contents of causal and counterfactual thoughts alone do not provide direct evidence for their functions because their contents may depend on cognitive as well as motivational factors. On this account, causal and counterfactual thinking may serve the same functions, even if they have different contents. We returned to this issue in the following chapter (Chapter 3), which compared the consequences of causal and counterfactual thoughts.

The findings of the first three experiments do not support the view that the primary function of counterfactual thinking is to test the necessity of potential necessary and sufficient causes. On this account, causal thoughts should focus on necessary and sufficient causes, whereas counterfactual thoughts may focus on necessary but insufficient causes, but otherwise causal and counterfactual thinking should have the same determinants. In contrast, the results show that there is a range of systematic differences in the activation and content of causal and counterfactual thoughts which are independent of necessity and sufficiency. The specific pattern of differences is consistent with the view that causal and counterfactual thinking serve different functions: causal thinking is directed towards prediction, whereas counterfactual thinking is directed towards prevention.

The consequences of causal and counterfactual thoughts

We further investigated the functions of causal and counterfactual thinking in the three experiments reported in Chapter 3, by comparing their consequences for judgements and emotions. We reported one experiment which provided the first comparison of the effects of causal and counterfactual thoughts on judgements of the preventability, controllability and predictability of an outcome (Experiment 4), a second experiment which provided the first examination of the influence of causal and counterfactual thinking on the hindsight bias for a bad outcome (Experiment 5), and a third experiment which provided the first comparison of the consequences of causal thoughts, upward counterfactual thoughts (about how an outcome could have been better) and downward counterfactual thoughts (about how an outcome could have been worse) for a range of judgements related to prevention and prediction, and for emotions (Experiment 6).

In Experiment 4, we compared the effects of causal and counterfactual thinking on participants' ratings of the preventability, controllability and predictability of a bad outcome. The results provide the first demonstration that counterfactual thoughts increase the perceived preventability of an outcome, and they support previous findings that counterfactual thoughts increase perceived control (e.g., McMullen et al., 1995). In contrast, causal thoughts had no effects on either perceived preventability or perceived control. Participants who generated counterfactual thoughts rated the outcome as more preventable than those who generated causal thoughts, consistent with the view that counterfactual thoughts are more concerned with prevention than causal thoughts. Neither counterfactual nor causal thoughts affected the perceived predictability of the outcome in Experiment 4, but a possible reason is that the outcome had a relatively low baseline level of perceived predictability.

We further examined the consequences of causal and counterfactual thinking for judgements related to prediction in Experiment 5, by comparing their effects on the hindsight bias, that is the tendency to overestimate the prior predictability of an outcome, once the outcome is known. We gave participants a scenario previously shown by Roeser and Olson (1996) to evoke the hindsight bias, and which also tended to evoke the hindsight bias in our participants. Participants who generated causal thoughts showed as strong a hindsight bias as controls, whereas those who generated counterfactual thoughts showed no hindsight bias. The findings support the view that causal thoughts are concerned with prediction, whereas counterfactual thoughts are concerned with prevention -- the hindsight bias seems to be a side effect of learning how an outcome can be predicted (e.g., Hawkins & Hastie, 1990; Wasserman et al., 1991), but if counterfactual thoughts show how an outcome can be prevented, they may decrease its perceived predictability, thereby eliminating the hindsight bias.

Experiment 6 extended our examination of the idea that causal thoughts are directed towards prediction, by comparing the consequences of causal and counterfactual thoughts for the perceived likelihood of a hypothetical *future* outcome as well as the perceived predictability of a past outcome. The experiment also further examined the idea that counterfactual thoughts are directed towards prevention, by comparing the consequences of causal and counterfactual thoughts for a range of judgements related to prevention, including perceived control, responsibility and blame, the preventability and controllability of a *future* outcome and intentions for future action. The experiment also examined the idea that causal and counterfactual thoughts have different emotional functions by comparing their consequences for emotions. We examined the role of direction in the consequences of counterfactual thoughts by instructing some participants to generate upward counterfactuals (about how an outcome could have been better),

and others to generate downward counterfactuals (about how an outcome could have been worse). We also examined the role of content in the consequences of causal and counterfactual thoughts by instructing some participants to generate internal thoughts (focusing on their own actions) and allowing others to generate general thoughts (focusing on whatever they wished).

First, the results support the view that causal thoughts are concerned with prediction -- causal thoughts tended to increase the perceived predictability of a past outcome compared to either upward or downward counterfactuals. Moreover, internal causal thoughts tended to increase the perceived likelihood of a better future outcome, whereas neither upward nor downward counterfactuals affected future likelihood judgements. Second, the findings support the view that upward counterfactual thoughts are concerned with prevention -- internal upward counterfactual thoughts increased judgements about prevention, whereas internal downward counterfactual thoughts had no effect. However, internal causal thoughts also increased judgements about prevention, which suggests that causal and counterfactual thoughts may have the same implications for prevention when they have the same contents, an issue which we returned to in the experiments reported in the following chapter (Chapter 4). Finally, the findings suggest that downward counterfactual thoughts are concerned with emotions rather than either prediction or prevention -- downward counterfactuals enhanced emotions, whereas causal thoughts and upward counterfactuals had no effects on emotions.

The findings reported in Chapter 3 do not support the view that the primary purpose of counterfactual thinking is to determine the necessity of potential causes. On this account, causal and counterfactual thoughts might have different consequences due to differences in their tendencies to focus on necessary and sufficient antecedents. However, the overall results of the three experiments suggest that causal and counterfactual thoughts tend to have different

consequences even when they both focus on antecedents which were individually necessary but insufficient for an outcome. The findings suggest that causal thoughts have consequences for prediction, upward counterfactual thoughts have consequences for prevention, and downward counterfactual thoughts have consequences for emotions.

Understanding and reasoning from causal and counterfactual assertions

We completed our investigation of the functions of causal and counterfactual thinking in Chapter 4, by reporting the results of three experiments which compared how people understand and reason from causal and counterfactual assertions with the same contents, which we compared to prediction assertions (phrased as indicative conditionals) and prevention assertions. We reported one experiment which provided the first examination of people’s interpretations of the meanings of the four sorts of assertions (Experiment 7), a second experiment which provided the first comparison of the sorts of inferences people endorse from the four sorts of assertions (Experiment 8), and a third experiment which investigated their latencies to endorse the inferences from the four sorts of assertions (Experiment 9).

The results of Experiment 7 suggest that people often interpret a causal assertion such as “Exercise that was excessive caused the occurrence of angina” to refer to the same possibilities as a counterfactual assertion about the same factual situation such as “If exercise had not been excessive, then angina would not have occurred”. Both assertions are often taken to mean that it was possible for both exercise that was excessive and angina to occur, or for neither exercise that was excessive nor angina to occur, whereas it was impossible for exercise that was excessive to occur and angina to not occur, or for exercise that was excessive to not occur and angina to occur. Likewise, a prediction assertion such as “If exercise was excessive, then angina occurred”

and a prevention assertion such as “Exercise that was not excessive prevented the occurrence of angina” are also often taken to refer to these same two possibilities and no other ones. This “biconditional” interpretation conveys information useful for both prediction and prevention.

Participants interpreted just over half of the causal and counterfactual assertions as biconditionals. However, they made “conditional” interpretations of a sizable minority of the assertions. On a conditional interpretation, causal and counterfactual assertions about the same factual situation are distinguished by their reference to one possibility which they do not hold in common. On a conditional interpretation, causal and prediction assertions both convey information useful for prediction but not prevention, whereas counterfactual and prevention assertions both convey information useful for prevention but not prediction. We can conclude that at least some causal assertions are interpreted in a way that conveys information about prediction but not prevention, whereas some counterfactual assertions are interpreted in a way that conveys information about prevention but not prediction. The findings do not support the view that counterfactual thoughts convey information about the necessity of potential necessary and sufficient causes. On this account, causal assertions should always be interpreted as biconditionals, whereas counterfactual assertions may be interpreted as either biconditionals or conditionals. In contrast, participants interpreted only 53% of causal assertions as biconditionals, and they made slightly more conditional interpretations of causal assertions than counterfactual assertions (25% vs 19%).

In Experiment 7, participants were asked to judge which situations had been possible from a list of explicitly presented situations, which may have forced them to think more explicitly about each situation than they would usually do. The aim of Experiments 8 and 9 was to examine the possibilities that *spontaneously* come to mind for causal and counterfactual

assertions, and for prediction and prevention assertions, by comparing the sorts of inferences people make from the four sorts of assertions. The results show a similar pattern of inferences from causal and prediction assertions, which differs in systematic ways from the pattern of inferences from counterfactual and prevention assertions. In Experiment 8, for affirmative assertions, participants made more affirmative (modus ponens) than negative (denial of the antecedent) inferences from causal and prediction assertions, but they made as many affirmative as negative inferences from counterfactual and prevention assertions. Experiment 9 replicated this pattern for the times taken to make the inferences. For affirmative assertions, participants made affirmative inferences faster than negative inferences from causal and prediction assertions, but they made affirmative and negative inferences at the same speed from counterfactual and prevention assertions.

The results of Experiments 8 and 9 suggest that causal and prediction assertions tend to be represented in the same way as each other, by an initial model corresponding to the affirmative case. This representation conveys information useful for prediction but not prevention. It represents explicitly only one situation in which the consequent occurs, and in that situation, the antecedent occurs, and so it suggests that the antecedent predicts the consequent. In contrast, it does not represent explicitly any situations in which the consequent does not occur, and so it conveys no information about how to prevent it. In contrast, counterfactual and prevention assertions tend to be represented in the same way as each other, by a more explicit set of models corresponding to both the affirmative and negative cases. Unlike the single initial model of causal and prediction assertions, this representation conveys information useful for prevention: the contrast between the affirmative and negative cases shows how the affirmative case can be prevented. The alternative view that counterfactual assertions convey information

about the necessity of potential necessary and sufficient causes would not predict the similarities we found between causal and prediction assertions on the one hand, and counterfactual and prevention assertions on the other.

The findings of our final three experiments suggest that there are similarities and differences in how people understand and represent causal and counterfactual assertions about the same factual situation. The differences are consistent with the view that causal and counterfactual thoughts tend to be concerned with different sorts of causal relations: causal thoughts are concerned with relations in which an antecedent predicts an outcome, whereas counterfactual thoughts are concerned with relations in which an antecedent prevents an outcome.

Implications

In this section, we consider the implications of the experiments, including their contributions to research, their limitations, their implications for the functions of causal and counterfactual thinking, and their implications for cognitive processes in causal and counterfactual thinking.

Contributions to research

Our experiments make an important contribution to research on causal and counterfactual thinking. Causal and counterfactual thinking have both been studied extensively in isolation from each other, but our experiments provide the first systematic *comparison* of the two. Only three previous studies have attempted to compare causal and counterfactual thoughts. Two studies analysed the content of participants' directed causal and counterfactual thoughts about fictional characters involved in car accidents and plane crashes (N'gbala & Branscombe, 1995; Mandel &

Lehman, 1996), and one study compared the consequences of causal and counterfactual thinking for the hindsight bias in football fans (Roese & Maniar, 1997). We provided the first comparison of the activation of causal and counterfactual thoughts (Experiments 1 and 2), the first investigation of the content of spontaneous causal and counterfactual thoughts (Experiment 1), and the first systematic examination of the content of directed causal and counterfactual thoughts (Experiment 3). We provided the first comparison of the effects of causal and counterfactual thoughts on judgements related to prevention (Experiments 4 and 6), the first comparison of their influence on judgements of the predictability of a bad outcome (Experiment 5), the first comparison of their consequences for predictions about a future outcome (Experiment 6), and the first comparison of their effects on emotions (Experiment 6). We provided the first comparison of people's understanding of causal and counterfactual assertions (Experiment 7), and the first comparison of inferences from causal and counterfactual assertions (Experiments 8 and 9).

We compared participants' causal and counterfactual thoughts using a variety of different dependent measures, including frequencies of spontaneous causal and counterfactual thoughts produced in a diary (Experiments 1 and 2), completions of causal or counterfactual sentence stems (Experiments 3, 4 and 5), ratings of the causal importance of antecedents (Experiment 3), listings of as many causal or counterfactual thoughts as come to mind (Experiment 6), judgements of whether situations were possible or impossible given causal and counterfactual assertions (Experiment 7), frequencies of inferences from causal and counterfactual assertions presented in a booklet (Experiment 8), and latencies of inferences from causal and counterfactual assertions presented on a computer (Experiment 9). We examined causal and counterfactual thinking about a range of different sorts of situations, including moving to a new town (Experiments 1 and 2), a car accident (Experiment 3), an academic failure (Experiment 4), a

historical scenario (Experiment 5), a recent negative life event (Experiment 6), and a variety of realistic causal relations (Experiments 7, 8 and 9). Moreover, we investigated causal and counterfactual thoughts about different sorts of outcomes, including bad and good, unexpected and expected, and controllable and uncontrollable outcomes (Experiment 1).

Limitations

Our experiments significantly extend the quantity and scope of past comparisons of causal and counterfactual thinking, but of course they do not provide anything near a complete comparison of causal and counterfactual thoughts. One limitation of our experiments is that they compared only a few specific attributes of causal and counterfactual thoughts. Our comparison of the activation of causal and counterfactual thoughts only examined the effects of outcome valence, outcome expectancy, outcome controllability and antecedent normality. Our comparison of the content of directed causal and counterfactual thoughts only investigated the effects of antecedent covariation and controllability. Our comparison of the consequences of causal and counterfactual thoughts only dealt with their consequences for judgements about prediction and prevention, and for emotions. Our comparison of how people understand and reason from causal and counterfactual assertions only examined which possibilities people believe the assertions refer to, and their tendencies to make the modus ponens and denial of the antecedent inferences from them. Further research is clearly needed in order to compare the many other aspects of causal and counterfactual thinking which we have not addressed.

In particular, we suggest that future studies should directly test the view that causal and counterfactual thoughts differ in their concern with necessary and sufficient causal relations (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12).

In this thesis, we have concentrated on systematically testing the view that causal and counterfactual thoughts differ in their concern with predictive and preventative causal relations (e.g., Mandel & Lehman, 1996; see Chapter 1, p. 12-13). We have already pointed out that several of our main results which support the “prediction and prevention” view cannot be easily explained by the “necessity and sufficiency” view. However, although our findings suggest that counterfactual thoughts are used to determine how unwanted outcomes could have been prevented, they do not refute the possibility that counterfactual thoughts may *also* be used to test the necessity of potential causes. The necessity and sufficiency view and the prediction and prevention view may not be incompatible. We suggest that experiments designed to pit the predictions of the two views against each other could provide stronger evidence regarding the relative validity of each account. For example, future studies may examine the effects of the perceived necessity and sufficiency of antecedents on the contents of causal and counterfactual thoughts. Some theorists have proposed that causal thoughts focus on sufficient causes, whereas counterfactual thoughts focus on necessary causes (N’gbala & Branscombe; see Chapter 1, pp. 26-27), but this suggestion has not been systematically tested.

A further limitation of our experiments is that they examined causal and counterfactual thoughts about only a few specific sorts of situations. The exception was Experiment 6, in which participants generated causal and counterfactual thoughts about a wide range of real-life events. In all of the other experiments, participants were required to think about particular situations. In Experiments 1 to 5, participants generated causal and counterfactual thoughts about imaginary scenarios. The scenarios were based on a range of different sorts of contexts, but only one scenario was used in each experiment, and so further research is needed to determine whether the findings of each experiment generalise to different sorts of scenarios. Moreover, further studies

are required to ensure that the findings generalise to real-life situations. However, it should be noted that past studies of counterfactual thinking about real-life events have in general supported the results of previous scenario studies (e.g., Davis et al., 1995; Gilovich & Medvec, 1994; Markman & Weary, 1996). Experiments 7 to 9 examined participants' representations of a limited range of causal and counterfactual assertions. Further research is needed to determine whether the findings generalise to people's representations of self-generated causal and counterfactual thoughts encompassing a broader range of contents. It should also be noted that our participants were recruited exclusively from student populations except in Experiment 9, in which the participants were members of the general public.

Finally, we acknowledge that some of the predicted findings were only marginally significant and there were some inconsistencies in the results of the experiments, perhaps as a result of differences in the materials used. For example, as expected, causal thinking resulted in higher predictability ratings than counterfactual thinking in Experiments 5 and 6. However, contrary to expectations, there were no differences in the predictability ratings of participants who generated causal and counterfactual thoughts in Experiment 4, perhaps due to the low baseline level of perceived predictability. Furthermore, as expected, there were differences in participants' tendencies to make the modus ponens and denial of the antecedent inferences from affirmative causal and prediction assertions in Experiments 8 and 9. However, contrary to expectations, these differences did not occur for negative assertions, perhaps due to the greater difficulty involved in processing negative assertions (e.g., Evans et al., 1993).

Despite the above limitations, the overall findings of our experiments allow us to draw some tentative conclusions about the functions of causal and counterfactual thinking and the

cognitive processes involved in causal and counterfactual thinking, and we will discuss these implications in the next two subsections.

Functions of causal and counterfactual thinking

Our overall findings suggest that causal and counterfactual thinking serve different functions: causal thoughts are concerned with prediction, upward counterfactuals are concerned with prevention and downward counterfactuals are concerned with emotions. We propose that these three functions may be united by the common goal of maintaining a sense of control.

In Chapter 1, we considered the evidence for similarities and differences in causal and counterfactual thinking. We concluded that causal and counterfactual thinking might be concerned with different sorts of causal relations, but this idea was highly speculative as only three previous studies had directly compared causal and counterfactual thoughts. The experiments we have reported in this thesis provide the first systematic comparison of causal and counterfactual thinking. The results show that there are specific patterns of similarities and differences in causal and counterfactual thoughts, and this finding is consistent with the view that they are concerned with different sorts of causal relations.

One view is that causal and counterfactual thinking differ in their concern with necessary and sufficient causal relations: causal thoughts may identify antecedents which were both necessary and sufficient to produce an outcome, and the function of counterfactual thoughts may be to test the necessity of hypothesised causes (e.g., Einhorn & Hogarth, 1986; Hilton, 1990; McGill & Klein, 1993; see Chapter 1, pp. 11-12). Our experiments were not designed to test this hypothesis, but our findings clearly contradict several of its predictions. For example, according to this view, causal and counterfactual thoughts should be activated in the same sorts of

situations. Causal and counterfactual thoughts should have the same sorts of contents, except that causal thoughts should focus on necessary and sufficient antecedents, whereas counterfactual thoughts may focus on necessary but insufficient antecedents. Any differences in the consequences of causal and counterfactual thoughts should depend on differences in their focus on necessary and sufficient antecedents. Causal assertions should be interpreted as biconditional relations, in which the antecedent is necessary and sufficient, whereas counterfactual assertions may be interpreted as conditional relations in which the factual antecedent is necessary but insufficient. None of these predictions is supported by our findings. In contrast, the results provide converging support for the view that causal and counterfactual thoughts differ in their concern with predictive and preventative causal relations (e.g., Mandel & Lehman, 1996; see Chapter 1, pp. 12-13).

Of course, prediction and prevention are not completely separate activities. The ability to predict which outcomes will occur as a result of different possible antecedents may also allow people to prevent bad outcomes. Likewise, understanding how a bad outcome could have been prevented may also allow people to better predict good and bad outcomes in the future. Prediction and prevention may complement each other (Mandel & Lehman, 1996), and this overlap may account for the fact that we found similarities as well as differences in causal and counterfactual thinking. For example, both causal and counterfactual thoughts are evoked by bad outcomes (Experiment 1), perhaps because the avoidance of bad outcomes requires both predictive and preventative knowledge. Moreover, both causal and counterfactual thoughts focus on antecedents which covary with an outcome (Experiment 3), perhaps because the presence of a covarying antecedent predicts an outcome, whereas the absence of a covarying antecedent prevents an outcome. Furthermore, both causal and counterfactual assertions are often

interpreted as biconditionals (Experiment 7), which may convey information about both prediction and prevention.

However, the pattern of differences we have found suggests that, despite this overlap, prediction and prevention are distinguished by their relations to causal and counterfactual thinking. For example, causal but *not* counterfactual thoughts are activated by unexpected (and therefore unpredicted) outcomes, and unusual (and therefore predictive) antecedents, whereas counterfactual but *not* causal thoughts are activated by controllable (and therefore preventable) outcomes (Experiments 1 and 2). Moreover, causal thoughts tend to increase judgements about prediction compared to counterfactual thoughts, whereas counterfactual thoughts tend to increase judgements about prevention compared to causal thoughts (Experiments 4, 5 and 6).

Furthermore, participants' initial mental models of causal and prediction assertions tend to represent explicitly only the affirmative case, which may convey information about prediction but not prevention, whereas their initial mental models of counterfactual and prevention assertions tend to represent explicitly both the affirmative and negative cases, which may convey information about prevention (Experiments 8 and 9).

The distinction we have made between predictive (causal) thinking and preventative (counterfactual) thinking is also consistent with findings from other research. For example, victims of traumatic events often spontaneously generate counterfactual thoughts about how they could have prevented the event without believing that they caused it (e.g., Davis et al., 1995). Furthermore, pre-school children spontaneously generate counterfactual alternatives more often in response to a question about how an outcome could have been prevented than to a question about why an outcome, which suggests that even young children can distinguish between predictive (causal) and preventative (counterfactual) thinking (Harris, German & Mills, 1996).

Of course, causal and counterfactual thinking may also serve functions other than prediction and prevention. For example, the results of Experiment 6 support previous suggestions that downward counterfactual thoughts serve an emotional function by helping people to feel better about unwanted outcomes (e.g., Markman et al., 1993; Roese, 1994). Upward counterfactual and causal thoughts had no effects on emotions in this experiment, which suggests that they are not concerned with emotions. However, upward counterfactual and causal thoughts may serve emotional functions in some circumstances. For example, participants who are directed to vividly imagine and focus on a counterfactual alternative, rather than compare it to reality, report more positive affect after generating upward counterfactuals and more negative affect after generating downward counterfactuals (McMullen, 1997). This finding suggests that upward counterfactuals can enhance emotions when they are “used for a purpose other than acting as a standard of comparison by which the past factual event is evaluated” (McMullen, 1997, p. 75). Causal thoughts may also enhance emotions in some circumstances, depending on their content. For example, people may feel better after attributing a bad outcome to external factors rather than their own actions (e.g., McFarland and Ross, 1982).

The predictive, preventative and emotional functions of causal and counterfactual thinking may all contribute to people’s more general aim of maintaining a sense of control over themselves and the world around them (e.g., Thompson, 1993). Although there are competing theories of *how* people generate causal thoughts (e.g., Einhorn & Hogarth, 1986), psychologists tend to agree that the ultimate function of causal thinking is to allow people to understand their environment so that they can predict *and control* future outcomes (e.g., Cheng & Novick, 1990; Heider, 1958; Kelly, 1972a). For example, according to Kelly (1972a, p. 22-23), “the purpose of causal analysis – the function it serves for the species and the individual – is effective control.

The attributor is not simply an attributor, a seeker after knowledge. His latent goal in gaining knowledge is that of effective management of himself and his environment”. Experiment 6 showed that both causal and counterfactual thoughts focusing on internal factors increase perceived control over an actual past outcome and a hypothetical future outcome. Moreover, even causal and counterfactual thoughts focusing on external factors may contribute to a more general sense of control, by allowing people to *understand* outcomes which challenge their assumption that unwanted outcomes are predictable and preventable (e.g., Davis & Lehman, 1995; Janoff-Bulman, 1992). Causal and upward counterfactual thoughts may serve distinct functions, prediction and prevention, but these functions may be united by the common goal of maintaining a sense of control.

Cognitive processes in causal and counterfactual thinking

Our experimental results suggest that causal and counterfactual thinking serve distinct functions. However, previous findings that counterfactual thoughts influence causal judgements suggest that there may be a close relation between the cognitive processes involved in causal and counterfactual thinking (e.g., Roese & Olson, 1996; Wells & Gavanski, 1989; see Chapter 1, p. 8). We suggest that causal and counterfactual thinking depend on the same sorts of underlying mental representations of causal relations, but causal assertions can be generated from the initial models of a causal relation, whereas counterfactual assertions require the initial mental models to be fleshed out to be more explicit.

People may understand causal relations by constructing models in their minds which represent the relevant possibilities (Johnson-Laird & Byrne, 1991; Goldvarg & Johnson-Laird, 2000). Consistent with this view, Experiment 7 showed that people often interpret a causal

relation to mean that two situations were possible: the factual situation in which both the antecedent and the consequent occurred, and the counterfactual situation in which neither the antecedent nor the consequent occurred. However, people's initial mental models of causal relations may represent *explicitly* only the factual situation, whereas other possibilities may be represented in an implicit form which may be fleshed out to be more explicit if necessary (Goldvarg & Johnson-Laird, 2000). Consistent with this view, the results of Experiments 8 and 9 suggest that people tend to initially represent causal assertions as a single possibility corresponding to the stated situation. However, the results indicate that people initially represent counterfactual assertions as a more explicit set of models corresponding to both the stated situation and its opposite. These findings suggest that people can produce causal assertions directly from their initial mental models of causal relations, which represent explicitly the factual situation in which the antecedent and consequent occurred. In contrast, in order to produce counterfactual assertions, people may have to first flesh out their initial mental models so that they also represent explicitly the counterfactual situation in which neither the antecedent nor the consequent occurred.

We suggest that people generate causal assertions directly from their initial mental models of causal relations, which represent explicitly only one possibility, whereas they generate counterfactual assertions by fleshing out their initial mental models so that they represent explicitly more than one possibility. This view suggests that it should be more difficult for people to generate counterfactual assertions than causal assertions -- people take longer to answer and they make more errors when a reasoning task requires them to flesh out implicitly represented information and keep multiple models in mind than when it can be performed using only an initial single explicit model (e.g., Johnson-Laird & Byrne, 1991). Consistent with this

view, young children can think about the causes of events before they can generate explicit counterfactual alternatives, perhaps because counterfactual thinking requires them to keep two possible situations in mind, which exceeds their working memory capacities (Robinson & Beck, 2000). Moreover, participants in our first two experiments produced about twice as many causal assertions as counterfactual assertions, perhaps because causal assertions are easier to generate. However, following controllable and unwanted outcomes, participants generated as many counterfactual assertions as causal assertions. We suggest that following controllable and unwanted outcomes, people are motivated to make the effort to flesh out their initial mental models of causal relations to include explicit counterfactual alternatives, because these alternatives suggest how such outcomes can be prevented in the future.

An alternative view is that causal thinking depends on prior counterfactual thinking (e.g., Lipe, 1991; Wells & Gavanski, 1989). Our experimental evidence suggests that people tend to engage in counterfactual thinking in order to determine how bad outcomes could have been prevented. However, people may *also* generate counterfactual thoughts in order to determine how good and bad outcomes were caused. Several philosophers have argued that people's most basic causal concepts depend on an awareness of counterfactual possibilities (e.g., Lewis, 1973a; 1973b; Mackie, 1974), and likewise, some psychologists have suggested that counterfactual information is central to causal thinking (e.g., Lipe, 1991; McGill & Klein, 1993; Wells & Gavanski, 1989). However, there is little evidence that people spontaneously think about counterfactual alternatives when they consider the cause of an outcome. For example, the availability of counterfactual thoughts has a greater effect on participants' causal judgements when they are *directed* to generate counterfactual thoughts before making the causal judgements, than when they make the causal judgements before being instructed to engage in counterfactual

thinking (Wells & Gavanski, 1989). Moreover, participants asked to generate questions they would like to have answered in order to assess a causal hypothesis produce very few counterfactual questions (Lipe, 1991).

In contrast, we suggest that counterfactual thoughts influence causal judgements, not because causal thoughts depend on prior counterfactual thoughts, but because counterfactual thoughts depend on causal information. For example, the counterfactual thought that Mr. Jones wouldn't have had a car accident if he had taken a different route depends on the knowledge that the *cause* of the accident was another person's driving, rather than Mr. Jones's driving. This view is consistent with some philosophical theories of the meaning of counterfactuals. For example, on one account, a counterfactual conditional asserts that there is a causal relation between its antecedent and its consequent, and so judgements about the truth of counterfactuals depend on knowledge of causal laws (e.g., Goodman, 1954/1983). Counterfactual thoughts may influence causal judgements because people need to bring to mind causal information in order to generate counterfactual assertions. Consistent with this view, Roese and Olson (1997) found that people answer a causal question faster after previously answering a related counterfactual question, but they do not answer the counterfactual question any faster after previously answering the causal question. Furthermore, the evidence that young children have a sophisticated understanding of causal relations before they are capable of engaging in explicit counterfactual reasoning suggests that counterfactual thinking may depend on causal knowledge, but causal thinking cannot depend on prior counterfactual thinking (e.g., Robinson & Beck, 2000).

But if the preparative benefits of counterfactual thinking are not a result of the role of counterfactual thoughts in determining the causes of past failures (pace Roese, 1997; Roese &

Olson, 1997), then why do people find counterfactual thinking so instructive? Tetlock and Belkin (1996, p. 4) conclude from their review of counterfactual argument in the social sciences that “counterfactual reasoning is a prerequisite to any form of learning from history”. Moreover, experimental evidence suggests that in everyday life, counterfactual thoughts help people to learn from personal mistakes and thereby improve their subsequent performance (e.g., Roese, 1994). We suggest that people learn from counterfactual thinking not because counterfactual thoughts determine causal judgements, but because counterfactual thoughts render explicit previously implicit causal knowledge (see also Kahneman, 1995). We propose that people may learn more from counterfactual thinking than from causal thinking because they can generate causal assertions directly from their initial mental models of causal relations, which show how the factual outcome came about, whereas they generate counterfactual assertions by fleshing out their initial mental models so that they *also* represent explicit counterfactual alternatives which show how the outcome can be prevented.

We have suggested that both causal and counterfactual thinking depend on mental models of causal relations, but of course people might represent their causal knowledge in a way which is quite different from the sorts of models we have described. However, our findings are consistent with the core tenets of the mental model theory (Johnson-Laird & Byrne, 1991): people understand an assertion by constructing internal models which represent possibilities consistent with the assertion (Experiment 7); people initially represent explicitly as few models as necessary to capture the meaning of the assertion (Experiments 8 and 9); and inferences which require less models to be kept in mind and which can be made from the initial representation of an assertion are easier -- and therefore made more often and faster -- than inferences which require multiple models to be kept in mind and the initial models to be fleshed out (Experiments

8 and 9). These principles not only account for our novel findings, they also account for many of the well-established phenomena of deductive reasoning (e.g., Johnson-Laird & Byrne, 1991; Johnson-Laird et al., 1992). Moreover, the specific patterns of inferences participants made from causal, counterfactual, prediction and prevention assertions in Experiments 8 and 9 would not be predicted by the competing rule-based theories of reasoning (e.g., Braine & O'Brien, 1998; Rips, 1994).

The mental model theory was originally proposed to account for deductive reasoning and language comprehension (Johnson-Laird, 1983), but it has been extended to provide a unified account of many other kinds of thinking (for a review see Oakhill & Garnham, 1996). It has been applied to how people think about causal relations, which may be represented in models of factual and counterfactual possibilities (e.g., Goldvarg & Johnson-Laird, 2000; Johnson-Laird & Byrne, 1991). It has also been extended to provide an account of how people generate counterfactual thoughts by altering aspects of their initial mental models of a situation (e.g., Byrne, 1996; 1997), which has been experimentally tested in the domains of the temporal order effect (Byrne et al., 2000) and the action effect (Byrne & McEleney, 1997; 2000). The accounts of causal and counterfactual thinking based on the mental model theory are recent developments, and additional research is needed to fully examine their assumptions. However, these accounts are consistent with the findings of our experiments, which provide the first systematic comparison of causal and counterfactual thinking. We suggest that the mental model theory provides a useful framework for further comparisons of causal and counterfactual thinking, which may provide more direct tests of its predictions.

Conclusions

Our experiments provide the first step towards systematically comparing causal and counterfactual thinking. The results show that there are systematic similarities and differences in the activation, content and consequences of causal and counterfactual thoughts, and in how people understand and reason from causal and counterfactual assertions. The findings suggest that causal thoughts are concerned with prediction, whereas counterfactual thoughts are concerned with prevention. We suggest that people may generate causal assertions directly from their initial mental models of causal relations, whereas they may generate counterfactual assertions by fleshing out their initial mental models so that they represent explicit counterfactual alternatives which show how an outcome can be prevented. We propose that what people gain from counterfactual thinking is not knowledge of the cause of an outcome, but a greater awareness of how they can intervene to prevent future failures. Causal and counterfactual thinking may be distinct activities that serve complementary functions, prediction and prevention, which may be united in the common goal of maintaining a sense of control over past and future outcomes.

References

- Abelson, R. P. & Black, J. B. (1986). Introduction. In J. A. Galambos, R. P. Abelson and J. B. Black (Eds.), *Knowledge structures*. Hillsdale, NJ: Erlbaum.
- Abramson, L. Y., Seligman, M. E. P. & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, 87, 49-74.
- Alloy, L. B. & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, 91, 1, 112-149.
- Anderson, C. A. (1983). Motivational and performance deficits in interpersonal settings: The effects of attributional style. *Journal of Personality and Social Psychology*, 45, 1136-1147.
- Au, T. K. (1984). Counterfactuals: In reply to Alfred Bloom. *Cognition*, 17, 289-302.
- Averill, J. A. (1982). *Anger and aggression*. New York: Springer-Verlag.
- Baker, A. G., Murphy, R. A. & Vallée-Tourangeau, F. (1996). Associative and normative accounts of causal induction: Reacting to versus understanding cause. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The Psychology of Learning and Motivation*, Vol. 34 (pp. 1-45). San Diego: Academic Press.
- Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.
- Berofsky, B. (1973). The counterfactual analysis of causation. *Journal of Philosophy*, 70, 568-569.

- Bettman, J. R. & Weitz, B. A. (1983). Attributions in the boardroom: Causal reasoning in corporate annual reports. *Administrative Science Quarterly*, 28, 165-183.
- Bohner, G., Bless, N., Schwarz, N. & Strack, F. (1988). What triggers causal attributions? The impact of valence and subjective probability. *European Journal of Social Psychology*, 18, 335-345.
- Bouts, P., Spears, R. & van der Pligt, J. (1992). Counterfactual processing and the correspondence between events and outcomes: Normality vs. value. *European Journal of Social Psychology*, 22, 387-96.
- Bower, G. H. & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, 247, 44-48.
- Bowerman, M. (1986). First steps in acquiring conditionals. In E. C. Traugott, A. ter Meulen, J. S. Reilly & C. A. Ferguson (Eds.), *On conditionals* (pp. 285-307). Cambridge, UK: Cambridge University Press.
- Braine, M. D. S. & O'Brien, D. P. (Eds.) (1998). *Mental Logic*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Branscombe, N. R., N'gbala, A., Kobryniewicz, D. & Wann, D. L. (1997). Self and group protection concerns influence attributions but they are not determinants of counterfactual mutation focus. *British Journal of Social Psychology*, 36, 387-404.
- Branscombe, N. R., Owen, S., Garstka, T. A. & Coleman, J. (1996). Rape and accident counterfactuals: Who might have done otherwise and would it have changed the outcome? *Journal of Applied Social Psychology*, 26, 12, 1042-1067.
- Bullock, M., Gelman, R. & Baillargeon, R. (1982). The development of causal reasoning. In W.

J. Friedman (Ed.), *The Developmental Psychology of Time* (pp. 209-254). Orlando, FL: Academic Press.

Bulman, R. J. & Wortman, C. B. (1977). Attributions of blame and coping in the "real world": Severe accident victims react to their lot. *Journal of Personality and Social Psychology*, 35, 351-363.

Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31, 61-83.

Byrne, R. M. J. (1996). Towards a model theory of imaginary thinking. In J. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science: Essays in honour of Phil Johnson-Laird* (pp. 155-174). Hove, UK: Psychology Press.

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In D. L. Medin (Ed.), *The Psychology of Learning and Motivation*, Vol. 37 (pp. 105-154). San Diego, CA: Academic Press.

Byrne, R. M. J., Culhane, R., & Tasso, A. (1995). The temporality effect in thinking about what might have been. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 385-390). Hillsdale, NJ: Lawrence Erlbaum Associates.

Byrne, R. M. J., Espino, O. & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language*, 40, 347-373.

Byrne, R. M. J. & McEleney, A. (1997). Cognitive processes in regret for actions and inactions. In M. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 73-78). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Byrne, R. M. J. & McEleney, A. (2000). Counterfactual thinking about actions and failures to act. *Journal of Experimental Psychology: Learning, Memory and Cognition* (in press).
- Byrne, R. M. J., Segura, S., Culhane, R., Tasso, A. & Berrocal, P. (2000). The temporality effect in thinking about what might have been. *Memory & Cognition*, 28, 2, 264-281.
- Byrne, R. M. J. & Tasso, A. (1994). Counterfactual reasoning: inferences from hypothetical conditionals. In A. Ram & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 124-129). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Byrne, R. M. J. & Tasso, A. (1999). Deductive reasoning from factual, possible and counterfactual conditionals. *Memory and Cognition*, 27, 726-740.
- Carpenter, P. (1973). Extracting information from counterfactual clauses. *Journal of Verbal Learning and Verbal Behaviour*, 12, 512-521.
- Carroll, J. S. & Payne, J. W. (1976). The psychology of parole decision process: A joint application of attribution theory and information processing psychology. In J. S. Carroll and J. W. Payne (Eds.), *Cognition and social behaviour* (pp. 13-32). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carroll, J. S. & Payne, J. W. (1977). Crime seriousness, recidivism risk and causal attributions in judgments of prison term by students and experts. *Journal of Applied Psychology*, 62, 595-602.
- Carroll, J. S., Perkowitz, W. T., Lurigio, A. J. & Weaver, F. M. (1987). Sentencing goals, causal attributions, ideology and personality. *Journal of Personality and Social Psychology*, 52, 107-118.

- Chan, D. & Chua, F. (1994). Suppression of valid inferences: syntactic views, mental models, and relative salience. *Cognition*, 53, 217-238.
- Chapman, L. J. & Chapman, J. P. (1967). Genesis of popular but erroneous diagnostic observations. *Journal of Abnormal Psychology*, 72, 193-204.
- Cheng, P. W. (1993). Separating causal laws from causal facts: pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The Psychology of Learning and Motivation* (Vol. 30, pp. 215-264). New York: Academic Press.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104, 367-405.
- Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.
- Cheng, P. W. & Nisbett, R. E. (1993). Pragmatic constraints on causal deduction. In R. E. Nisbett (Ed.), *Rules for Reasoning* (pp. 207-227). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cheng, P. W. & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, 58, 545-567.
- Cheng, P. W. & Novick, L. R. (1991). Causes versus enabling conditions. *Cognition*, 40, 83-120.
- Chisholm, R. M. (1946). The contrary-to-fact conditional. *Mind*, 55, 289-307.
- Christensen-Szalanski, J. J. & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behaviour and Human Decision Processes*, 48, 147-168.

- Clary, E. G. & Tesser, A. (1983). Reactions to unexpected events: The naïve scientist. *Personality and Social Psychology Bulletin*, 9, 609-620.
- Copi, I. M. (1982). *Introduction to logic*. New York: Macmillan Publishing.
- Corrigan, R. & Denton, P. (1996). Causal understanding as a developmental primitive. *Developmental Review*, 16, 162-202.
- Cummins, D. D. (1995). Naïve theories and causal deduction. *Memory and Cognition*, 23, 5, 646-658.
- Davis, C. G. & Lehman, D. R. (1995). Counterfactual thinking and coping with traumatic life events. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 355-74). Mahwah, NJ: Lawrence Erlbaum Associates.
- Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C. & Ellard, J. H. (1996). Self-blame following a traumatic event: The role of perceived avoidability. *Personality and Social Psychology Bulletin*, 22, 557-567.
- Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C. & Thompson, S. C. (1995). The undoing of traumatic life events. *Personality and Social Psychology Bulletin*, 21, 109-124.
- Dodge, K. A. & Crick, N. R. (1990). Social information processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, 16, 8-22.
- Ducasse, C. J. (1968). *Truth, knowledge and causation*. New York: Humanities Press.

- Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, 31, 674-685.
- Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 3-19.
- Eiser, J. R., Van der Plight, J., Raw, M. & Sutton, S. R. (1985). Trying to stop smoking: Effects of perceived addiction, attributions for failure and expectancy of success. *Journal of Behavioural Medicine*, 8, 4, 321-341.
- Evans, J. St. B. T., Newstead, S. E. & Byrne, R. M. J. (1993). *Human Reasoning: The Psychology of Deduction*. Hove, UK: Lawrence Erlbaum Associates.
- Fiedler, K. & Semin, G. R. (1988). On the causal information conveyed by different interpersonal verbs: The role of implicit sentence context. *Social Cognition*, 6, 21-39.
- Fillenbaum, S. (1974). Information amplified: Memory for counterfactual conditionals. *Journal of Experimental Social Psychology*, 102, 1, 44-49.
- Fischhoff, B. (1975). Hindsight $\neq$ foresight: The effect of outcome knowledge on judgement under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288-299.
- Fiske, S. T. & Taylor, S. E. (1991). *Social cognition*. Reading, MA: Addison-Wesley.
- Foersterling, F. (1985). Attributional retraining: A review. *Psychological Bulletin*, 98, 495-512.
- Foersterling, F. & Groenvald, A. (1983). Ursachzuschreibungen fur ein Wahlergebnis: Eine Überprüfung von Hypothesen der Attributionstheorie in einer Feldstudie anhand die niedersächsischen Kommunalwahl [Attributions for election results: A study of

attributional hypotheses in a field study of Lower Saxony community elections].

Zeitschrift für Socialpsychologie, 14, 262-269.

Gavanski, I. & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314-25.

Gelman, S. & Kalish, C. (1993). Categories and causality. In R. Pasnak & M. Howe (Eds.), *Emerging themes in cognitive development, Vol. II: Competencies* (pp. 3-32). New York: Springer-Verlag.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology*, 44, 1110-1126.

Gilovich, T., & Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology*, 67, 357-365.

Gioia, D. A. & Sims, H. P. (1986). Cognition-behaviour connections: Attribution and verbal behaviour in leader-subordinate interactions. *Organizational Behavior and Human Decision Processes*, 37, 197-229.

Giroto, V., Legrenzi, P. & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78, 111-133.

Gleicher, F., Boninger, D. S., Strathman, A., Armor, D., Hetts, J. & Ahn, M. (1995). With an eye toward the future: The impact of counterfactual thinking on affect, attitudes and behaviour. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 283-304). Mahwah, NJ: Lawrence Erlbaum Associates.

Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A. & Sherman, S. J.

(1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin*, 16, 2, 284-295.

Goldvarg, Y. & Johnson-Laird, P. N. (1999). Reasoning with causal relations. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society* (pp. 191-195). Mahwah, NJ: Lawrence Erlbaum Associates.

Goldvarg, Y. & Johnson-Laird, P. N. (2000). Naive causality: a mental model theory of causal meaning and reasoning. *Manuscript submitted for publication*.

Goodman, N. (1983). *Fact, fiction and forecast*. Cambridge, MA: Harvard University Press.
(original work published 1954)

Goswami, U. (1992). *Analogical reasoning in children*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Harré, R. & Madden, E. H. (1975). *Causal powers: A theory of natural necessity*. Oxford: Basil Blackwell.

Harris, P. L., German, T. & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61, 233-259.

Hart, H. L. A. & Honoré, A. M. (1985). *Causation and the law*. Oxford: Clarendon Press.
(Original work published 1959)

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, 46, 44-56.

Hawkins, S. A. & Hastie, R. (1990). Hindsight: Biased judgements of past events after the outcomes are known. *Psychological Bulletin*, 107, 311-327.

- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hewstone, M. & Jaspers, J. M. F. (1987). Covariation and causal attribution: A logical model of the intuitive analysis of variance. *Journal of Personality and Social Psychology*, 53, 663-672.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1-46). San Diego, CA: Academic Press.
- Hilton, D. J. (1990). Conversational processes and causal explanations. *Psychological Bulletin*, 107, 65-81.
- Hilton, D. J. & Erb, H. (1996). Mental models and causal explanation: Judgements of probable cause and explanatory relevance. *Thinking and Reasoning*, 2, 4, 273-308.
- Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, 93, 136-153.
- Hume (1978). *A treatise of human nature*. Oxford: Oxford University Press. (Original work published 1739)
- Hume (1999). *An enquiry concerning human understanding*. Oxford: Oxford University Press. (original work published 1748)
- Janoff-Bulman, R. (1992). *Shattered assumptions: Towards a new psychology of trauma*. New York: Free Press.
- Jaspers, J. M. F., Hewstone, M. R. C. & Fincham, F. D. (1983). Attribution theory and research: The state of the art. In J. M. F. Jaspers, F. D. Fincham & M. R. C. Hewstone (Eds.),

Attribution theory: Conceptual, developmental and social dimensions (pp. 1-36). San Diego, CA: Academic Press.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction*. Hove, UK, and Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. & Byrne, R. M. J. (2000). Conditionals: a theory of meaning, pragmatics and inference. *Manuscript submitted for publication*.

Johnson-Laird, P. N., Byrne, R. M. J. & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99, 418-439.

Jones, E. E. & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 2 (pp. 219-266). New York: Academic Press.

Jones, E. E. & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behaviour. In E. E. Jones, D. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behaviour* (pp. 11-26). Morristown, NJ: General Learning Press.

Kahneman, D. (1995). Varieties of counterfactual thinking. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 375-396). Mahwah, NJ: Lawrence Erlbaum Associates.

Kahneman, D. & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136-153.

Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208). New York: Cambridge University Press.

Kahneman, D. & Varey, C. A. (1990). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology*, 59, 1101-1110.

Kanazawa, S. (1992). Outcome or expectancy? Antecedents of spontaneous causal attribution. *Personality and Social Psychology Bulletin*, 18, 6, 659-668.

Kant, I. (1929). *Critique of pure reason*. (N. K. Smith, Trans.). London: MacMillan. (Original work published 1781)

Kassin, S. M. (1979). Consensus information, prediction and causal attribution: a review of the literature and issues. *Journal of Personality and Social Psychology*, 37, 1966-1981.

Keane, M. T. (1997). What makes an analogy difficult?: The effects of order and causal structure in analogical mapping. *Journal of Experimental Psychology: Language, Memory and Cognition*, 23, 946-967.

Kelly, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation, 1967* (pp. 192-238). Lincoln, NE: University of Nebraska Press.

Kelly, H. H. (1972a). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behaviour* (pp. 1-26). Morristown, NJ: General Learning Press.

Kelly, H. H. (1972b). Causal schemata and the attribution process. In E. E. Jones, D. E.

Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behaviour* (pp. 151-174). Morristown, NJ: General Learning Press.

Kelly, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28, 107-128.

Kelly, H. H. & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, 31, 457-503.

Kernis, M. H., Zucherman, M. & McVay, E. (1988). Motivational factors affecting performance: The impact of perceived locus of causality. *Personality and Social Psychology Bulletin*, 14, 3, 524-535.

Kim, J. (1973). Causes and counterfactuals. *Journal of Philosophy*, 70, 570-572.

Klauer, K. C. & Migulla, G. (1995). Spontanes kontrafaktisches Denken [Spontaneous counterfactual processing]. *Zeitschrift für Sozialpsychologie*, 26, 34-42.

Klauer, K. C., Jacobsen, T. & Migulla, G. (1995). Counterfactual processing: A test of a hierarchical correspondence model. *European Journal of Social Psychology*, 25, 577-595.

Knight, R. T. & Grabowecky, M. (1995). Escape from linear time: Prefrontal cortex and conscious experience. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1357-1371). Cambridge, MA: MIT Press.

Kun, A. (1978). Evidence for preschoolers' understanding of causal direction in extended causal sequences. *Child Development*, 49, 218-222.

Landman, J. (1987). Regret and elation following action and inaction: Affective responses to

positive versus negative outcomes. *Personality and Social Psychology Bulletin*, 13, 524-536.

Landman, J. & Manis, J. D. (1992). What might have been: Counterfactual thought concerning personal decisions. *British Journal of Psychology*, 83, 473-477.

Lau, R. R. (1984). Dynamics of the attribution process. *Journal of Personality and Social Psychology*, 46, 1017-1028.

Lau, R. R. & Russel, D. (1980). Attributions in the sports pages: A field test of some current hypotheses in attribution research. *Journal of Personality and Social Psychology*, 39, 29-38.

Leddo, J. & Abelson, R. O. (1986). The nature of explanations. In J. A. Galambos, R. P. Abelson and J. B. Black (Eds.), *Knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Legrenzi, P., Girotto, V. & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision-making. *Cognition*, 49, 37-66.

Lewis, D. (1973a). Causation. *Journal of Philosophy*, 70, 556-67.

Lewis, D. (1973b). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, D. (1986). *On the plurality of worlds*. New York: Basil Blackwell.

Lipe, M. G. (1991). Counterfactual reasoning as a framework for attribution theories. *Psychological Bulletin*, 108, 3-18.

Lombard, L. B. (1990). Causes, enablers and the counterfactual analysis. *Philosophical Studies*,

59, 195-211.

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. London: Oxford University Press.

Macrae, C. N. & Milne, A. B. (1992). A curry for your thoughts: Empathic effects on counterfactual thinking. *Personality and Social Psychology Bulletin*, 18, 5, 625-630.

Mandel, D. R. & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, 71, 3, 450-463.

Mandel, D. R. & Lehman, D. R. (1998). Integration of contingency information in judgement of cause, covariation and probability. *Journal of Experimental Psychology: General*, 127, 269-285.

Markman, K. D. & Weary, G. (1996). The influence of chronic control concerns on counterfactual thought. *Social Cognition*, 14, 292-316.

Markman, K. D., Gavanski, I., Sherman, S. J. & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, 29, 87-109.

Markman, K. D., Gavanski, I., Sherman, S. J. & McMullen, M. N. (1995). The impact of perceived control on the imagination of better and worse possible worlds. *Personality and Social Psychology Bulletin*, 21, 588-595.

McArthur, L. Z. (1972). The how and what of why: Some determinants and consequences of causal attributions. *Journal of Personality and Social Psychology*, 22, 171-188.

McArthur, L. Z. & Post, D. L. (1977). Figural emphasis and person perception. *Journal of*

Experimental Social Psychology, 13, 520-535.

- McCloy, R. (2000). *Thinking about what might have been: Cognitive processes in counterfactual and semifactual thinking about controllable events*. Ph.D. thesis: University of Dublin, Ireland.
- McCloy, R. & Byrne, R. M. J. (1999). Thinking about what might have been: If only, even if, causality and emotions. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society* (pp. 373-377). Mahwah, NJ: Lawrence Erlbaum Associates.
- McCloy, R. & Byrne, R. M. J. (2000). Cognitive processes in counterfactual thinking about controllable events. *Memory & Cognition* (in press).
- McFarland, C. & Ross, M. (1982). Impact of causal attributions on affective reactions to success and failure. *Journal of Personality and Social Psychology*, 43, 937-946.
- McGill, A. L. (1989). Context effects in judgments of causation. *Journal of Personality and Social Psychology*, 57, 189-200.
- McGill, A. L., & Klein, J. G. (1993). Contrastive and counterfactual thinking in causal judgment. *Journal of Personality and Social Psychology*, 64, 897-905.
- McMullen, M. N. (1997). Affective contrast and assimilation in counterfactual thinking. *Journal of Experimental Social Psychology*, 33, 77-100.
- McMullen, M. N., Markman, K. D. & Gavanski, I. (1995). Living in neither the best nor the worst of all possible worlds: Antecedents and consequences of upward and downward counterfactual thinking. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 133-167). Mahwah, NJ: Lawrence Erlbaum Associates.

Meyers-Levy, J. & Maheswaran, D. (1992). When timing matters: The influence of temporal distance on consumers' affective and persuasive responses. *Journal of Consumer Research*, 19, 424-433.

Michotte, A. (1946). *La perception de la causalité* [The perception of causality]. Paris: Vrin.

Mill, J. S. (1996). *A system of logic raciocentive and inductive*. London: Routledge. (Original work published 1843)

Miller, D. T. & Gunasegaram, S. (1990). Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology*, 59, 1111-1118.

Miller, D. T. & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, 12, 513-519.

Miller, G. & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge: Cambridge University Press.

Moreno-Ríos, S. M., García-Madruga, J. A. & Byrne, R. M. J. (2000). Semifactual and counterfactual conditionals. *Manuscript submitted for publication*.

Morris, M. W. & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, 102, 2, 331-335.

Musashi, M. (1974). *A book of five rings*. Woodstock, NY: Overlook. (Original work written 1645)

N'gbala, A. & Branscombe, N. R. (1995). Mental simulation and causal attribution: When

simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, 31, 139-162.

N'gbala, A. & Branscombe, N. R. (1997). When does action elicit more regret than inaction and is counterfactual mutation the mediator of this effect? *Journal of Experimental Social Psychology*, 33, 324-343.

Nasco, S. A. & Marsh, K. L. (1999). Gaining control through counterfactual thinking. *Personality and Social Psychology Bulletin*, 25, 5, 556-568.

Niedenthal, P. M., Tangney, J. P. & Gavanski, I. (1994). 'If only I weren't' versus 'If only I hadn't': Distinguishing shame from guilt in counterfactual thinking. *Journal of Personality and Social Psychology*, 67, 4, 585-95.

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Oakhill, J. & Garnham, A. (Eds.) (1996). *Mental models in cognitive science: Essays in honour of Phil Johnson-Laird*. Hove, UK: Psychology Press.

Pyszczynski, T. A. & Greenberg, J. (1981). Role of disconfirmed expectancies in the instigation of attributional processing. *Journal of Personality and Social Psychology*, 40, 31-38.

Regan, D. T., Strauss, E. & Fazio, R. H. (1974). Liking and the attribution process. *Journal of Experimental Social Psychology*, 10, 385-397.

Reichenbach, H. (1956). *The Direction of Time*. Berkeley, CA: University of California Press.

Rescher, N. (1995). Thought experimentation in Presocratic philosophy. In N. Rescher (Ed.), *Essays in the history of philosophy* (pp. 27-37). Adershot, UK: Avebury.

- Rips, L. J. (1994). *The Psychology of Proof*. Cambridge, MA: MIT Press.
- Ritov, I. & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioural Decision Making*, 3, 263-277.
- Robinson, E. J. & Beck, S. (2000). What is difficult about counterfactual reasoning? In P. Mitchell & K. J. Riggs (Eds.), *Children's Reasoning and the Mind* (pp. 101-119). Hove, UK: Psychology Press.
- Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, 66, 5, 805-818.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133-148.
- Roese, N. J. & Hur, T. (1997). Affective determinants of counterfactual thinking. *Social Cognition*, 15, 4, 274-290.
- Roese, N. J., Hur, T. & Pennington, G. L. (1999). Counterfactual thinking and regulatory focus: Implications for action versus inaction and sufficiency versus necessity. *Journal of Personality and Social Psychology*, 77, 6, 1109-1120.
- Roese, N. J. & Maniar, S. D. (1997). Perceptions of purple: Counterfactual and hindsight judgments at Northwestern Wildcats football games. *Personality and Social Psychology Bulletin*, 23, 1245-1253.
- Roese, N. J. & Olson, J. M. (1993). The structure of counterfactual thought. *Personality and Social Psychology Bulletin*, 19, 3, 312-319.
- Roese, N. J. & Olson, J. M. (1995). Counterfactual thinking: A critical overview. In N. J. Roese

and J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 1-59). Mahwah, NJ: Lawrence Erlbaum Associates.

Roese, N. J. & Olson, J. M. (1996). Counterfactuals, causal attributions and the hindsight bias. *Journal of Experimental Social Psychology*, 32, 197-227.

Roese, N. J. & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 29 (pp. 1-59). San Diego, CA: Academic Press.

Ross, L. (1977). The intuitive scientist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 10 (pp. 174-220). New York: Academic Press.

Sanna, L. J. & Turley, K. J. (1996). Antecedents to spontaneous counterfactual thinking: Effects of expectancy violation and outcome valence. *Personality and Social Psychology Bulletin*, 22, 9, 909-919.

Schachter, S. & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review*, 69, 379-399.

Schkade, D. A. & Kilbourne, L. M. (1991). Expectancy-outcome consistency and hindsight bias. *Organisational Behaviour and Human Decision Processes*, 49, 105-123.

Schoeneman, T. J., van Uchelen, C., Stonebrink, S. & Cheek, P. R. (1986). Expectancy, outcome and event type: Effects on retrospective reports of attributional activity. *Personality and Social Psychology Bulletin*, 12, 353-362.

Shanks, D. R. (1989). Selectional processes in causality judgment. *Memory & Cognition*, 17, 27-34.

- Sherman, S. J. (1991). Thought systems for the past as well as the future. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Advances in social cognition*, Vol. 4 (pp. 173-195). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47, 1-51.
- Shustack, M. W. & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Siegler, R. S. & Liebert, R. M. (1974). Effects of contiguity, regularity and age on children's causal inference. *Developmental Psychology*, 10, 574-579.
- Sim, D. L. H. & Morris, M. (1998). Representativeness and counterfactual thinking: The principle that antecedent and outcome correspond in magnitude. *Personality and Social Psychology Bulletin*, 24, 6, 596-609.
- Spellman, B. A. & Mandel, D. R. (1999). When possibility informs reality: Counterfactual thinking as a cue to causality. *Current Directions in Psychological Science*, 8, 120-123.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27, 76-105.
- Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98-112). Oxford: Basil Blackwell.
- Staudenmayer, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.) *Reasoning: Representation and process* (pp. 55-79). New York: Wiley.

- Stevenson, R. J. & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology*, 48A, 613-643.
- Storms, M. D. (1973). Videotape and attribution process: Reversing actors' and observers' point of view. *Journal of Personality and Social Psychology*, 27, 165-175.
- Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam, Netherlands: North-Holland.
- Swain, M. (1978). A counterfactual analysis of event causation. *Philosophical Studies*, 34, 1-19.
- Sweeney, P. D., Anderson, K. & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology*, 50, 974-991.
- Taylor, S. E. & Fiske, S. T. (1975). Point of view and perceptions of causality. *Journal of Personality and Social Psychology*, 32, 439-445.
- Taylor, S. E. & Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 11 (pp. 249-288). New York: Academic Press.
- Taylor, S. E. & Pham, L. B. (1996). Mental simulation, motivation and action. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behaviour* (pp. 219-235). London: Guilford Press.
- Tetlock, P. E. & Belkin, A. (Eds.) (1996). *Counterfactual thought experiments in world politics: Logical, methodological, and psychological perspectives*. Princeton, NJ: Princeton University Press.

- Thompson, S. C. (1993). Naturally occurring perceptions of control: A model of bounded flexibility. In G. Weary, F. Gleicher & K. L. Marsh (Eds.), *Control motivation and social cognition* (pp. 74-93). New York: Springer-Verlag.
- Thompson, V. A. & Byrne, R. M. J. (2000). Reasoning counterfactually: Making inferences about things that didn't happen. *Manuscript submitted for publication*.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22, 6, 742-758.
- Thorndike, E. L. (1905). *The elements of psychology*. New York: Mason-Henry.
- Trabasso, T. & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, 24, 595-611.
- Turnbull, W. & Slugoski, B. R. (1988). Conversational and linguistic processes in causal attribution. In D. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 66-93). Brighton, UK: Harvester Press.
- Tversky, A. & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.) *Progress in social psychology, Vol. 1* (pp. 49-72). Hillsdale, NJ: Erlbaum.
- von Wright, G. H. (1973). On the logic and epistemology of the causal relation. In P. Suppes (Ed.), *Logic, Methodology and Philosophy of Science, IV* (pp. 293-312). Amsterdam, Netherlands: North-Holland.
- Wasserman, D., Lempert, R. O. & Hastie, R. (1991). Hindsight and causality. *Personality and Social Psychology Bulletin*, 17, 30-35.
- Weiner, B. (1985a). "Spontaneous" causal thinking. *Psychological Bulletin*, 97, 1, 74-84.

- Weiner, B. (1985b). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer Verlag.
- Weiner, B., Russel, D. & Lerman, D. (1978). Affective consequences of causal ascriptions. In J. H. Harvey, W. J. Ickes & R. F. Kidd (Eds.), *New Directions in Attribution research*, Vol. 2 (59-88). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wells, G. L. & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, 56, 161-169.
- Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology*, 53, 421-430.
- White, P. E. (1989). A theory of causal processing. *British Journal of Psychology*, 80, 431-454.
- Williams, C. W., Lees-Haley, P. R. & Price, J. R. (1996). The role of counterfactual thinking and causal attribution in accident-related judgments. *Journal of Applied Social Psychology*, 26, 2100-2112.
- Wilson, K. V. (1956). A distribution-free test of analysis of variance hypotheses. *Psychological Bulletin*, 1, 96-101.
- Wong, P. T. P. & Weiner, B. (1981). When people ask "why" questions and the heuristics of attributional search. *Journal of Personality and Social Psychology*, 40, 650-663.
- Zeelenberg, M. (1996). *On the importance of what might have been : Psychological perspectives on regret and decision making*. Ph.D. thesis: University of Amsterdam, Netherlands.

Appendix 1 Sample experimental materials

(a) Experiment 1

Page 1:

Thank you very much for agreeing to take part in these experiments. There are two separate experiments in this booklet. They are not intelligence tests, and there are no “good” or “bad” answers. Please take your time, and work through the tasks in the order in which they are presented. Please do not look ahead, or go back over previous answers. Your responses will be completely confidential. Please record your age and gender below:

Age:

Gender: M..... F.....

Page 2:

The following story describes a real-life experience. Please read it slowly and carefully, and try to imagine yourself in the situation described. Try to experience the events as if they are really happening to you, and imagine what you would think.

 You’re moving house to start a new job in a different city. The night before you leave, you write down your thoughts about the move in your diary: ...I’ve got mixed feelings about moving to a place where I know hardly anyone - my friends and social life are so important to me.

unexpected (bad outcome) / expected (good outcome): *But I’m sure it will be easy to settle in to the new town - I’ve never had any trouble making new friends.*

expected (bad outcome) / unexpected (good outcome): *I’m sure it will be hard to settle in to the new town - I’ve always had trouble making new friends.*

A lot happens in your first two weeks in the new town. During your first week at work, a staff dinner is held.

controllable: *You decide to go because you want to get to know your colleagues.*

uncontrollable: *You have to go because your boss has asked all the staff to be there.*

You enjoy the evening and meet a lot of people. That weekend, your next-door neighbours invite you to a party.

Most of the people who live on your road will be there.

controllable: *However, that evening you decide to go to the cinema instead.*

uncontrollable: *However, that evening you're extremely ill with the flu, so you can't go.*

The next week,

controllable: *you decide to ring an old friend who lives in the town and ask him to show you around. You arrange to go out with him the following evening*

uncontrollable: *you happen to bump into an old friend who lives in the town and he insists on showing you around. He takes you out the following evening*

and he introduces you to a lot of his friends. A few days later, a colleague tells you there's a membership vacancy at her sports club. You think joining would be a good way to meet people,

controllable: *but then you decide to spend the money on a new stereo instead.*

uncontrollable: *but there's no way you can afford the membership fee.*

Six weeks after the move, things have turned out

unexpected: *nothing like*

expected: *exactly as*

you had expected.

bad: *You haven't made any real friends in the new town and you feel very lonely and isolated. You are very upset*

good: *You've made a lot of good friends in the new town, and you feel quite happy and at home. You are very pleased*

unexpected: *and very surprised.*

expected: *but not surprised.*

Before turning the page, please take a moment to consider the events that have occurred and imagine the thoughts you might have about them.

Page 3:

That night, while thinking over the events of the past six weeks, you realise that you haven't written in your diary since the night before you moved. In the space below, jot down what you might write in your diary about your current situation and the events leading up to it. You may turn back and re-read the story first if you wish. Please try to fill the page.

Dear Diary...

(b) Experiment 2

Page 1:

Thank you very much for agreeing to take part in this experiment. The experiment is not an intelligence test, and there are no “good” or “bad” answers. Please take your time, and work through the tasks in the order in which they are presented. Please do not look ahead, or go back over previous answers. Your responses will be completely confidential. Please record your age and gender below:

Age:

Gender: M..... F.....

Page 2:

The following story describes a real-life experience. Please read it slowly and carefully, and try to imagine yourself in the situation described. Try to experience the events as if they are really happening to you, and imagine what you would think.

You're moving house to start a new job in a different city. The night before you leave, you write down your thoughts about the move in your diary: ...I've got mixed feelings about moving to a place where I know hardly anyone - my friends and social life are so important to me. But I'm sure it will be easy to settle in to the new town - I've never had any trouble making new friends...

A lot happens in your first two weeks in the new town. During your first week at work, a staff dinner is held. You have to go because your boss has asked all the staff to be there,

normal: *as he always*

unusual: *which he almost never*

does. You enjoy the evening and meet a lot of people. That weekend, your next-door neighbours invite you to a party. Most of the people who live on your road will be there. However, you suffer from very

normal: *frequent*

unusual: *occasional*

migraines, and as very

normal: *often*

unusual: *rarely*

occurs, that evening you have a really bad one, so you can't go. The next week, as

normal: *regularly*

unusual: *very unusually*

happens, you happen to bump into an old friend who lives in the town and he insists on showing you around. He takes you out the following evening and he introduces you to a lot of his friends. A few days later, a colleague tells you there's a membership vacancy at her sports club. You think joining would be a good way to meet people. However,

normal: *as you would never have that kind of money,*

unusual: *although you would very rarely not have enough money, at that time*

there's no way you can afford the membership fee.

Six weeks after the move, things have turned out nothing like you had expected. You haven't made any real friends in the new town and you feel very lonely and isolated. You are very upset, and very surprised.

Before turning the page, please take a moment to consider the events that have occurred and imagine the thoughts you might have about them.

Page 3:

That night, while thinking over the events of the past six weeks, you realise that you haven't written in your diary since the night before you moved. In the space below, jot down what you might write in your diary about your current situation and the events leading up to it. You may turn back and re-read the story first if you wish. Please try to fill the page.

Dear Diary...

(c) Experiment 3

Page 1:

Thank you very much for agreeing to take part in these experiments. There are two separate experiments in this booklet. They are not intelligence tests, and there are no “good” or “bad” answers. Please take your time, and work through the tasks in the order in which they are presented. Please do not look ahead, or go back over previous answers. Your responses will be completely confidential. Please record your age and gender below:

Age:
Gender: M..... F.....

Page 2:

Please read the following story slowly and carefully and then answer the questions that follow:

Mr. Jones is 47 years old, the father of three and a successful banking executive. He was recently involved in a serious car accident. On the day of the accident, Mr. Jones did not leave work at his usual time. He

time-controllable / route-uncontrollable: *decided*

time-uncontrollable / route controllable: *had*

to leave earlier than usual

time-controllable / route-uncontrollable: *as it was an exceptionally sunny day and he wanted to get out and enjoy it.*

time-uncontrollable / route-controllable: *to do some important errands for his wife, who was sick at home.*

Mr. Jones did not drive home by his usual route. He

time-controllable / route-uncontrollable: *had*

time-uncontrollable / route controllable: *decided*

to drive along the coast road

time-controllable / route-uncontrollable: *to do some important errands for his wife, who was sick at home.*

time-uncontrollable / route controllable: *as it was an exceptionally sunny day and he wanted to enjoy the view.*

Mr. Jones knew that accidents were much more likely to happen

time-covarying / route-non-covarying: *at that time than at the time he usually drove home, because speeding was so much more common at that time.*

time-non-covarying / route-covarying: *on the coast road than on the route he usually took, because speeding was so much more common on the coast road.*

So, although he was aware that

time-covarying / route-non-covarying: *the coast road was safer than the route he usually took,*

time-non-covarying / route-covarying: *the roads were safer at that time than at the time he usually drove home,*

Mr. Jones drove especially carefully.

The accident happened when Mr. Jones was about half-way home. He slowed down as he approached a bend in the road. At that moment a sports car, driven by a young man named Mark Smith, came around the corner at very high speed. Smith lost control, attempted to swerve to avoid the oncoming traffic, and crashed into the side of Mr. Jones’s car. Mr. Jones was seriously injured and spent three weeks recovering in hospital. As commonly happens in such situations, he often thought about the accident for some time afterwards.

Page 3:

Q.1 Please rate from 0 (not at all causal) to 8 (very causal) the extent to which each of the following factors caused

Mr. Jones’ accident:

	not at all causal			moderately causal				very causal	
a) the time Mr. Jones left work	0	1	2	3	4	5	6	7	8
b) the route Mr. Jones took	0	1	2	3	4	5	6	7	8
c) Mr. Jones’ driving	0	1	2	3	4	5	6	7	8
d) Mark Smith’s driving	0	1	2	3	4	5	6	7	8
e) Mrs. Jones’ illness	0	1	2	3	4	5	6	7	8

Q.2 Please write down four ways that Mr. Jones' accident could have been avoided "if only....":

If only.....

If only.....

Q.3 How much was the likelihood of Mr. Jones having an accident increased by:

b) the route he took home not at all 1 2 3 4 5 6 7 very much

Q.4 How much control did Mr. Jones have over:

b) the route he took home	no control	1	2	3	4	5	6	7	total control
---------------------------	------------	---	---	---	---	---	---	---	---------------

240

(d) Experiment 4

Page 1:

Thank you very much for agreeing take part in this study. This study is concerned with how people think about everyday events. It is not a test of intelligence, personality, or any other characteristic, and there are no right or wrong answers. You will be asked to read a short story and to answer some questions. Please take your time and give the answers that come to mind most easily. It is important to complete the tasks in the order in which they are presented. Please do not look ahead, or go back and change your answers once you have written them. If you have a question, you can ask the researcher, but please don't consult your neighbour. Please record your age and gender below:

Age:years

Gender: M..... F.....

Page 2:

The following story describes a real-life experience. Please read it slowly and carefully, and try to imagine yourself in the situation described. Try to experience the events as if they are really happening to you, and imagine what you would think.

You are a student in your first year at University. You have to pass some end of year exams in order to continue into second year. Students are randomly assigned to be examined by either Dr. Smith in May or Dr. Jones in June. You are initially assigned to Dr. Smith in May, but you ask the course director if you can be reassigned to Dr. Jones in June, so that you'll have more time to prepare for the exams. She agrees.

You sit your exams in June and you think they go fairly well. In July, both Dr. Smith and Dr. Jones post their students' results on the notice board. You find your name on Dr. Jones's list and learn that you have failed and will not be allowed into second year. You are very disappointed. Looking more closely at the lists, you notice that most of the other students examined by Dr. Jones also failed the year, whereas very few of those examined by Dr. Smith failed. You are very upset.

counterfactual: *You think that you would not have failed the year if...*

causal: *You think that you failed the year because...*

control: *Please write a one-sentence summary of the above story on the lines below:...*

Before turning the page, please take a moment to imagine what you would think and feel if you were really in this situation.

Page 3:

Please answer each of the questions below by circling one of the numbers from 1 to 9.

(i) To what extent could you have prevented failing the year?

Not at all preventable							Totally preventable	
1	2	3	4	5	6	7	8	9

(ii) How much control did you have over whether or not you failed the year?

Not at all controllable							Totally controllable	
1	2	3	4	5	6	7	8	9

(iii) How predictable was it that you would fail the year?

Not at all predictable							Totally predictable	
1	2	3	4	5	6	7	8	9

(e) Experiment 5

Page 1 (English version):

Thank you very much for agreeing take part in this study. This study is concerned with how people think about everyday events. It is not a test of intelligence, personality, or any other characteristic, and there are no right or wrong answers. You will be asked to read a short story and to answer some questions. Please take your time and give the answers that come to mind most easily. It is important to complete the tasks in the order in which they are presented. Please do not look ahead, or go back and change your answers once you have written them. If you have a question, you can ask the researcher, but please don't consult your neighbour. Please record your age and gender below:

Age:years

Gender: M..... F.....

Page 1 (Italian version):

Ti ringraziamo per la tua collaborazione. Questo lavoro riguarda il modo in cui le persone pensano agli eventi quotidiani. Non si tratta di un test di intelligenza, né di personalità e non ci sono risposte giuste o sbagliate da fornire. Dovrai semplicemente leggere una breve storia e rispondere ad alcune domande. Fornisci la risposta che ti viene in mente più facilmente. Non hai limiti di tempo per farlo. Completa il compito nell'ordine in cui ti viene presentato, non saltare pagine e non cambiare la risposta una volta che l'hai fornita. Se hai delle domande, falle alla persona che ha distribuito il questionario. Per favore, non parlare con il tuo vicino. Indica qui sotto i tuoi dati personali:

Età:anni

Sesso: M..... F.....

Page 2 (English version):

During the first month of the First World War (1914-1918), hundreds of thousands of lives were lost during the Battle of the Frontiers. At this time, the Germans fought the French and the British along the border between

Germany and France. In these early weeks, the Germans quickly acquired a reputation for being ruthless and harsh, particularly as they burned and terrorised villages in Belgium and France.

During the battle, a small British force accidentally encountered a much larger German force. The British were forced to retreat into a village, putting all the inhabitants into extreme danger of attack. The village was small and peaceful. It was inhabited mainly by farmers, and among them many children. Unfortunately, the British commander, Lt. Dorian Moore, was slow, inexperienced, and indecisive. Moore ordered his troops to stay in the village, and to simply wait for a German attack. This spelled certain doom for the village.

One British soldier was Thomas Jensen. He was only 20 years old, and although somewhat shy and unsure of himself, he was bright and had already involved himself in several minor military decisions for the British. He saw a way of luring the Germans away from the village, thereby saving the village from destruction. The more he worked on his plan, the more he became convinced that he could save the village, as well as his fellow soldiers. The problem was: since Jensen was so young, it was not certain that Moore would listen to him.

Jensen faced a difficult decision. As he saw it, he had three alternatives: 1) He could simply forget his plan and follow the orders Moore had already given. 2) Since he was so well-liked, he could organise his fellow soldiers to disobey Moore's orders, and to follow his plan instead. 3) He could carefully explain his plan to another officer, and have that officer tell Moore. With luck, Moore might order the troops to follow his plan.

Of these three choices, choice 1 was the least attractive (since it meant certain doom for the village). Choice 2 was riskiest to Jensen himself (he could be tried and shot for mutiny if his plan failed) but it could save the village (assuming that the troops were willing to follow Jensen's plan). Choice 3 was safer for Jensen (his conduct could not be questioned) and it could also save the village (assuming Moore would change his orders and instruct the troops to follow Jensen's plan). After a sleepless night, Jensen decided on choice 3.

counterfactual / causal / outcome-control: *The outcome was that Moore rejected Jensen's plan and strictly reaffirmed his original orders to the troops. The village was destroyed.*

Page 2 (Italian version):

Nei primi mesi della Prima Guerra Mondiale (1914-1918), la cosiddetta Battaglia delle Frontiere causò centinaia di migliaia di vittime. Tale battaglia venne combattuta, al confine tra Francia e Germania, dagli eserciti

francese e inglese contro l'esercito tedesco. Nelle prime settimane di combattimento le truppe tedesche si guadagnarono la reputazione di essere spietate, dato che seminavano terrore e distruzione nei villaggi francesi e belgi da loro conquistati.

Durante questa battaglia un piccolo distaccamento inglese venne a contatto con una forza tedesca molto più grande. Gli inglesi furono costretti a ritirarsi in un villaggio, i cui abitanti vennero perciò a trovarsi in una situazione di estremo pericolo. Questo villaggio era piccolo e pacifico, abitato com'era da contadini, molti dei quali erano solo dei bambini. Sfortunatamente il comandante delle truppe inglese, il tenente Dorian Moore, era lento, privo di esperienza e indeciso. Moore ordinò alle sue truppe di fermarsi nel villaggio e di aspettare l'attacco tedesco. Ciò produsse dei sinistri presagi circa la sorte del villaggio.

Uno dei soldati inglesi si chiamava Thomas Jensen. Aveva solo vent'anni e, benché fosse un po' timido e insicuro, era molto intelligente ed era già stato coinvolto in parecchie decisioni militari del suo reparto. Jensen intravide la possibilità di distogliere i tedeschi dal villaggio, salvandolo così dalla distruzione. Più Jensen pensò al suo piano, più si convinse della possibilità di salvare sia il villaggio che i suoi commilitoni. C'era però un problema: dato che Jensen era molto giovane, non era detto che il tenente Moore l'avrebbe ascoltato.

Jensen aveva di fronte a sé una decisione molto difficile. Per quel che gli sembrava, aveva solo tre alternative. (1) Poteva semplicemente dimenticare il suo piano e seguire gli ordini che il tenente Moore aveva già impartito. (2) Dato che era così ben voluto dagli altri, poteva convincere i suoi commilitoni a disobbedire agli ordini di Moore e a seguire, invece, il suo piano. (3) Poteva spiegare in modo dettagliato il suo piano ad un altro ufficiale, il quale ne poteva poi parlare con Moore. Con un po' di fortuna Moore avrebbe potuto ordinare ai soldati di seguire il piano elaborato da Jensen.

Di queste tre scelte, la (1) era la meno attraente, dato che significava condannare il villaggio alla distruzione. La scelta (2) era la più rischiosa per Jensen (che poteva essere processato e fucilato per ammutinamento, qualora il suo piano fosse fallito), ma poteva salvare il villaggio (assumendo che le truppe avrebbero seguito il suo piano). La scelta (3) era la più sicura per Jensen (la cui condotta non avrebbe potuto essere messa sotto accusa) e poteva anche salvare il villaggio (assumendo che Moore avrebbe cambiato i propri piani, ordinando alle truppe di seguire quello di Jensen). Dopo una notte insonne, Jensen decise per la scelta (3).

counterfactual / causal / outcome-control: *Il risultato fu che Moore rifiutò il piano di Jensen e ordinò alle truppe di seguire rigidamente il suo piano iniziale. Il villaggio venne distrutto.*

Page 3 (English version):

1. Please complete the following description of the story you have just read:

The story is about the misfortune of a village during the First World War. The village

counterfactual: *might have been saved if...*

causal: *was destroyed because...*

control: *was...*

2. counterfactual / causal / outcome-control: *Some people taking part in this study were asked to read the story but were not told the outcome. Please try to put yourself in their shoes and answer the following questions as if you do not know the outcome:*

(i) Please rate your level of agreement with the statements below by circling one of the numbers from 1 (strongly disagree) to 9 (strongly agree):

It was predictable in advance that the village would be *saved*

1	2	3	4	5	6	7	8	9
Strongly disagree			Neutral			Strongly agree		

It was predictable in advance that the village would be *destroyed*

1	2	3	4	5	6	7	8	9
Strongly disagree			Neutral			Strongly agree		

(ii) Please judge the likelihood of each of the following outcomes by writing a probability value of from 0 to 100% next to each outcome. The TOTAL of all your probability judgements should equal 100%.

The village is *saved*.....%

The village is *destroyed*.....%

Page 3 (Italian version):

1. Ti chiediamo di completare la seguente descrizione della storia che hai appena letto:

La storia riguarda il destino sfortunato di un villaggio, durante la Prima Guerra Mondiale. Il villaggio

counterfactual: *avrebbe potuto essere salvato se...*

causal: *venne distrutto perché...*

control: *fu...*

2. counterfactual / causal / outcome-control: *Ad alcune persone che hanno preso parte a questa ricerca è stata presentata la storia che hai appena letto, ma senza la conclusione. Cerca di metterti nei panni di queste persone e di rispondere alle seguenti domande come se tu ignorassi la conclusione della storia:*

(i) Indica il tuo grado di accordo con le frasi che seguono, facendo un segno su uno dei numeri che vanno da 1 (totalmente in disaccordo) a 9 (totalmente d'accordo):

Si poteva prevedere in anticipo che il villaggio sarebbe stato *salvato*

1	2	3	4	5	6	7	8	9
Totalmente in disaccordo				Neutro		Totalmente d'accordo		

Si poteva prevedere in anticipo che il villaggio sarebbe stato *distrutto*

1	2	3	4	5	6	7	8	9
Totalmente in disaccordo				Neutro		Totalmente d'accordo		

(ii) Stima la probabilità di ognuno degli esiti presentati qui sotto, indicando, per ogni esito, un valore di probabilità compreso tra 0% e 100%. La somma totale dei tuoi giudizi di probabilità dovrebbe essere 100%.

Il villaggio viene *salvato*.....%

Il villaggio viene *distrutto*.....%

Thank you *very* much

Gender: M.....

F.....

Please take a moment

The event should have the following characteristics:

- Some examples of the types of events you might recall are doing badly in an exam or competition, the end of a relationship, a fight with a friend or relative, being involved in an accident or crime, a financial or material loss, etc..

In the space below, please write a brief description of the event and its outcome:...

How upsetting was the event at the time? (Please circle one of the numbers from 1 to 7)

Not at all upsetting	1	2	3	4	5	6	7	Extremely upsetting
----------------------	---	---	---	---	---	---	---	---------------------

Following negative life

Following negative life events, people often think about

general upward counterfactual: *how the outcome could have turned out better.*

internal upward counterfactual: *how, if they had acted differently, the outcome could have turned out better.*

general downward counterfactual: *how the outcome could have turned out even worse.*

internal downward counterfactual: *how, if they had acted differently, the outcome could have turned out even worse.*

general causal: *why the outcome turned out as it did.*

internal causal: *how their own actions might explain why the outcome turned out as it did.*

control: *the event a lot.*

Please think carefully about the event you have described, and in the space below, write down as many

general upward counterfactual: *ways that the outcome could have turned out better*

internal upward counterfactual: *ways that you could have made the outcome better*

general downward counterfactual: *ways that the outcome could have turned out worse*

internal downward counterfactual: *ways that you could have made the outcome worse*

general causal: *causal explanations for the outcome*

internal causal: *ways that you might have caused or enabled the outcome*

control: *thoughts about the outcome*

as come to mind. Please take your time, write down everything that occurs to you, and feel free to elaborate on your opinions....

When you have finished writing, and before turning the page, please read over your list of

general / internal upward counterfactual: *ways that the outcome could have turned out better*

general / internal downward counterfactual: *ways that the outcome could have turned out worse*

general / internal causal: *causal explanations for the outcome*

control: *thoughts about the outcome*

slowly and carefully, thinking about their meaning.

Page 4:

Please answer each of the following questions by circling one of the numbers from 1 to 7.

1. Your current feelings about the event

To what extent does thinking about the event *now* make you feel:

	Neutral							
Regret	1	2	3	4	5	6	7	Satisfaction
Unhappiness	1	2	3	4	5	6	7	Happiness
Disappointment	1	2	3	4	5	6	7	Relief
Depression	1	2	3	4	5	6	7	Elation
Guilt/shame	1	2	3	4	5	6	7	Pride
Pessimism	1	2	3	4	5	6	7	Optimism

2. Your current thoughts about the event

How much control do you think you had over the outcome?

No control	1	2	3	4	5	6	7	Total control
------------	---	---	---	---	---	---	---	---------------

To what extent do you think you were responsible for the outcome?

Not at all responsible	1	2	3	4	5	6	7	Totally responsible
------------------------	---	---	---	---	---	---	---	---------------------

To what extent were do you think you were to blame for the outcome?

Not at all to blame	1	2	3	4	5	6	7	Totally to blame
---------------------	---	---	---	---	---	---	---	------------------

How predictable do you think the outcome was in advance?

Not at all predictable	1	2	3	4	5	6	7	Totally predictable
------------------------	---	---	---	---	---	---	---	---------------------

Page 5:

3. Your thoughts about the future

Imagine that in the future, you are in a *similar* situation where there is a possibility of the same *type* of negative outcome occurring again (e.g., if you described an exam failure, another exam is approaching; if you described the end of a relationship, another relationship is deteriorating).

How likely do you think it is that things would turn out better this time?

Very unlikely 1 2 3 4 5 6 7 Very likely

How much control do you think you would have over the outcome?

No control 1 2 3 4 5 6 7 Total control

How likely do you think it is that you could prevent the same type of negative outcome from occurring again?

Very unlikely 1 2 3 4 5 6 7 Very likely

How likely do you think it is that you would act differently from the way you acted last time?

Very unlikely 1 2 3 4 5 6 7 Very likely

(g) Experiment 7

Page 1:

Thank you very much for agreeing to take part in this experiment on everyday reasoning. You will be asked to read statements about past events and to indicate what was possible and what was impossible, assuming that each statement is true. This is not a test of intelligence or personality, and there are no good or bad answers. However, you should take your time and think carefully about your answers.

Page 2:

Please read these instructions very carefully before working through the questions that follow:

Imagine any two events, each of which may or may not occur, for example, drinking a certain liquid and becoming ill. Varying the occurrence versus non-occurrence of both events results in four possible scenarios:

- You drink the liquid and you become ill.
- You drink the liquid and you do not become ill.
- You do not drink the liquid and you become ill.
- You do not drink the liquid and you do not become ill.

However, if there is some *relationship* between the events, for example, if drinking or not drinking the liquid *causes* or *prevents* becoming ill, then it may be that only some of these scenarios are *really possible* and the others are *really impossible*.

On the following pages, you will be given statements about relationships between pairs of past events. The statements themselves need not be true. In fact, some of them are true and some of them are false.

Regardless of whether you think that a statement is true or false, your task is to decide which scenarios were really possible and which were really impossible *if the statement is true*.

1. If exercise had not been excessive, then angina would not have occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,

and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Exercise that is excessive occurs and angina occurs.
- Exercise that is excessive occurs and angina does not occur.
- Exercise that is excessive does not occur and angina occurs.
- Exercise that is excessive does not occur and angina does not occur.

Page 3:

2. Reactivity that was not excessive prevented the occurrence of damage to the reactor.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Reactivity that is excessive occurs and damage to the reactor occurs.
- Reactivity that is excessive occurs and damage to the reactor does not occur.
- Reactivity that is excessive does not occur and damage to the reactor occurs.
- Reactivity that is excessive does not occur and damage to the reactor does not occur.

3. Exercise that was excessive caused the occurrence of angina.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Exercise that is excessive occurs and angina occurs.
- Exercise that is excessive occurs and angina does not occur.
- Exercise that is excessive does not occur and angina occurs.
- Exercise that is excessive does not occur and angina does not occur.

4. If genes were diverse, then new diseases developed.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Genes are diverse and new diseases develop.

.....
- Genes are diverse and new diseases do not develop.

.....
- Genes are not diverse and new diseases develop.

.....
- Genes are not diverse and new diseases do not develop.

.....

Page 4:

5. A childhood that was happy caused the occurrence of amnesia.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- A childhood that is happy occurs and amnesia occurs.

.....
- A childhood that is happy occurs and amnesia does not occur.

.....
- A childhood that is happy does not occur and amnesia occurs.

.....
- A childhood that is happy does not occur and amnesia does not occur.

.....

6. A childhood that was not happy prevented the occurrence of amnesia.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- A childhood that is happy occurs and amnesia occurs.

.....
- A childhood that is happy occurs and amnesia does not occur.

.....
- A childhood that is happy does not occur and amnesia occurs.

.....
- A childhood that is happy does not occur and amnesia does not occur.

.....

7. If solar energy had not been used, then global warming would not have occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Solar energy is used and global warming occurs.
- Solar energy is used and global warming does not occur.
- Solar energy is not used and global warming occurs.
- Solar energy is not used and global warming does not occur.

Page 5:

8. If the childhood had not been happy, then amnesia would not have occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- A childhood that is happy occurs and amnesia occurs.
- A childhood that is happy occurs and amnesia does not occur.
- A childhood that is happy does not occur and amnesia occurs.
- A childhood that is happy does not occur and amnesia does not occur.

9. Genes that were diverse caused the development of new diseases.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Genes are diverse and new diseases develop.
- Genes are diverse and new diseases do not develop.
- Genes are not diverse and new diseases develop.
- Genes are not diverse and new diseases do not develop.

10. Exercise that was not excessive prevented the occurrence of angina.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Exercise that is excessive occurs and angina occurs.

- Exercise that is excessive occurs and angina does not occur.
- Exercise that is excessive does not occur and angina occurs.
- Exercise that is excessive does not occur and angina does not occur.

Page 6:

11. If reactivity had not been excessive, then damage to the reactor would not have occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Reactivity that is excessive occurs and damage to the reactor occurs.
- Reactivity that is excessive occurs and damage to the reactor does not occur.
- Reactivity that is excessive does not occur and damage to the reactor occurs.
- Reactivity that is excessive does not occur and damage to the reactor does not occur.

12. If the childhood was happy, then amnesia occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- A childhood that is happy occurs and amnesia occurs.
- A childhood that is happy occurs and amnesia does not occur.
- A childhood that is happy does not occur and amnesia occurs.
- A childhood that is happy does not occur and amnesia does not occur.

13. If reactivity was excessive, then damage to the reactor occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Reactivity that is excessive occurs and damage to the reactor occurs.

- Reactivity that is excessive occurs and damage to the reactor does not occur.
- Reactivity that is excessive does not occur and damage to the reactor occurs.
- Reactivity that is excessive does not occur and damage to the reactor does not occur.

Page 7:

14. Non-use of solar energy prevented the occurrence of global warming.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Solar energy is used and global warming occurs.
- Solar energy is used and global warming does not occur.
- Solar energy is not used and global warming occurs.
- Solar energy is not used and global warming does not occur.

15. If exercise was excessive, then angina occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Exercise that is excessive occurs and angina occurs.
- Exercise that is excessive occurs and angina does not occur.
- Exercise that is excessive does not occur and angina occurs.
- Exercise that is excessive does not occur and angina does not occur.

16. Reactivity that was excessive caused the occurrence of damage to the reactor.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Reactivity that is excessive occurs and damage to the reactor occurs.
- Reactivity that is excessive occurs and damage to the reactor does not occur.

Reactivity that is excessive does not occur and damage to the reactor occurs.

Reactivity that is excessive does not occur and damage to the reactor does not occur.

Page 8:

17. Use of solar energy caused the occurrence of global warming.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

Solar energy is used and global warming occurs.

Solar energy is used and global warming does not occur.

Solar energy is not used and global warming occurs.

Solar energy is not used and global warming does not occur.

18. Genes that were not diverse prevented the development of new diseases.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

Genes are diverse and new diseases develop.

Genes are diverse and new diseases do not develop.

Genes are not diverse and new diseases develop.

Genes are not diverse and new diseases do not develop.

19. If solar energy was used, then global warming occurred.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

Solar energy is used and global warming occurs.

Solar energy is used and global warming does not occur.

Solar energy is not used and global warming occurs.

Solar energy is not used and global warming does not occur.

Page 9:

20. If genes had not been diverse, then new diseases would not have developed.

Please place a *tick* beside the scenarios that were *possible* if the above statement is true,
and place a *cross* beside the scenarios that were *impossible* if the above statement is true:

- Genes are diverse and new diseases develop.
- Genes are diverse and new diseases do not develop.
- Genes are not diverse and new diseases develop.
- Genes are not diverse and new diseases do not develop.

Thank you very much for your participation!

To finish, please answer the following questions for the records of the experiment:

- 1. What is your age?
.....years
- 2. What is your gender?
Male..... Female.....
- 3. Have you ever studied logic?
Yes..... No.....
- 4. Have you studied the psychology of reasoning beyond undergraduate level?
Yes..... No.....

(h) Experiment 8

Page 1:

Thank you very much for agreeing to take part in this study. This study is concerned with how people think about everyday events and situations. It is not a test of intelligence, personality, or any other characteristic, and there are no right or wrong answers. However, it is important to take your time and think carefully about your answers. Please carefully read the instructions below before working through the problems that follow. If you have any questions, you can ask the researcher.

Instructions:

On the following pages, you will be asked to make inferences from information about everyday events and situations. In each case, you will be presented with two factual statements, and you will be asked to say what, if anything, follows. Here is an example:

All Trinity students are intelligent.
Paul is a Trinity student.
What, if anything, follows?
.....

In this example, if you think that the two statements, “All Trinity students are intelligent” and “Paul is a Trinity student” necessarily imply that Paul is intelligent, then you should write down “Paul is intelligent” as your answer. If you think that the statements necessarily imply that Paul is not intelligent, then you should write “Paul is not intelligent”. If you think that the statements imply nothing definite about Paul’s intelligence, then you should write “Nothing follows”.

The problems that you will complete will have the same form as the example shown above, but in each case, you will be asked to draw conclusions from different statements. It is important to read each set of statements very carefully, and to think about your answers. There are 16 problems in total, and it is essential that you complete all of them. Please complete them in the order in which they are presented. Do not look ahead, or go back over previous answers.

Page 2:

1.

If fertiliser is put on the plants, then they grow quickly.

Fertiliser is put on the plants.

What, if anything, follows?

.....
2.

Studying hard prevents John from not getting good marks.

John does not study hard.

What, if anything, follows?

.....
3.

Not depressing the brake causes the car not to slow down.

The brake is not depressed.

What, if anything, follows?

.....
4.

If Jenny does not turn on the air conditioning, then she does not feel cool.

Jenny turns on the air conditioning.

What, if anything, follows?

.....

Page 3:

5.

Putting fertiliser on the plants causes them to grow quickly.

Fertiliser is put on the plants.

What, if anything, follows?

.....
6.

If John studies hard, then he gets good marks.

John does not study hard.

What, if anything, follows?

.....

7. If the brake were not depressed, then the car would not slow down.

The brake is not depressed.

What, if anything, follows?

.....

8. Not turning on the air conditioning causes Jenny to not feel cool.

Jenny turns on the air conditioning.

What, if anything, follows?

.....

Page 4:

9. If fertiliser were put on the plants, then they would grow quickly.

Fertiliser is put on the plants.

What, if anything, follows?

.....

10. Studying hard causes John to get good marks.

John does not study hard.

What, if anything, follows?

.....

11. Not depressing the brake prevents the car from slowing down.

The brake is not depressed.

What, if anything, follows?

.....

12. If Jenny were not to turn on the air conditioning, then she would not feel cool.

Jenny turns on the air conditioning.

What, if anything, follows?

.....

Page 5:

13. Putting fertiliser on the plants prevents them from not growing quickly.

Fertiliser is put on the plants.

What, if anything, follows?

.....

14. If John were to study hard, then he would get good marks.

John does not study hard.

What, if anything, follows?

.....

15. If the brake is not depressed, then the car does not slow down.

The break is not depressed.

What, if anything, follows?

.....

16. Not turning on the air conditioning prevents Jenny from feeling cool.

Jenny turns on the air conditioning.

What, if anything, follows?

.....

Page 6:

Thank you very much for your participation!

To finish, please answer the following questions for the records of the experiment:

- 1 What is your age?
 years
- 2. What is your gender?
Male..... Female.....
- 3. Have you ever studied logic?
Yes..... No.....
- 4. Have you studied the psychology of reasoning beyond undergraduate level?
Yes..... No.....

(i) Experiment 9

Verbal instructions:

This study is concerned with how people think about everyday events and situations. It is not a test of intelligence, but it is important to try to concentrate and think carefully about your answers. The study is made up of a number of problems that you work through at your own pace. For each problem, the screen will display a sentence, followed by another sentence, followed by a question. When you have finished reading a sentence and you want to continue to the next screen, you press this key marked "NEXT". When you want to answer a question, you press one of these three answer keys marked "Y" for "yes", "N" for "no", and "CT" for "can't tell". For example, the screen might first display a sentence like this: "All sports cars drive fast". When you've read and understood the sentence you press this key marked "next". Then it gives you the second sentence, "Clare's car is not a sports car". Again, you press "NEXT". Then it gives you a question, "Does Clare's car drive fast?". To answer the question you have to think about the meaning of the two sentences that preceded it, "All sports cars drive fast" and "Clare's car is not a sports car". If you think that the two sentences, "All sports cars drive fast" and, "Clare's car is not a sports car" necessarily imply that Clare's car drives fast, then you should press "Y" for "yes". If you think that "All sports cars drive fast" and "Clare's car is not a sports car" necessarily imply that Clare's car does not drive fast, then you should press "N" for "no". If you think that you can't tell anything definite about whether or not Clare's car drives fast just from those two statements, then you should press "CT" for "can't tell". One thing that it's important to remember is that you should assume that the statements are true. So for example, you might think that it's not true that all sports cars drive fast. But when you're answering the question you should assume that it is true that all sports cars drive fast and that Clare's car is not a sports car, and decide whether or not Clare's car drives fast based just on those two facts. So that's one problem, and then you just press "next" whenever you're ready to continue to the next problem. If you want to take a short break or ask a question, you can stop between problems, so when it says "press next to continue" you can wait a few seconds before you continue. But it's very important that you don't stop in the middle of a problem.

Demonstration and practice problems:

PRESS "NEXT" FOR A DEMONSTRATION

All sports cars drive fast.
Clare's car is not a sports car.

QUESTION

Does Clare's car drive fast?

PRESS "NEXT" TO CONTINUE

END OF DEMONSTRATION

PRESS "NEXT" TO PRACTICE ON SOME EXAMPLE PROBLEMS

Some people from Zog drink wine.
Phil is not from Zog.

QUESTION

Does Phil drink wine?

PRESS "NEXT" TO CONTINUE

Genya earns more than Orlando.
Orlando earns more than Ruth.

QUESTION

Does Ruth earn more than Genya?

PRESS "NEXT" TO CONTINUE

All kangaroos are red.
Most red animals are dangerous.

QUESTION

Are any kangaroos dangerous?

PRESS "NEXT" TO CONTINUE

Only people over 18 or accompanied by an adult may enter the bar.
Vittorio is not over 18.

QUESTION

May Vittorio enter the bar?

PRESS "NEXT" TO CONTINUE

None of the athletes are Frenchmen.
All of the Frenchmen are teachers.

QUESTION

Are any of the athletes teachers?

PRESS "NEXT" TO CONTINUE

Juan is taller than Valerie.

Juan is not taller than Antonio.

QUESTION

Is Valerie taller than Antonio?

PRESS "NEXT" TO CONTINUE

All of the dogs eat biscuits.

Arthur is not a dog.

QUESTION

Does Arthur eat biscuits?

PRESS "NEXT" TO CONTINUE

Only students who fail the exam are not allowed to graduate.

Rachel does not fail the exam.

QUESTION

Is Rachel allowed to graduate?

END OF PRACTICE - WELL DONE!

PRESS "NEXT" TO START THE TRIAL

Example experimental problems:

Turning on the air conditioning causes Kate to feel cool.

Kate turns on the air conditioning.

QUESTION

Does Kate feel cool?

PRESS "NEXT" TO CONTINUE

If a person were to not eat sweets, then they would not get cavities.

Rose does not eat sweets.

QUESTION

Does Rose get cavities?

PRESS "NEXT" TO CONTINUE

There being bad weather conditions prevents the plane from not crashing.

There are bad weather conditions.

QUESTION

Does the plane crash?

PRESS "NEXT" TO CONTINUE

Not breathing asbestos causes a person to not develop lung cancer.

Dave breathes asbestos.

QUESTION

Does Dave develop lung cancer?

PRESS "NEXT" TO CONTINUE

If there were an increase in unemployment, then the crime rate would go up.

There is an increase in unemployment.

QUESTION

Does the crime rate go up?

PRESS "NEXT" TO CONTINUE

Not changing the oil regularly causes the engine to not last for 10 years.

The oil is not changed regularly.

QUESTION

Does the engine last for 10 years?

PRESS "NEXT" TO CONTINUE

If someone takes ibuprofen, then they become nauseous.

Pat does not take ibuprofen.

QUESTION

Does Pat become nauseous?

PRESS "NEXT" TO CONTINUE

If a person were to not eat a lot of fat, then they would not develop heart disease.

Mick eats a lot of fat.

QUESTION

Does Mick develop heart disease?

PRESS "NEXT" TO CONTINUE

If a person is exposed to radiation, then they develop cancer.

QUESTION

Anne is exposed to radiation.

Does Anne develop cancer?

PRESS "NEXT" TO CONTINUE

There being traces of PCB's in the water prevents the fish from not dying.

There are not traces of PCB's in the water.

QUESTION

Do the fish die?

PRESS "NEXT" TO CONTINUE

If the brake were depressed, then the car would slow down.

The brake is not depressed.

QUESTION

Does the car slow down?

PRESS "NEXT" TO CONTINUE

If there is not snow overnight, then the number of car accidents does not increase.

There is snow overnight.

QUESTION

Does the number of car accidents increase?

PRESS "NEXT" TO CONTINUE

There being a low pressure system causes it to rain.

There is not a low pressure system.

QUESTION

Does it rain?

PRESS "NEXT" TO CONTINUE

Not studying hard prevents John from getting good marks.

John does not study hard.

QUESTION

Does John get good marks?

PRESS "NEXT" TO CONTINUE

If fertiliser is not put on the plants, then they do not grow quickly.

Fertiliser is not put on the plants.

QUESTION

Do the plants grow quickly?

PRESS "NEXT" TO CONTINUE

There not being high interest rates prevents many businesses from going bankrupt.

There are high interest rates.

QUESTION

Do many businesses go bankrupt?

END OF TRIAL – WELL DONE!

Appendix 2 Sample causal and counterfactual thoughts from Experiment 1

Causal thoughts

General internal: “It’s my own fault for not taking all the opportunities to make new friends”

General external: “Due to unfortunate circumstances beyond my control I have been unable to establish myself in the local social scene”

Specific inhibitory: “Being sick for the neighbours’ dinner party was bad luck and therefore I don’t know any of my colleagues”

Specific facilitative: “Things have turned out well, mainly because I went to the staff dinner”

Counterfactual thoughts

General internal: “I wish I had tried harder”

General external: “If only I could slough off the strains of maturity”

Specific inhibitory: “I should have joined the sports club instead of spending the money on that stereo - I would have got to know so many people”

Specific facilitative: “I could have tried to make my own new friends instead of just continuing to go out with my old ones”

Appendix 3 Experiment 6 t-tests

Preventability

Upward vs downward: $t(103) = 0.89, p = 0.38$

Upward vs causal: $t(102) = 0.49, p = 0.29$

Upward vs control: $t(61) = 1.31, p = 0.19$

Downward vs causal: $t(105) = 1.36, p = 0.18$

Downward vs control: $t(60) = 0.58, p = 0.57$

Causal vs control: $t(64) = 1.7, p = 0.09$

General upward vs internal upward: $t(45) = 2.15, p = 0.04$

General upward vs general downward: $t(56) = 0.79, p = 0.43$

General upward vs internal downward: $t(54) = 0.57, p = 0.35$

General upward vs general causal: $t(61) = 0.55, p = 0.58$

General upward vs internal causal: $t(43) = 2.07, p = 0.04$

General upward vs control: $t(60) = 0.32, p = 0.75$

Internal upward vs general downward: $t(46) = 1.24, p = 0.22$

Internal upward vs internal downward: $t(40) = 2.78, p < 0.01$

Internal upward vs general causal: $t(46) = 1.57, p = 0.12$

Internal upward vs internal causal: $t(39) = 0.07, p = 0.94$

Internal upward vs control: $t(46) = 2.41, p = 0.02$

General downward vs internal downward: $t(51) = 1.35, p = 0.38$

General downward vs general causal: $t(60) = 0.26, p = 0.8$

General downward vs internal causal: $t(45) = 1.24, p = 0.22$

General downward vs control: $t(57) = 1.07, p = 0.29$

Internal downward vs general causal: $t(55) = 1.1, p = 0.28$

Internal downward vs internal causal: $t(38) = 2.65, p = 0.01$

Internal downward vs control: $t(54) = 0.23, p = 0.82$

General causal vs internal causal: $t(44) = 1.54, p = 0.13$

General causal vs control: $t(61) = 0.85, p = 0.4$

Internal causal vs control: $t(44) = 2.32, p = 0.02$

Predictability

Upward vs downward: $t(103) = 0.14, p = 0.89$

Upward vs causal: $t(102) = 1.81, p = 0.07$

Upward vs control: $t(58) = 0.5, p = 0.62$

Downward vs causal: $t(105) = 1.89, p = 0.06$

Downward vs control: $t(62) = 0.61, p = 0.55$

Causal vs control: $t(59) = 0.96, p = 0.34$

General upward vs internal upward: $t(43) = 0.2, p = 0.84$

General upward vs general downward: $t(57) = 0.1, p = 0.92$

General upward vs internal downward: $t(49) = 0.26, p = 0.56$

General upward vs general causal: $t(61) = 1.53, p = 0.13$

General upward vs internal causal: $t(42) = 0.83, p = 0.41$

General upward vs control: $t(60) = 0.37, p = 0.71$

Internal upward vs general downward: $t(44) = 0.1, p = 0.93$

Internal upward vs internal downward: $t(43) = 0.07, p = 0.95$

Internal upward vs general causal: $t(42) = 1.62, p = 0.11$

Internal upward vs internal causal: $t(39) = 0.96, P = 0.34$

Internal upward vs control: $t(45) = 0.54, p = 0.61$

General downward vs internal downward: $t(50) = 0.16, p = 0.58$

General downward vs general causal: $t(57) = 1.58, p = 0.12$

General downward vs internal causal: $t(43) = 0.9, p = 0.37$

General downward vs control: $t(58) = 0.45, p = 0.65$

Internal downward vs general causal: $t(48) = 1.64, p = 0.1$

Internal downward vs internal causal: $t(43) = 1, p = 0.32$

Internal downward vs control: $t(51) = 0.59, p = 0.56$

General causal vs internal causal: $t(41) = 0.49, p = 0.63$

General causal vs control: $t(60) = 1.1, p = 0.28$

Internal causal vs control: $t(44) = 0.48, p = 0.6$

Future likelihood

Upward vs downward: $t(103) = 0.91, p = 0.37$

Upward vs causal: $t(102) = 0.82, p = 0.41$

Upward vs control: $t(57) = 0.63, p = 0.53$

Downward vs causal: $t(105) = 0.07, p = 0.94$

Downward vs control: $t(59) = 0.1, p = 0.92$

Causal vs control: $t(60) = 0.04, p = 0.97$

General upward vs internal upward: $t(38) = 0.89, p = 0.38$

General upward vs general downward: $t(57) = 1.69, p = 0.42$

General upward vs internal downward: $t(49) = 0.4, p = 0.69$

General upward vs general causal: $t(60) = 0.16, p = 0.87$

General upward vs internal causal: $t(49) = 3.2, p < 0.01$

General upward vs control: $t(58) = 0.97, p = 0.33$

Internal upward vs general downward: $t(40) = 0.57, p = 0.5$

Internal upward vs internal downward: $t(41) = 0.49, p = 0.63$

Internal upward vs general causal: $t(43) = 0.97, p = 0.34$

Internal upward vs internal causal: $t(33) = 1.69, p = 0.1$

Internal upward vs control: $t(43) = 0.01, p = 0.99$

General downward vs internal downward: $t(51) = 1.15, p = 0.26$

General downward vs general causal: $t(59) = 1.7, p = 0.09$

General downward vs internal causal: $t(48) = 1.23, p = 0.4$

General downward vs control: $t(58) = 0.63, p = 0.45$

Internal downward vs general causal: $t(54) = 0.51, p = 0.61$

Internal downward vs internal causal: $t(42) = 2.42, p = 0.02$

Internal downward vs control: $t(58) = 0.63, p = 0.53; t(53) = 0.52, p = 0.61$

General causal vs internal causal: $t(51) = 3.05, p < 0.01$

General causal vs control: $t(61) = 1.05, p = 0.3$

Internal causal vs control: $t(50) = 1.88, p = 0.07$

Emotion

Upward vs downward: $t(102) = 2.72, p < 0.01$

Upward vs causal: $t(100) = 0.47, p = 0.64$

Upward vs control: $t(69) = 1.56, p = 0.12$

Downward vs causal: $t(99) = 3.32, p < 0.01$

Downward vs control: $t(75) = 4, p < 0.01$

Causal vs control: $t(61) = 1.22, p = 0.23$

General upward vs internal upward: $t(32) = 0.5, p = 0.62$

General upward vs general downward: $t(53) = 1.53, p = 0.13$

General upward vs internal downward: $t(43) = 2.24, p = 0.03$

General upward vs general causal: $t(61) = 0.9, p = 0.37$

General upward vs internal causal: $t(40) = 0.28, p = 0.21$

General upward vs control: $t(60) = 1.74, p = 0.09$

Internal upward vs general downward: $t(40) = 1.65, p = 0.11$

Internal upward vs internal downward: $t(41) = 2.23, p = 0.03$

Internal upward vs general causal: $t(31) = 0.14, p = 0.89$

Internal upward vs internal causal: $t(37) = 0.23, p = 0.82$

Internal upward vs control: $t(33) = 0.78, p = 0.25$

General downward vs internal downward: $t(50) = 0.75, p = 0.26$

General downward vs general causal: $t(51) = 2.32, p = 0.01$

General downward vs internal causal: $t(47) = 1.6, p = 0.12$

General downward vs control: $t(54) = 2.99, p < 0.01$

Internal downward vs general causal: $t(41) = 2.97, p < 0.01$

Internal downward vs internal causal: $t(44) = 2.25, p = 0.03$

Internal downward vs control: $t(44) = 3.57, p < 0.01$

General causal vs internal causal: $t(39) = 0.47, p = 0.64$

General causal vs control: $t(60) = 0.9, p = 0.37$

Internal causal vs control: $t(42) = 1.21, p = 0.23$

Appendix 4 Experiment 7 content effects

Biconditional interpretations

Exercise vs genes: 53% vs 58%, Wilcoxon's $z = 0.62$, $n = 25$, $p = 0.53$

Exercise vs childhood: 53% vs 51%, Wilcoxon's $z = 0.28$, $n = 25$, $p = 0.78$

Exercise vs solar: 53% vs 59%, Wilcoxon's $z = 0.91$, $n = 25$, $p = 0.36$

Exercise vs reactivity: 53% vs 55%, Wilcoxon's $z = 0.47$, $n = 25$, $p = 0.64$

Genes vs childhood: 58% vs 51%, Wilcoxon's $z = 0.70$, $n = 25$, $p = 0.49$

Genes vs solar: 58% vs 59%, Wilcoxon's $z = 0.38$, $n = 25$, $p = 0.70$

Genes vs reactivity: 58% vs 55%, Wilcoxon's $z = 0.76$, $n = 25$, $p = 0.44$

Childhood vs solar: 51% vs 59%, Wilcoxon's $z = 1.10$, $n = 25$, $p = 0.27$

Childhood vs reactivity: 51% vs 55%, Wilcoxon's $z = 0.57$, $n = 25$, $p = 0.57$

Solar vs reactivity: 59% vs 55%, Wilcoxon's $z = 0.59$, $n = 25$, $p = 0.57$

Conditional interpretations

Exercise vs genes: 28% vs 14%, Wilcoxon's $z = 2.48$, $n = 25$, $p = 0.01$

Exercise vs childhood: 28% vs 25%, Wilcoxon's $z = 0.71$, $n = 25$, $p = 0.48$

Exercise vs solar: 28% vs 19%, Wilcoxon's $z = 1.87$, $n = 25$, $p = 0.06$

Exercise vs reactivity: 18% vs 35%, Wilcoxon's $z = 1.48$, $n = 25$, $p = 0.14$

Genes vs childhood: 14% vs 25%, Wilcoxon's $z = 2.34$, $n = 25$, $p = 0.02$

Genes vs solar: 14% vs 19%, Wilcoxon's $z = 1.69$, $n = 25$, $p = 0.09$

Genes vs reactivity: 14% vs 35%, Wilcoxon's $z = 3.21$, $n = 25$, $p < 0.01$

Childhood vs solar: 25% vs 19%, Wilcoxon's $z = 1.24$, $n = 25$, $p = 0.21$

Childhood vs reactivity: 25% vs 35%, Wilcoxon's $z = 1.82$, $n = 25$, $p = 0.07$

Solar vs reactivity: 19% vs 35%, Wilcoxon's $z = 2.71$, $n = 25$, $p < 0.01$

Appendix 5 List of publications

Published papers:

McEleney, A. & Byrne, R. M. J. (1998). Spontaneous counterfactual thoughts and explanations. In J. Dunnion, G. O'Hare, S. O Nuallain, R. Reilly & B. Smyth (Eds.), *Artificial Intelligence and Cognitive Science '98* (pp. 205-210). Dublin, Ireland: University College Dublin. **(summary of selected data from Experiment 1)**

McEleney, A. & Byrne, R. M. J. (1999). Consequences of counterfactual reasoning and causal reasoning. In S. Bagnara (Ed.), *European Conference on Cognitive Science '99* (pp. 199-205). Siena, Italy: University of Siena. **(summary of selected data from Experiment 6)**

McEleney, A. & Byrne, R. M. J. (2000). Counterfactual thinking and causal explanation. In J. García-Madruga, N. Carriedo & M. J. González-Labra (Eds.), *Mental Models in Reasoning* (in press). Madrid, Spain: UNED. **(summary of selected data from Experiment 1)**

Papers submitted for publication:

McEleney, A. & Byrne, R. M. J. (2000). Counterfactual thinking and causal explanation. *Manuscript submitted for publication.* **(Experiments 1, 2 and 3)**

McEleney, A., Byrne, R. M. J. & Girotto, V. (2000). Counterfactual and causal thinking: Consequences for prediction, prevention and emotions. *Manuscript submitted for publication.* **(Experiments 4, 5 and 6)**

McEleney, A. & Byrne, R. M. J. (2000). Counterfactual and causal reasoning: Inferences about prediction and prevention. *Manuscript submitted for publication.* **(Experiments 7, 8 and 9)**