

Counterfactual Thinking

Neal J Roese, Simon Fraser University, Vancouver, British Columbia, Canada
James M Olson, University of Western Ontario, Canada

CONTENTS

Introduction

Typology

Psychological consequences

Determinants of counterfactual thinking

Conclusion

Mental constructions of alternatives to facts or events. These thoughts of 'what might have been' are linked to a variety of emotional and judgmental consequences.

INTRODUCTION

0504.001 The term 'counterfactual' means contrary to established facts or actual events. Counterfactual thinking typically involves imaginative speculation about alternatives to past outcomes: that is, about what might have been. Counterfactuals often (though not always) take the form of conditional propositions, containing the dual components of antecedent and consequent. In everyday cognition, counterfactual thinking usually targets personal goals and desires, such that individuals focus on actions that might have brought about particular desired ends (e.g. 'If I had studied harder, I would have earned a higher grade'). Counterfactuals can also be deployed in everyday speech as arguments ('If not for Gorbachev, the Soviet Union and the Cold War would have persisted into the twenty-first century') or invitations to further speculation and elaboration, e.g., 'What if President Kennedy hadn't been assassinated?' Counterfactuals have intrigued philosophers throughout the twentieth century because of their implications for logic and epistemology, but more recently counterfactual thinking has inspired psychological research because such thought processes influence a wide range of emotional, judgmental, and behavioural outcomes.

0504.002 The form and content of counterfactuals is limitless, and although they may conjure the bizarre and the fantastic, everyday counterfactual thinking is mundane. Indeed, an essential feature seems to be that counterfactuals preserve the integrity of the world as we know it, altering but one or two specific features, then unfurling immediate consequences against a backdrop that is essentially the

same as actuality. Thus, one might wonder how the Second World War might have unfolded had Hitler attacked and defeated the British at Dunkirk rather than allowing them to escape, but background features, such as the previous history of Europe, the power of the respective nations' armaments, and for that matter the laws of physics, remain unchanged. Given this rule of restricted alteration, a key theoretical focus has been to specify which finite features of reality are perceived to be more changeable, or mutable, as opposed to the infinite background features that remain constant within one's mind. The sections below on determinants of counterfactual thought are descriptions of these patterns.

Counterfactual thinking is a rule-bound creative act, and as such has been construed as a principal ingredient of consciousness and language. Hofstadter (1985), for example, argued that a comprehensive attempt to create artificial intelligence must include some facility for production of counterfactuals that operates in a manner similar to that of human cognition. A further elaboration of this theme is that counterfactuals are constrained by reality because they are functional; that is, they often provide useful prescriptions for how a goal might have been achieved in the past, and hence how it might yet be achieved in the future (Roese, 1994).

TYPOLOGY

Counterfactuals have been classified in two main ways: direction and structure. Direction refers to whether the counterfactual specifies a state that is better than actuality (an upward counterfactual) or worse than actuality (a downward counterfactual). Counterfactuals are also described in terms of structure of their phrasing. The counterfactual antecedent may be an addition of some feature

not in fact present (an additive counterfactual), or it may remove a feature that was present (a subtractive counterfactual). These two typologies have proved effective in delineating a variety of theoretical relations, described below.

PSYCHOLOGICAL CONSEQUENCES

Causation

0504.005 Counterfactuals are intimately related to causal inferences. Causation may be defined as a relation between two variables (objects, states, etc.) in which one produces or generates changes in the other. A counterfactual conditional nearly always implies causation. Counterfactual conditionals denote an antecedent-consequent pair that diverges from a related, factual antecedent-consequent pair, thereby satisfying the logic of J. S. Mill's method of difference for inferring causation. For example, the observation that a match held motionless remains bereft of flame might be followed by the counterfactual supposition that 'if the match had struck a hard surface, it would have ignited.' The mental alteration of but one feature of actuality (striking as opposed to not striking the match), when accompanied by the imagined consequential variation in ignition, provides the basis for inferring that the antecedent of match strike causally influences ignition. The logic of the method of difference is the same as the covariation criterion for causation that forms the theoretical platform for many theories of causal attribution, in that counterfactuals present one datum, albeit imagined, that may be added to a set of divergent background observations. Although absence of covariation may be used to rule out causation, presence of covariation is not in itself sufficient to infer causation. Therefore, the same problems of induction that bedevil formal analyses of causation apply similarly to counterfactual reasoning (Spellman and Mandel, 1999). Nearly all psychological consequences of counterfactual thinking appear to be rooted either in this causal inference mechanism or in a contrast effect mechanism.

Contrast Effects

0504.006 In comparative judgment, the juxtaposition of one object with a second can render judgments of the features of the latter more extreme. Thus, as demonstrated in classical psychophysics experiments, an object may be judged to be heavier after holding a lighter object, a colour may be deemed darker if

set against a lighter background, and so on. Counterfactual comparisons may similarly influence emotional appraisals of specific outcomes by making them, in contrast, seem better or worse. Thus, upward counterfactuals make an actual event seem less favorable, whereas downward counterfactuals make an actual event seem more favourable. This contrast effect underlies a variety of effects of counterfactual thinking on social judgment.

Social Judgment

A variety of social judgmental consequences of counterfactual thinking have been mapped; five are detailed here. 0504.007

First, counterfactuals influence emotion, typically making emotional reactions more extreme (by way of a contrast effect) than would otherwise have been the case. Regret is an affective state predicated on upward counterfactual thinking and is the subject of much research in its own right. Counterfactual-induced affective changes can then influence judgment further. For example, in responses to victimization, inferring that a victim's misfortune could easily have been averted might create greater sympathy for the victim, but also greater recommendations for monetary compensation to the victim (Miller and McFarland, 1986). 0504.008

Second, counterfactuals influence likelihood estimates in at least two ways, both rooted in the causal inference mechanism. The mental simulation of an alternative antecedent event can make future, similar events seem more likely. This would occur to the extent that the prior action is controllable and presumed to be sufficient to have brought about a favorable outcome. That is, the individual might intend to perform the action in the future to bring about a desired goal, in part because the individual infers that performing it in the past would have brought about that desired goal in the past (Roese, 1994). Counterfactuals can also make past events seem more predictable (the hindsight bias) to the extent that they clarify causal linkages, i.e. specify how an event was brought about and thus how it might have been improved or negated (Roese and Olson, 1996). For example, a student who reacts to a poor grade with the counterfactual, 'If only I had studied harder, I would have performed better', has used the counterfactual to articulate the causal power of studying to influence performance. This causal inference may then form the basis of a behavioral intention to study more thoroughly for the next examination, which then yields beliefs in the heightened probability of future success. 0504.009

0504.010 Third, and drawing directly on the previous description of heightened likelihood estimates, counterfactuals can heighten perceived control, again by way of causal inferences. To the extent that a desired event is seen to be attainable had one only acted in a certain way, it confers a belief in personal control (Nasco and Marsh, 1999). In other words, one may generalize from the specific instance of having been able to effect positive outcomes ('If I had studied harder, I would have performed better') to the beliefs regarding global personal efficacy ('I can accomplish many things with a little extra effort').

0504.011 Fourth, counterfactuals can influence decision-making. If a decision is made but an alternative decision might have brought about clearly better rewards, the resulting emotion of regret may compel changes in decision-making strategy, altering the course of subsequent behavior. Research on cognitive dissonance theory specifies conditions under which individuals alter appraisals as a function of postdecisional regret, but theory linking dissonance to counterfactuals is underdeveloped.

0504.012 Fifth, counterfactuals can make observers suspicious. If an event occurs but is surprising because it is easy to imagine it occurring differently, an observer might be more suspicious regarding ulterior goals of the actor than in cases in which it is easy to imagine the event occurring in many similar ways, even if the probability of event occurrence remains constant. Take the example of a child who loves chocolate-chip cookies: the child is permitted to have just one cookie before dinner, but is required to select the cookie with eyes closed from a jar containing one chocolate-chip cookie and nine oatmeal cookies. If the child happens to select the coveted chocolate-chip cookie, an observer might suspect that the child had peeked. If, however, the cookie jar contained ten chocolate-chip cookies and ninety oatmeal cookies, suspicion might be reduced as there are ten similar ways for the coveted cookie to be selected without intent. Even though the probability of selecting the chocolate chip cookie is identical in both cases, ease of generation of alternatives differs and results in variation in suspicion (Miller *et al.*, 1989).

DETERMINANTS OF COUNTERFACTUAL THINKING

Activation

0504.013 When does counterfactual thinking occur? A principal trigger is negative affect resulting from an undesirable outcome (Sanna and Turley, 1996). When things go wrong, people often ruminate about how the outcome could have been avoided. Thus, thoughts about 'what might have been' are more common following defeats than victories, failures than successes, and penalties than rewards. A second activator of counterfactual thinking is surprise resulting from an unexpected event. Unexpected occurrences violate implicit predictions and thereby attract attention, which induces consideration of why the outcome occurred. A third trigger of counterfactual thinking is a near miss, or an event that almost occurred. When something nearly happens, it seizes the perceiver's imagination. An athlete who finishes second by a hair's breadth in a 100 m race is likely to experience vivid thoughts about the counterfactual outcome of winning, whereas finishing a distant second evokes fewer thoughts of hypothetical victory.

0504.014 These triggers of counterfactual thinking correspond to situations where this activity is most useful. As noted earlier, counterfactual thinking provides causal information about an outcome. What kinds of outcomes are most important to understand? Negative outcomes demand comprehension for survival reasons (prevention). Unexpected outcomes, by definition, indicate failures of prediction. Outcomes that almost occurred might occur in the future. Thus, these triggers reflect adaptive coping and support a functional view of counterfactual thinking (Roese, 1994).

Content

0504.015 Of the infinite number of possible alternatives to reality, which does the mind select for consideration? That is, what are the typical contents of mental reconstructions? Researchers have identified several qualities that render events or antecedents more mutable. As Hofstadter (1985: p. 239) argued, there are natural 'fault lines' of the mind along which reality is cognitively cleaved.

0504.016 One variable influencing the content of counterfactual thoughts is the normality of the antecedents to an event (Kahneman and Miller, 1986). When considering alternative possibilities, perceivers often focus on unusual things preceding an outcome, rather than routine aspects of the situation,

with the mental reconstruction transforming the unusual antecedent into a more normal form. For example, a student who spends less time than usual studying for an examination and performs poorly is likely to think, 'If only I had studied more, I would have done better', even though many other mutations are also theoretically possible (e.g. 'If only the test had been easier').

0504.017

A second feature of antecedents that increases the probability that they will be selected for counterfactual mutation is controllability. Perceivers are more likely to mutate controllable than uncontrollable aspects of a situation. For example, following a car accident at high speed on a slippery winter road, the driver is more likely to think 'If only I had driven more slowly' than 'If only it hadn't been snowing'. Serial position also influences counterfactual content. Typically, the most recent antecedents are mutated. A missed shot at the buzzer of a one-point loss in basketball is more likely to be altered than preceding misses, even though all misses were equally responsible for the outcome. If, however, several antecedents constitute a causal chain, then early events are likely to be mutated. For example, if a truck blows a tire and hits a car, which then runs into a school bus injuring some children, perceivers will think 'If only the truck hadn't blown a tire', rather than 'If only the car hadn't hit the bus'.

CONCLUSION

0504.018

Counterfactual thinking, or thoughts of alternatives to past outcomes, is a common feature of everyday mental life. It exerts a variety of effects on emotion and judgment, and is thought to do so primarily through underlying mechanisms rooted in causal inference effects or contrast effects.

References

- Kahneman D and Miller DT (1986) Norm theory: comparing reality to its alternatives. *Psychological Review* **93**: 136–153.
- Miller DT and McFarland C (1986) Counterfactual thinking and victim compensation: a test of norm theory. *Personality and Social Psychology Bulletin* **12**: 513–519.
- Miller DT, Turnbull W and McFarland C (1989) When a coincidence is suspicious: the role of mental simulation. *Journal of Personality and Social Psychology* **57**: 581–589.
- Nasco SA and Marsh KL (1999) Gaining control through counterfactual thinking. *Personality and Social Psychology Bulletin* **25**: 556–568.

- Roese NJ (1994) The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology* **66**: 805–818.
- Roese NJ and Olson JM (1996) Counterfactuals, causal attributions, and the hindsight bias: a conceptual integration. *Journal of Experimental Social Psychology* **32**: 197–227.
- Sanna LJ and Turley KJ (1996) Antecedents to spontaneous counterfactual thinking: effects of expectancy violation and outcome valence. *Personality and Social Psychology Bulletin* **22**: 906–919.
- Spellman BA and Mandel DR (1999) When possibility informs reality: counterfactual thinking as a cue to causality. *Current Directions in Psychological Science* **8**: 120–123.

Further Reading

- Ferguson N, ed. (1997) *Virtual History: Alternatives and Counterfactuals*. London: Picador.
- Gilovich T and Medvec VH (1995) The experience of regret: what, when, and why. *Psychological Review* **102**: 379–395.
- Harris PL, German T and Mills P (1996) Children's use of counterfactual thinking in causal reasoning. *Cognition* **61**: 233–259.
- Kahneman D and Tversky A (1982) The simulation heuristic. In: Kahneman D, Slovic P and Tversky A (eds) *Judgment Under Uncertainty: Heuristics and Biases*, pp. 201–208. New York: Cambridge University Press.
- Lewis D (1973) *Counterfactuals*. Cambridge, MA: Harvard University Press.
- Roese NJ and Olson JM, eds (1995) *What Might Have Been: The Social Psychology of Counterfactual Thinking*. Mahwah, NJ: Erlbaum.
- Sanna LJ, Turley-Ames KJ and Meier S (1999) Mood, self-esteem, and simulated alternatives: thought-provoking affective influences on counterfactual direction. *Journal of Personality and Social Psychology* **76**: 543–558.
- Tetlock PE and Belkin A, eds (1996) *Counterfactual Thought Experiments in World Politics*. Princeton, NJ: Princeton University Press.

Glossary

- Additive counterfactual** A counterfactual structured around the addition of an antecedent that was not part of factual reality.
- Direction of comparison** A comparison target may be compared to a referent that is evaluatively better (upwards) or worse (downwards).
- Downward counterfactual** A counterfactual specifying a state that is evaluatively worse than what actually occurred.
- Hindsight bias** Tendency to judge events to be more predictable, knowable, and certain in hindsight than in foresight.
- Mutability** Extent to which a feature or attribute is perceived to be changeable; actions tend to be mutable whereas laws of physics are immutable.

Psychophysics Branch of psychology dealing with mathematical descriptions of simple perceptual judgments.

Subtractive counterfactual A counterfactual structured around the deletion of an antecedent that was part of factual reality.

Upward counterfactual A counterfactual specifying a state that is evaluatively better than what actually occurred.

Keywords: (Check)

counterfactual; regret; undoing; simulation; mutability

Encyclopedia of Cognitive Science - author queries

Article 504 [Roese & Olson]

Affiliations

Please check that authors' affiliations are correct.

Introduction

paragraph 3, sentence 2, 'Hofstadter (1985)' - not in reference list.

Determinants of counterfactual thinking

Content

paragraph 1, last sentence 'Hofstadter' - see above.