Reasoning Counterfactually in Chinese: Picking up the Pieces

David Yeh (d-yeh1@northwestern.edu)
Department of Psychology, Northwestern University
2029 N. Sheridan Road Evanston, IL 60208-2710 USA

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, Northwestern University
2029 N. Sheridan Road Evanston, IL 60208-2710 USA

Abstract

We review the controversy concerning whether the lack of a clear counterfactual marker in Chinese results in a deficiency in counterfactual reasoning (Au, 1983, 1984; Bloom, 1981, 1984; Liu, 1985). We describe a study in which we compared two kinds of counterfactual assertions. The results showed an accuracy advantage for English speakers over Chinese speakers when specific contextual information was required to detect the counterfactual, but not on other counterfactual sentences. Implications for language and thought are discussed.

Introduction

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference

--Robert Frost

Counterfactual thought refers to a mode of thinking that is literally contrary to fact. In practice, counterfactual thought often supports speculation as to what might have been or what could have happened had some detail or event in the past occurred differently; as in the assertion, “If I had practiced more, I might have been a rock star.”

Roese (2005) points out that counterfactual reasoning is employed in many cognitive processes. It permits the exploration of causation and, in some cases, the acknowledgement of blame; it occurs in reflection upon opportunity cost; it can lead one to experience feelings of regret; and finally it provides benchmarks and standards by which we recognize deviation from normalcy (Kahneman & Miller, 1986; Roese, 2005). The counterfactual provides us with an alternative to the current state of affairs that is grounded in logical chains of causation. It allows us to judge reality by comparing it to what could have very easily happened instead, as is demonstrated above in the excerpt from Robert Frost’s The Road Not Taken.

Despite this apparent pervasiveness, the counterfactual is undeniably a somewhat complex form of thought. As Byrne (2002) notes, counterfactual thinking requires simultaneously holding two different construals of a given state of affairs. Harris (2000) discussed counterfactual thinking as an instance of theoretical thought, distinguishing it from empirical thought. In reviewing Luria’s studies of Uzbekistani peasants in the 1930’s, he considers that their failure to deal with logical reasoning from premises stemmed in large measure from their habit of adopting the empirical mode, in which one reasons about things that are factually true in the world. Counterfactual thinking and other instances of the theoretical mode may in part be culturally learned. Indeed, some evidence suggests that mastery of the counterfactual mode occurs rather late in cognitive development (Markovitz & Vachon, 1989; Wing & Scholnick, 1986; Guttentag & Ferrell, 2004).

What are the cultural practices that foster counterfactual thinking? One obvious candidate is the linguistic forms that serve to signal a needed shift to counterfactual reasoning. For example, in English the past subjunctive can be used, as in, “Had he not gone to the forum that day, Caesar would have lived longer.” An if-clause can be added to make the counterfactual more salient, as in, “If he had not gone to the forum that day, Caesar would have lived longer.”

These forms as demonstrated above signal the need for counterfactual reasoning in English. It is natural, then, to ask whether comparable forms exist in other languages; and if not, whether there are accompanying differences in counterfactual reasoning. We can distinguish a weak and a strong view of what these effects might be. On the weak view (the signaling view) the clarity with which a standard signal indicates a shift to counterfactual reasoning allows the listener to make this shift more immediately, thereby increasing processing efficiency. This weak view posits that there is no general deficiency in counterfactual ability resulting from language, but that differences in how clearly a language marks the need for counterfactual reasoning may influence processing efficiency, particularly in nontransparent contexts.

The strong (habitual thought) view includes the assumption of the weak view that having a standard cue marking the necessity for counterfactual reasoning can increase the efficiency with which speakers of this language shift to that mode when appropriate. However, it also assumes that having this kind of convenient cue permits more frequent use of the counterfactual mode. By that same token, speakers of a language that either lacks a standard cue or employs an ambiguous cue will be less likely to engage in counterfactual reasoning. This view essentially argues that
speakers of a language with a standard counterfactual cue will simply be better at counterfactual reasoning.

The strong version of the idea that language could affect the understanding and use of the counterfactual was developed in large part by Alfred Bloom (1981, 1984). Bloom observed that unlike English with its use of the past subjunctive, Chinese lacks any syntactic, semantic or intonational cue that distinctly signals a shift to the counterfactual mode. Consider the following English example and its Chinese equivalent:

(1) “If I had gone to the movies that evening, I could not have had dinner with my mom.”

The counterfactual mode is signaled by having the verb of the suppositional clause in the past perfect tense and that of the main clause in the conditional form. The Chinese equivalent is

(2) “Ru guo wo na tian wan shan qu kan le dien ying, wo jiu bu neng gen wo ma qu chi wan fan.”

This translates roughly to the English: If that night I go watch (past particle) a movie, I then cannot accompany my mom to go eat dinner. The detailed gloss is

‘Ru guo wo na tian wan shang
‘If I that day evening
qu kan le dian ying, wo jiu bu neng
go watched movie, I then not able
gen wo ma qu chi wan fan.’
‘with my mom go eat dinner.’

In Chinese, which lacks verb inflections, a counterfactual is signaled by comparing the tense information—e.g., the past tense particle (le) after the verb—with contextual information as to whether the event actually occurred. The second clause is simply a consequence clause; it has no internal marker of counterfactuality. To detect counterfactuality, the hearer must compare the sentential assertion with context. (The Chinese ru guo, like the English if, is ambiguous; both can mark hypothetical reasoning (such as syllogistic reasoning) as well as counterfactual reasoning.)

Bloom hypothesized that because English had a specific syntactic structure for the counterfactual, its speakers may develop a corresponding schema specific to counterfactual reasoning. He further hypothesized that, because Chinese does not grammatically differentiate between counterfactuals and other implicational relations, Chinese speakers may never develop such a specialized schema, and thus may be less efficient or competent at reasoning counterfactually than English speakers.

To test his theory, Bloom (1981) ran a study using two versions of a counterfactual story about a fictional philosopher named Bier. These texts were written in English by Bloom and then translated to Mandarin Chinese by native Chinese speakers. There were two versions of the story, which differed in the salience of the counterfactual logic employed. Version 2 logically permitted both a counterfactual and a factual interpretation, thus minimizing the salience of counterfactual logic. In Version 3, the only reasonable interpretation was counterfactual, and thus this version maximized the salience of the counterfactual logic. (Version 1, used in pilot studies, will not be discussed here.) These stories in English followed the abstract structure, “A was not the case. If A had been the case, B would have been the case; C would have been the case; D would have been the case; E would have been the case.”

The striking results of these studies seemed to offer convincing support for Bloom’s hypothesis. For Bier Version 2, which afforded both a factual and counterfactual interpretation, only 6-7% of Taiwanese participants interpreted the story counterfactually, as compared to 98% of American participants given the English version of the same text. Results from Bier Version 3 indicated that increasing the salience of counterfactual logic led to an increase in counterfactual responses from Chinese speakers.

Bloom also found that Chinese-English bilinguals performed considerably better on the English version of Version 3 than Chinese monolingual speakers given the Chinese version of the same story. Bloom found this to be the most compelling evidence for the effect of language on counterfactual thought and argued that “...for many, if not most, of the bilinguals in the study, the counterfactual mode of thought remains associated in their minds with the English linguistic world, activated more readily when cognitive processing is elicited by that linguistic world rather than by their native Chinese” (Bloom, 1981, pp. 31-32).

Bloom’s paper incited considerable controversy, most notably by Terry Au. Although she was sympathetic to Bloom’s theoretical assertions, Au argued that his study itself had many serious flaws. Upon reading over the Bier story, Au (1983) found that the Chinese translations of the texts were not as idiomatically written as the original English versions, a difference that could have contributed to the lower rates of counterfactual interpretation among Bloom’s Taiwanese participants.

Au (1983) created stories that were idiomatically written and found that Chinese bilinguals had little difficulty understanding the counterfactual logic of her stories in either language insofar as they were written idiomatically. Bloom countered by arguing that Au’s stories were highly concrete and too simple to provide a good test. Au also asked native Chinese speakers to translate Bloom’s original Chinese Version 3 to an English version as unidiomatically worded as it was in Chinese. She found that English-speaking participants presented with this text showed low rates of counterfactual interpretation, much like Bloom’s Chinese participants.

Au (1984) argued more specifically that Bloom had misused two crucial Chinese adverbs of contingency that occur in counterfactuals—jiu (“then”) and cai (“then and only then”)—such that the counterfactual intention of Bloom’s assertions was not properly signaled. Au rewrote the texts and presented them to students in the Hong Kong Anglo-Chinese
secondary school system. She found that 88% responded counterfactually, as compared to only 23% of the Chinese-speaking participants in Bloom’s study. Au concluded that there was no convincing evidence that the lack of a distinct counterfactual marker in Chinese hinders its speakers’ ability to reason counterfactually.

Liu (1985) noted, however, that there were still some methodological discrepancies between the Au and Bloom studies. While Bloom’s Chinese-speaking participants had come from Taiwan, Au’s came from an Anglo-Chinese track of secondary education in Hong Kong, where English is a more prominent influence. Also, Liu noted that despite some discussion of the importance of concreteness and complexity of the stimuli, neither Bloom nor Au had manipulated these variables. Liu thus ran the reading comprehension tasks used by Bloom and Au using participants from Taiwanese secondary schools. In a separate task, she varied stimuli on an abstract-concrete spectrum using items from Paivio, Yuille and Madigan’s (1968) list of nouns rated for concreteness, imagery, meaningfulness and frequency.

In her reading comprehension trials Liu, like Au, found that Chinese speakers displayed no difficulties with reasoning in the counterfactual mode as long as texts were written idiomatically. Her analysis of the effects of concreteness and complexity of stimuli produced no significant effects. However, post-hoc analyses suggested that the intelligibility of a counterfactual statement may be influenced by topic familiarity. Chinese participants seemed to perform better when the counterfactual dealt with things encountered on a daily basis rather than things rarely encountered.

Several lines of evidence point to the importance of context for comprehension of counterfactuals in Chinese. Au (1983) noted the importance of providing the factual background to be negated by an unfamiliar Chinese counterfactual. Such information is not necessary for the processing of a familiar Chinese counterfactual nor is it necessary for the processing of both familiar and unfamiliar English counterfactuals.

Bloom (1984) also acknowledged that receivers of information might draw upon their world knowledge and familiarity with a given context in a manner independent of linguistic restraints in order to construct an enriched interpretation of that context and subsequently formulate a response to it.

We thus proposed to directly manipulate the degree to which context made it easy to determine counterfactuality. To do this we presented counterfactual statements in either transparent or non-transparent contexts. We define these terms as follows: In the transparent contexts the information needed to detect the counterfactual could be drawn from general world knowledge, including facts about well-known public figures and historical events. Such knowledge is well entrenched and highly accessible. In contrast, in the non-transparent contexts the topic or story line was relatively novel or unpredictable. Although the general area might have been well understood or even familiar to the reader, the specific events described were not; thus, the information needed to identify the assertion as counterfactual had to be derived from comparisons with the specific story context.

Consequently, for transparent stories, the suppositional clause by itself was sufficient to reveal the counterfactuality of the assertion. For example, the suppositional clause

*If antibiotics had never been discovered...*

can be understood as counterfactual even without the rest of the story. In contrast, for the non-transparent stories, a more detailed comparison with context was required. For example, one story concerned Michael, a college student whose academic diligence so alienates his girlfriend that she eventually leaves him. Here, the counterfactual suppositional clause

*If Michael had gone out with his girlfriend that night...*

can only be interpreted as counterfactual by attending to the specific context provided (since obviously college students do sometimes go out with their girlfriends and sometimes not). In all cases, the information necessary to detect the counterfactual was present; what differed was the accessibility and immediacy of that information.

Our logic here is that, on the signaling view (the weak view), Chinese speakers should be more affected than English speakers by the degree to which factuality information is readily accessible. This is because detection of a counterfactual in Chinese requires some judgment of factuality. If the judgment can be made within the supposition (by referencing general world knowledge) then processing can be quite efficient. However, when this is not possible (in the non-transparent cases), Chinese speakers must compare the sentence with the context in order to determine its counterfactuality. Thus Chinese speakers should show difficulty with non-transparent counterfactuals, relative to their performance with transparent counterfactuals and relative to English speakers in both conditions.

On the signaling view, counterfactuals whose proper interpretation is transparent without external contextual support should not be hindered by the ambiguity of Chinese grammar in expressing counterfactuals. On the strong (habitual thought) view, Chinese speakers should perform worse than English speakers on all types of counterfactuals.

As an added measure, we recorded participants’ response time to answer questions. This allowed us to assess the possibility that Chinese speakers are less efficient at counterfactual processing, as suggested by Bloom. Such an effect would be manifest as longer reaction times amongst native Chinese speakers (relative to English speakers) on non-transparent counterfactuals, and, on the strong view, on transparent counterfactuals as well.

**Methods**

**Participants**

The Chinese speaking participants were volunteers recruited from top 5 universities in the Taipei area. Ages of participants ranged from 19-27. Participants had on average 11 years of
formal English training (exposure to English training was unavoidable for this age bracket and education level, as English classes are part of the standard curriculum starting in Taiwanese middle schools). However, participants’ self reports indicated that they did not use English proficiently on a daily basis. Participants’ average self-ratings for English fluency was a 2.5 on a 7-point scale (7 being most fluent). When asked how they believed their fluency in English compared to their fluency in Chinese, participants gave an average self-rating of 1.5 (7 indicating that English fluency was on par with Chinese fluency). 84 participants were tested in total, with the results from 5 participants being omitted from analysis due to problems when testing. The data from 3 additional participants were discarded due to low accuracy (criteria elaborated below).

The English-speaking participants were 30 undergraduate students at Northwestern University, who participated in the study for class credit. The data of 3 participants were omitted due to low accuracy.

Materials and Procedure
We composed 4 experimental stories. Two concerned a topic area allowing for transparent counterfactual assertions and two for non-transparent counterfactual assertions. To clarify, our use of this label “non-transparent” is meant to indicate that the counterfactuality of the assertion in this story cannot be locally interpreted as such simply by resorting to one’s world knowledge. To ensure non-transparency, we created fictional topics. Both transparent and non-transparent stories contained information presented both factually and counterfactually. The stories had an average length of 192 words (sd = 26) in English and 315 words (sd = 45) in Chinese. The design included two within-subject variables—factuality of the statement and transparency of the content—and one between-subject variable—English or Chinese.

The structure of the stories was based on the stimuli used by Bloom and Au. The first half of the story consisted of factually presented information; this was followed by a counterfactual antecedent and four counterfactual consequences. The counterfactual statements were unique in that they were not based on one another nor were they merely counterfactual reiterations of things stated in the factual half. Participants were presented with each story and allowed as much time as necessary to finish reading it. They then advanced to a series of eight true/false statements by pressing the space bar; at this time the story would disappear from the screen. The eight statements were presented in random order. Of the eight statements, four were based on information presented factually; four were based on information presented counterfactually. Within each of these two categories, two statements were true, two were false.

Participants were instructed to identify these statements as either true or false according to the context of the story by pressing the arrow keys. This process was repeated for each subsequent story.

Four additional distracter stories were dispersed at random between experimental stories. These stories were slightly shorter than experimental stories, and were followed by only 3 true/false statements to be evaluated by participants.

The experiment was run on the Macromedia Authorware platform which recorded the accuracy with which each participant responded to the T/F statements as well as the time it took for the participant to respond to each statement. The study was presented on an IBM ThinkPad notebook for Taiwanese participants and Dell PCs for American participants at Northwestern University. All instructions were presented on this program and participants were debriefed at the conclusion of the study with a closing slide.

Participants were informed that they were participating in a study examining the effects of literary techniques on comprehension and memory. There were not informed of the bilingual dimension of the study.

Removing Outliers
We discarded one non-transparent counterfactual statement, as both American and Taiwanese participants performed significantly below chance. Further analysis and interviews with participants revealed that the question was in fact faulty.

Participants’ data were discarded on the following bases: those whose overall accuracy was more than two SDs from the mean and individual responses from the remaining participants taking longer than 10,000 ms. This resulted in discarding 6.8% of the data from Taiwanese participants and 11.5% of the data from American participants.

Results
Accuracy
In non-transparent contexts, English speakers demonstrated significantly greater comprehension accuracy for counterfactual statements than did Chinese speakers (m = 91.52% for English, m = 76.69% for Chinese). Such a difference in accuracy was not observed for counterfactual statements in transparent contexts (m = 96.09% for English, m = 95.56% for Chinese). This pattern is consistent with the signaling hypothesis, but not with the strong habitual thought hypothesis (which predicts poorer performance on all counterfactuals).

Taiwanese and American participants performed equally accurately for transparent (m = 91.78% for Chinese, m = 94.53% for English) and non-transparent (m = 91.61% for Chinese, m = 89.84% for English) factual statements; thus there was no overall difference in comprehension accuracy.

A repeated-measures ANOVA revealed a significant 3-way interaction of language, transparency and factuality, F(1, 106) = 16.27, p < .001, as well as all possible 2-way interactions between the three variables—transparency by language: F(1, 106) = 4.46, p < .05; factuality by language: F(1, 106) = 8.80, p < .01; transparency by factuality: F(1, 106) = 15.88, p < .001. The analysis also showed significant effects of language, F(1, 106) = 12.441, p < .01, and transparency, F(106, 1) = 37.387, p < .001, but no main effect of factuality.
As illustrated in the figure 1, Taiwanese and American participants differed significantly in accuracy of response only when presented with a counterfactual statement concerning a non-transparent topic area, $t (106) = 4.42$, $p < .001$. Whereas Chinese speakers in this condition performed at only 76.7% accuracy, English speakers had 91.5% accuracy. Differences in response accuracy between all other conditions were non-significant.

**Reaction Time**

Our measures of response time showed no significant difference between Chinese-speaking and English-speaking participants in the time required to respond to either factual (m = 3568 ms for Chinese, m = 3647 ms for English) or counterfactual (m = 3342 ms for Chinese, m = 3536 ms for English) statements in the transparent condition. The same similarity held true for the time required to respond to factual (m = 3404 ms for Chinese, m = 3500 ms for English) and counterfactual (m = 3855 ms for Chinese, m = 4179 ms for English) statements in the non-transparent condition.

A repeated-measures ANOVA failed to reveal a significant 3-way interaction. The ANOVA did reveal significant effects of transparency, $F (1, 106) = 10.59$, $p < .01$, and factuality, $F (1, 106) = 12.21$, $p < .01$ as well as a significant two-way interaction of transparency and factuality, $F (1, 106) = 51.14$, $p < .001$. Neither transparency nor factuality interacted significantly with language. In fact, English speakers took nonsignificantly longer than Chinese speakers to respond in all conditions.

**Discussion**

Although Chinese speakers’ overall comprehension accuracy was equal to that of English speakers, they showed a specific difficulty with counterfactuals concerning non-transparent topic areas—that is, with counterfactual statements that were not readily identifiable as such by drawing from world knowledge. For these counterfactuals, English-speaking subjects could take advantage of the presence of a clear syntactic marker. However, Chinese-speaking participants had to match the sentence with contextual information to discover that it was counterfactual.

Our findings differ somewhat from those of prior researchers. We found an accuracy disadvantage for Chinese speakers for non-transparent counterfactuals. The narrative stimuli used by Bloom and Au in their respective studies both fall into our category of non-transparent, as they are both fictional stories. Yet Liu found that native Chinese speakers displayed no difficulty on these counterfactual stories when written idiomatically. How do we reconcile these findings?

There are two major differences in our methods. First, in our study, although participants were self-paced in reading the stories, the stories disappeared once test questions appeared. In contrast, in the prior research, participants were able to reference the stories while answering test questions. This seems likely to have resulted in increased accuracy. Indeed, Vorster & Schuring (1989) argue that “the Bloom-Au response-elicitation procedure contains a systematic bias toward responses suggesting counterfactual thinking.”

The second difference is in the time spent per story. None of our subjects took longer than 20 minutes to complete our task, which required them to read 4 experimental stories (of comparable length to the Bloom’s and Au’s stories, as well as 4 additional distracter stories only slightly shorter in length), and to answer a total of 44 T/F questions. In comparison, Liu, the only experimenter to make a note of the time her participants spent, states that most of her participants needed 25 minutes to complete 2 stories and answer a total of 4 multiple-choice questions.
We emphasized in our instructions that the task was being timed, and participants may have felt some pressure to complete the task quickly. In addition, participants could not re-read the text once they began answering questions. These two factors may have resulted in lower accuracy on the non-transparent counterfactuals than was found by Au and Liu.

An interesting follow-up to these studies would be to limit the time allowed for reading the story. If, as proposed, Chinese speakers need more extensive contextual processing, they should show an even greater relative disadvantage in accuracy than was found here. It would also be valuable to run our studies using the look-back methods of Bloom, Au and Liu while recording processing time. Such a design would allow us to assess the possibility of a speed-accuracy tradeoff. When allowed to consult the story as needed during the test phase, Chinese speakers’ accuracy should increase for non-transparent counterfactuals, but with a corresponding increase in processing time.

Our data provide support for the signaling hypothesis, that counterfactual reasoning amongst Chinese speakers will be compromised in some situations—namely, those in which general world knowledge is insufficient to indicate the counterfactual nature of the premise. However, our data do not support the strong conclusion that the lack of a standard signal in Chinese results in less practice and therefore in lower fluency in counterfactual reasoning. Chinese speakers showed no deficiency with transparent counterfactuals, as would have been expected had they suffered from a general deficiency in counterfactual reasoning.

Our data do not support a strong Whorfian position that language determines the ability to reason counterfactually. Rather, we conclude that Chinese speakers may be disadvantaged when counterfactuals must be detected with respect to specific current context. But this conclusion has cognitive consequences of its own. For if detecting a counterfactual in Chinese requires more active processing of context, and therefore more cognitive resources than are needed in English, then (1) the likelihood of missing a counterfactual is greater (as in our results) and (2) detection of counterfactuality may be delayed, resulting in the need for reprocessing previous information, with a concomitant risk of other kinds of inaccuracies. These results thus suggest a possible effect of language on the efficiency of everyday counterfactual processing.

Acknowledgments

This study was supported by NSF-ROLE award 27010NW/M1 // REC-0337360 and by the CLC group and URGC at Northwestern University. We thank the Cognition and Language group for intellectual support, Stefan Kaufmann and Caitlin Fausey for linguistic and conceptual insights, Kathleen Braun for help in conducting the study, Ingrid Ke for assistance in translations, and Yeh Ling, Xiao Ru, A Di and Wei Lian for assistance in recruiting Taiwanese participants.

References


Byrne, R.M.J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences, 6*, 426-431.


