Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 58 (2008) 609–626
Journal of Memory and Language
www.elsevier.com/locate/jml

Anomalies in real and counterfactual worlds: An eye-movement investigation

Heather J. Ferguson *, Anthony J. Sanford

Department of Psychology, University of Glasgow, Glasgow, Scotland G12 8QQ, UK

Received 5 December 2006; revision received 14 June 2007
Available online 25 July 2007

Abstract

Counterfactual reasoning is valid reasoning arising from premises that are true in a hypothetical model, but false in actuality. Investigations of counterfactuals have concentrated on reasoning and production, but psycholinguistic research has been more limited. We report three eye-movement studies investigating the comprehension of counterfactual information. Prior context depicted a counterfactual world (CW), or real world (RW), while a second sentence was manipulated to create RW anomalous continuations, where events included a violation of RW knowledge, and RW congruent continuations, where the events described were congruent with RW knowledge. Results showed that RW violations can be 'neutralised' within an appropriate pre-specified CW context, and RW congruent items can lead to the experience of an anomaly following an inconsistent CW context. Importantly, there was also evidence in all three studies for early processing difficulty with RW violations regardless of prior context, indicating that a proposition is rapidly evaluated against real-world knowledge, just prior to the accommodation of a proposition into a counterfactual world representation. We discuss the results in terms of a variety of accounts of the nature of counterfactual worlds.

Keywords: Counterfactual reasoning; Eye-movements; Discourse processing; Anomalies

Counterfactual reasoning, an understanding of events that are counter to reality, or false, is an essential ingredient of our everyday cognition. Counterfactual situations are frequently depicted through language, yet surprisingly little is known of how they are processed during reading or listening. In this paper, we attempt an exploration of counterfactual processing during reading. Counterfactuals are cases of possibly valid reasoning from premises that are false in actuality (Fauconnier & Turner, 2003), and require the comparison of reality to a model-based alternative. People understand a counterfactual statement, such as, *If money grew on trees then we'd all be millionaires* by keeping in mind two possibilities from the outset: the conjecture, *money grows on trees* and *we are all millionaires*, and the presupposed facts, *money does not grow on trees* and *we are not all millionaires* (Byrne & Tasso, 1999). The counterfactual thus requires that a person represent false information that is temporarily supposed to be true. Linguistic analyses have catalogued a number of ways in which counterfactual worlds may be triggered, including modal terms such as *could*, and *might*, and *if-then* constructions. It is also known that tense influences the plausibility of counterfactual interpretation

* Corresponding author. Fax: +44 (0) 141 330 4606.
E-mail address: h.ferguson@psy.gla.ac.uk (H.J. Ferguson).

0749-596X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2007.06.007

(e.g., Cowper, 1999; Kratzer, 1991). In the present paper, we rely on *If-then* constructions that clearly can signal a counterfactual world for consideration.

There has been a very large amount of research on reasoning with counterfactuals (c.f., Byrne, 2002), and on what sort of constraints there are on the kinds of counterfactual thoughts people are likely to generate in a variety of circumstances (e.g., Byrne, 1997; Kahneman & Miller, 1986; Markman & Tetlock, 2000). Counterfactuals are ubiquitous in cognitive activities, ranging from simple imagination beyond reality, and fantasy (e.g., Sternberg & Gastel, 1989) to the exploration of possibilities in deductive reasoning (e.g., Byrne & Tasso, 1999; Johnson-Laird & Byrne, 2002). They serve important social functions, for instance in reflecting on past events with negative outcomes [the "if-only.." effect; of Kahneman and Tversky (1982); see also Byrne, 2007; Johnson-Laird and Byrne, 1991; Kahneman, 1995].

In contrast to research within the framework of reasoning and its social concomitants, there has been very little research on how counterfactuals are understood during language comprehension, for instance of what kinds of representations they set up. One approach is that of mental spaces, described by Fauconnier (1985, 1997). Mental spaces are defined as structured, incremental sets that include elements and relationships between them, with availability for new elements to be added and new interactions between the elements to be created. Mental spaces, and the relationships between them, are a way of specifying an interpretation of a discourse. According to Fauconnier, two mental spaces are produced in the case of counterfactual conditionals; one is the reality space and the other is the counterfactual hypothetical space. He sees counterfactuality as a case of forced incompatibility between these two spaces, since what is true in the counterfactual space is false in the reality space. Although Fauconnier presents some very interesting analyses of what is entailed with counterfactual worlds, his analyses do not really provide any basis for predicting how propositions are processed with respect to real world and counterfactual world spaces.

A similar psychological account of reasoning, the mental model theory, has been proposed (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). This theory has a "core" *extensional* account of conditionals, making a conditional 'if p then q' logically equivalent to 'not-p or q'. Consequently, in the case of counterfactual conditionals, it is proposed that both factual and counterfactual possibilities are represented by the reader. An alternative view that has gained increasing interest was initiated by Ramsey (1931), who proposed that when comprehending a conditional statement, people "hypothetically add p to their stock of knowledge and argue on that basis about q". This practice is commonly known as the *Ramsey test*. Recent literature has challenged the mental model theory (Evans & Over, 2004; Evans, Over, & Handley, 2002). As an alternative, authors suggest a suppositional theory where a conditional of the form "if p then q" directs attention to possibilities following from p, and *not* to "not-p or q" possibilities. Therefore, counterfactual statements should be evaluated with respect to suppositional or hypothetical possibilities first.

The present paper is an attempt to examine the role played by real-world (factual) knowledge, and inferences from counterfactual worlds during on-line comprehension of simple statements. We illustrate the problem with a simple example. In the real world, it is anomalous to say (1):

(1) If the cat is hungry, the owner could feed the cat carrots and it would happily gobble them down.

If a counterfactual world is set up through a statement like (2), then statement (1) is not anomalous with respect to that counterfactual world, although it remains so with respect to the real world.

(2) It would be great if cats were vegetarian.

According to the mental model theory, people have to keep in mind both the conjecture *If cats were vegetarians then (1)*, and the presupposed facts that *cats are not vegetarian and do not like carrots* (e.g., Byrne & Tasso, 1999). Similarly, according to Fauconnier (1985, 1994), two spaces reflecting the real and the counterfactual world are set up. However, according to the suppositional theory, people would hypothetically suppose that *cats are vegetarians* and then judge their degree of confidence in *feeding cats a bowl of carrots* given that supposition. If the conditional probability was high, they would confidently believe the statement and accept it. Conversely, if the conditional probability was low, they would have doubts about the statement and either reject it or initiate further inferences in order to determine whether it could be consistent with the counterfactual scenario. Although it is undoubtedly true that ultimately a proper appreciation of counterfactuals requires knowledge about both real and counterfactual worlds, it is unclear whether the two would both be present simultaneously in a representation of the discourse model associated with the introduction of a counterfactual situation, or whether there would be a sequential process, in which the counterfactual was temporarily accepted as the true world, and sometime later the consequences of this are tested against the true world for inference. This immediately gives rise to a processing question: can something that is anomalous given our real-world knowledge be "neutralised" as an anomaly if it is consistent within a pre-specified counterfactual world context? According to the model theories of rea-