

The Psychology of Counterfactual Thinking

Edited by David R. Mandel,
Denis J. Hilton, and Patrizia Catellani

First published 2005
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Ave, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group

© 2005 David R. Mandel, Denis J. Hilton and Patrizia Catellani
selection and editorial matter; the contributors their contributions

Typeset in Garamond by Wearset Ltd, Boldon, Tyne and Wear
Printed and bound in Great Britain by MPG Books Ltd, Bodmin

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data

A catalog record for this book has been requested

ISBN 0-415-32241-3

Introduction

One of the great debates in the history of science concerned the reversibility of time. Classical dynamicists such as Isaac Newton viewed time as essentially reversible. Simultaneously invert the velocities of all the parts of a system and the system goes backwards in time, much like playing a movie in reverse creates the visual impression of backward time travel. The science of heat, thermodynamics, told a different story. According to the second law, the entropy of the universe was always increasing. The universe as a whole was aging irreversibly, and much like the constituents of a burnt piece of paper could never be reassembled into a clean, white sheet, time cannot go backwards either in practice or in principle.

In many ways, people's subjective experience of time reconciles these disparate perspectives. Most individuals (at least in Western cultures) are aware of a forward direction to "real" time, moving from the past through the present to the future. In this type of real time, people are always literally "in the present." However, in the present, people can do a remarkable thing. They can travel forwards or backwards in subjective time. Indeed, even at the present moment, one can imagine oneself in the past thinking about how some event in the further past might have been different! Likewise, one can imagine oneself in the future replaying a "past" episode that, in fact, has never happened yet.

According to Endel Tulving, this ability for "mental time travel" is largely due to the fact that humans have episodic memory, a capacity to recollect events experienced in one's past. Without this cognitive capacity, Tulving has argued, humans would be unable to form a stable concept of "the self" over time and, even more significantly, the myriad of human cultures that have existed and that exist today would never have been able to evolve. The story of human cognition is all the more remarkable because not only do humans have the ability to travel in subjective time, they also have the ability to effect changes to the historical record as they time-travel — to imagine actions with consequences in these (at least seemingly) possible worlds. The capacity for humans to explore and be influenced by the counterfactual worlds they construct is a truly outstanding evolutionary feat — one that has propelled our species far beyond even the most formidable

2 Introduction

powers of retrospection. The present volume explores the psychological bases of this remarkable feat of human cognition: *counterfactual thinking*.

Strictly speaking, counterfactuals refer to thoughts or statements that include at least some premises believed to be contrary to fact. According to this broad definition, counterfactuals do not require a temporal reference. For instance, one might say (counterfactually), "If all circles were squares, then all spheres would be cubes." The "contrariness to fact" aspect of counterfactuals has long been of interest to logicians, such as Nelson Goodman, David Lewis, and Richard Stalnaker, who have sought to explain how knowledge could be derived from false conditional premises. Psychologists, on the other hand, have paid greater attention to counterfactual thinking that focuses specifically on how *the past* might, could, would, or even should have turned out differently. These researchers have been intrigued by the psychologically compelling nature of *what if* and *if only* thoughts, and the propensity for people to mentally time-travel. They have directed their attention to two broad questions. First, what are the affective, motivational, cognitive, and social determinants of counterfactual thinking? Second, what are the functional and psychological consequences of counterfactual thinking? The contributions to this volume bring together a collection of contemporary theoretical insights and descriptions of recent empirical research that bears on these two overarching questions.

Traveling backward in time, the sustained attempt to address these questions can be traced back to 1982 when Danny Kahneman and Amos Tversky published a brief, but thought-provoking, chapter on what they called the "simulation heuristic." There they proposed that, in addition to using the availability heuristic to arrive at judgments (namely, a strategy by which the frequency or likelihood of an event is judged by the ease with which similar instances could be recalled from memory), people often rely on a more constructive process in which they mentally simulate a model of a temporally extended episode and then examine its contents and implications. They proposed that such models are often constructed with the aim of examining how a past outcome might have been "undone" or, more generally, what the consequences of a slight change to the historical record would likely have been (i.e., an intuitive "what if?" analysis).

Kahneman and Tversky offered a number of key proposals that set the research agenda for many years to come. One idea was that counterfactual simulations are normality-restoring. People tend to undo outcomes that they perceive as abnormal. Moreover, they tend to undo abnormal outcomes by mentally mutating antecedents that they similarly perceive as being abnormal under the circumstances. Rarely do people undo events by making what Kahneman and Tversky called "uphill changes" – mutations that involve mentally deleting normal antecedents or mentally inserting abnormal ones. Another important idea that Kahneman and Tversky advanced was that the ease of undoing had important consequences for how people responded emotionally and judgmentally to the actual events. Thus, a flight missed by five

minutes will likely evoke more disappointment and perhaps more self-recrimination than a flight missed by thirty minutes because it is easier to imagine having been on the flight in the former case than in the latter. The possibility of "making the flight" seems "closer" to reality when the flight was "just missed" than when it was missed a half hour ago.

Both of these key proposals, among others, were elaborated in Danny Kahneman and Dale Miller's 1986 exposition of *norm theory*. According to norm theory, judgmental and affective reactions to events are largely influenced by the standards of comparison that they recruit. These cognitive reference points, Kahneman and Miller proposed, were not merely based on a priori assessments of an event's likelihood, but rather were based on alternatives that were "post-computed" – that is, mentally constructed on the fly in response to feelings of surprise that often accompany expectancy violations. The notion of "norm restoration" had a profound influence on researchers who sought to explain the cognitive rules by which people mentally reconstructed the past using counterfactual thought experiments. Many of these "mutability constraints," such as preference for mutating exceptional rather than routine antecedents, actions rather than inaction, and proximal rather than distal events, were understood in terms of this "normality principle." Moreover, Kahneman and Miller's *emotional amplification hypothesis*, which states that emotional responses to events are contrasted away from the affective direction of the counterfactual reference point, highlighted an important fact: counterfactuals tend to have a direction, either being *upward* (i.e., better than reality) or *downward* (i.e., worse than reality), but rarely are they horizontal (i.e., just "different").

Spurred by these seminal contributions, the next decade witnessed the first major wave of sustained research and theory development. Many of the core insights from this period were described in a thought-provoking 1995 volume edited by Neal Roese and Jim Olson and entitled *What Might Have Been: The Social Psychology of Counterfactual Thinking*. Perhaps foremost among these developments was the idea that counterfactual thinking has a functional basis. On the one hand, it was proposed that upward counterfactual thinking served a preparatory function by allowing individuals to explore the causal bases of past outcomes, especially those that deviated from expectation and that had negative consequences. On the other hand, downward counterfactual thinking was proposed to regulate affective responses by making people feel better about reality upon realizing how "it could have been worse."

Also taking shape during this first wave of research was the influential idea that counterfactual thinking plays a key role in how people select the causes of past events. Again, the idea can be traced back to Kahneman and Tversky's seminal chapter, although similar (and, indeed, more elaborate) proposals had been made by the legal philosophers H.L.A. Hart and A.M. Honoré and by the philosopher J.L. Mackie. The basic proposal was that one of the ways that people would assign causal status to an outcome would be

to run a counterfactual test in which the proposed cause was negated in simulation. If the simulation ran on in such a way that the outcome was subsequently undone, it would lend support to the proposed cause. If the outcome remained intact in the simulation, then the proposed cause could be ruled out by this kind of thought experiment.

In the decade since the publication of Roese and Olson's edited volume, a second wave of counterfactual research has emerged. Although this work continued to be influenced by earlier theoretical assumptions and over a decade's worth of empirical research, a number of important new themes were evident. First, the methodology for doing counterfactual thinking research steadily matured. The use of "vignette studies," cleverly pioneered by Kahneman and Tversky, had been a favored method for exploring the determinants and consequences of counterfactual thinking in a substantial number of later studies. Over time, however, some researchers started to move beyond the use of scenario studies by exploring the counterfactuals that people reported in response to real events that they had either experienced or had spent a considerable amount of time thinking about. The contributions to the present volume by Dhimi, Mandel and Souza (Chapter 10) and by Tetlock and Henik (Chapter 12) clearly illustrate this thrust. In the former case, the authors examine the counterfactual thoughts of hundreds of sentenced prisoners in response to their arrest, conviction and sentencing. In the latter case, the authors examine the use of counterfactual arguments by experts on world politics and history.

There were also important theoretical developments taking shape. First, the idea that counterfactual thinking influences the causal selection process faced serious challenges. Part I on "Counterfactuals, Causality, and Mental Representation" is largely devoted to contemporary views on this issue. In Chapter 1, Mandel presents an overview of research on the relation between causal and counterfactual explanation. He summarizes the major critiques to the "counterfactual thinking influences causal selection" accounts that held sway in the first wave of counterfactual research, and then he sketches a new "judgment dissociation" theory of the relation between counterfactual and causal thinking. In Chapter 2, Spellman, Kincannon and Stose propose an alternative theoretical account of the causal selection process. They posit that causal selection is based primarily on subjective assessments of probability, but they also importantly explain how counterfactual thinking, in turn, can influence those probability judgments. In Chapter 3, Hilton, McClure and Slugoski contrast two main approaches to the study of causal judgment: an ahistorical approach that describes how people discover the causal relations between different *types* of events (as in the natural sciences), and a historical approach that describes how people discover the causes of particular effect occurring within an unfolding chain of events (as in the legal case approach). Contributing to the latter approach, they offer a typology of causal chains that expands on the "causal-temporal" distinction invoked in earlier counterfactual research.

A second theoretical development came from cognitive psychologists who had long been interested in the process of mental representation. Notably, Ruth Byrne and her colleagues initiated a program of research in the mid-1990s that offered a representational account of counterfactual thinking in terms of *mental models theory* she and Phil Johnson-Laird had developed (building on earlier ideas by Kenneth Craik). According to this account, by default, people represent only the true possibilities in a given case. However, because counterfactuals prompt individuals to represent false possibilities alongside a model of what is believed to be true, these thoughts can trigger different patterns of inference than their factual counterparts. The final chapter of Part I, Chapter 4 by Walsh and Byrne, reviews the key contributions that mental models theorists have made to the counterfactual thinking literature.

As noted earlier, one of the key objectives of the second wave of counterfactual thinking research was to develop an account of the functional (and possibly dysfunctional) bases of counterfactual thinking. The consensus that emerged from earlier work was, essentially, that upward counterfactuals help people learn and better prepare for the future but make them feel bad in the process by pointing to ways that things could have been better, whereas downward counterfactuals make people feel good by showing them how bad it could have been, but don't teach them very much. Since then, however, there have been important revisions and expansions of this functionalist account. Accordingly, we have devoted Part II on "Functional Bases of Counterfactual Thinking" to three contributions that explore some of the more recent and intriguing functionalist themes.

In Chapter 5, Markman and McMullen propose that the functions of upward and downward counterfactual thinking described in earlier accounts are characteristic of an *evaluative* mode of thinking in which counterfactual representations are contrasted against their factual counterparts. According to their account, however, people also generate counterfactuals in a more experiential, *reflective* mode of thinking in which little attention is devoted to what actually happened. In this mode, they propose, upward counterfactuals can improve affect and downward counterfactuals can function as "wake-up calls" that prompt preparatory responses. In Chapter 6, Segura and Morris examine the role of counterfactual thinking in experiential learning at individual and organizational levels. Segura and Morris unpack the learning cycle into three stages — evaluating outcomes, inducing rules, and implementing actions — and they examine the role that counterfactual simulations can play at each of these distinctive stages. In Chapter 7, Galinsky, Liljenquist, Kray, and Roese strike out into new functionalist territory by proposing not only that counterfactual thinking facilitates learning, but that it also plays a crucial role in how people formulate meaning in their everyday lives. They suggest that by thinking about the many ways in which events could have happened differently, people often conclude that personally significant events must have happened as they did for "good reason."

6 Introduction

A fourth notable area of theoretical development during the second wave of counterfactual research has been on the relation between counterfactual thinking and emotion. Part III on "Counterfactual Thinking and Emotion" examines how counterfactual thinking contributes both to the experience of both luck and regret. In Chapter 8, Teigen argues that people typically consider themselves "lucky" when avoiding a disaster and "unlucky" when a small mistake or accident turns out to have disproportionate consequences. Using a number of intriguing examples, Teigen examines three different mechanisms that elicit these two feelings, and related ones such as gratitude. In Chapter 9, Zeelenberg and van Dijk review their work on regret that, for example, identifies circumstances in which people can feel more regret after inaction than after actions. They then turn to the comparative nature of regret, and show that better forgone outcomes that are similar to the actual outcome are more likely to cause regret than those that are not. This effect is principally observed in people with low social comparison orientations, as those with strong tendencies to make social comparisons with others tend to feel regret regardless of whether the forgone outcome is similar to the target outcome or not.

Part IV on "Counterfactual Thinking in the Context of Crime, Justice, and Political History" offers three examples of how the extension of counterfactual research to applied domains may not only add to the external validity and methodological rigor of this research area, but also contribute to important theoretical developments. Focusing attention on counterfactual thinking in sentenced prisoners, Dhani, Mandel and Souza (Chapter 10) expand our knowledge of the influence of counterfactuals on emotions and attributional judgments. They show that the influence of counterfactual thinking on guilt is not only due to affective contrast, but that it is also mediated by attributions of self-blame. Moreover, they show that the situational context to which counterfactual thinking refers (e.g., committing a crime, being arrested, convicted, or sentenced) has an important moderating effect on the relation between upward counterfactual availability and anger. In Chapter 11, Catellani and Milesi examine counterfactual thinking in the context of rape cases and they show how social contextual factors may influence counterfactuals and, as a consequence, social judgments. Whereas norm theory focused primarily on intrapersonal norms, Catellani and Milesi show that *social* norms triggered by the actors in a given case can also constrain the content of counterfactuals. They also show how communication-related goals and the mutual expectancies of people involved in interpreting the event (e.g., during a trial) may constrain the generation, expression, and evaluation of counterfactuals. In Chapter 12, attention shifts to the historical and political domain, where Tetlock and Henik show that well established views of reality, such as political ideologies, also constrain which counterfactuals will be entertained as plausible. Demonstrating a key function of counterfactual thinking — *belief-system defense* — Tetlock and Henik show that, in reasoning about a wide range of historical events, experts selectively invoke

second-order "even if" counterfactuals and they are more likely to challenge connecting principles when they are presented with counterfactual arguments that fly in the face of their own preferred theories than when presented with belief-confirming counterfactuals.

Over little more than two decades, research on the psychology of counterfactual thinking has made important strides. The contributions to our book highlight in different ways the methodological, empirical, and theoretical advances that have been made in recent years. Projecting forward in subjective time, we have little doubt that the next few years will continue to reveal important new insights into the psychology of counterfactual thinking.

Finally, several thanks are in order. The idea for this book emerged following a stimulating conference on counterfactual thinking that we organized in May 2001, in Aix-en-Provence, France. If it were not for the financial support of the European Association for Experimental Social Psychology (EAESP), we probably would not have been able to organize the meeting and, in turn, this book might never have been. Thus, we thank EAESP for its generous support. We should also note that most of the contributors to our book also attended the conference. We thank Peter Robinson, the editor of the Research Monographs in Social Psychology series, who encouraged us early on to submit a proposal for this book. Finally, we sincerely thank Yeliz Ali, Katherine Carpenter, Terry Clague, Joe Whiting, and the rest of the staff at Routledge and Taylor & Francis, who have done a superb job of assisting us from start to finish.

*David Mandel
Denis Hilton
Patrizia Catellani*