

Correcting Causal Explanations in Memory*

Hollyn M. Johnson

University of Michigan
330 Packard Rd., Rm. 220B
Ann Arbor, MI 48104
hollyn.johnson@um.cc.umich.edu
(313) 764-4253

Colleen M. Seifert

University of Michigan
330 Packard Rd., Rm. 214D
Ann Arbor, MI 48104
seifert@umich.edu
(313) 763-0210

Abstract

Several lines of research have suggested that information previously integrated into memory can influence inferences and judgments, even when more recent information discredits it. A first experiment tested the prediction that information providing causal structure, versus being mentioned but otherwise unintegrated into the account, would lead to more influence, and found that subjects used both discredited and valid information affording causal structure to make inferences, but not incidentally mentioned information with the same content. Experiment 2 found that when a plausible causal alternative accompanied the correction, subjects showed less influence from the discredited information than when the correction simply negated earlier information. The findings suggest that the continued influence of discredited information depends on the causal structure it affords.

In the course of understanding an event, one may often learn "facts" that later turn out to be false or unfounded; ideally, one would want to then eliminate any effects of this discredited information on future reasoning and understanding processes. However, many previous studies (e.g., Ross, Lepper, & Hubbard, 1975; Wyer & Budesheim, 1987) have found that despite instructions to disregard certain information, it still results in continued influence on subjects' judgments and inferences, as compared to subjects who never received that information. This continued influence effect occurs despite evidence that subjects do notice the discrediting (Carretta & Moreland, 1983; Ross, et al., 1975; Wilkes & Leatherbarrow, 1988), and that they do make the connection between the disregard

instruction and the information it refers to (Johnson & Seifert, 1992; Wilkes & Leatherbarrow, 1988). In addition, while some studies have suggested that the continued influence results from a failure to successfully edit pre-correction inferences based on the eventually-discredited information (Anderson, Lepper, & Ross, 1980; Ross, et al., 1975; Wyer & Budesheim, 1987), Johnson and Seifert (1992) demonstrated that the continued influence effect can also arise due to inferences that subjects could have made only *after* the correction, when the information had already been discredited.

Why might discredited information continue to influence later reasoning? One possible explanation is that subjects may have used the information being discredited to construct a connected causal chain (Van den Broek, 1990), providing a basis for further inferences about the account (Graesser & Clark, 1985; Haviland & Clark, 1974; Kintsch, 1988). When subjects have no alternate way to structure an account coherently, they may continue to rely on the discredited information and base further inferences on it. In making an inference, one may search one's representation of the account for an antecedent that satisfies causal constraints such as sufficiency and operativity (Trabasso & van den Broek, 1985), and may find the account hard to structure or extend without using the discredited information. Thus, one may retrieve the discredited information and use it because it provides connections that structure the account, which might otherwise remain fragmented. If this explanation holds, less influence from discredited information would be expected when a correction presents an alternative that would also provide causal structure to the account; one could incorporate this alternative and still maintain a coherent understanding.

However, an alternate hypothesis is that the sheer *availability* of discredited information in memory can promote inferences consistent with it. Mentioning this information may activate key concepts in memory, resulting in inferences based on these concepts, rather than on other alternatives, due to their increased availability (Tversky & Kahneman, 1973). In contrast, control subjects who do not hear mention of the discredited information would not have these concepts

* This work was supported by the Office of Naval Research under contract N00014-91-J-1128 to the University of Michigan. Thanks to Jennifer Coe, Andrew Jeung, and Srinika Narayan for assistance in conducting the experiments. Thanks also to Hal Arkes for comments on an earlier draft.

activated. Therefore, they would be less likely to produce inferences consistent with them at the assessment.

To examine the causal role versus the availability hypothesis, two experiments were conducted, using a series of reports on a fire investigation, taken from Wilkes and Leatherbarrow (1988).¹ Experiment 1 manipulated whether the reports presented target information as causally implicated, or incidentally mentioned it within the reports' context, but without causal implications. If influence occurs due to the information's role in the account's causal structure, rather than because of mere activation in memory, then materials that mention the information in a non-causal fashion should not promote inferences consistent with it. Experiment 2 examined the causal role hypothesis further by testing whether subjects show less continued influence from discredited information when the correction replaces it with a plausible causal alternative, rather than just negating previous information.

Experiment 1: Availability vs. Causality

Experiment 1 tested whether the continued influence effect depends on the role the information plays in the account, or whether simply having it available in memory is sufficient to account for the effect. First, in a discredited causal reference group, subjects read that a short circuit occurred near a closet reportedly containing volatile materials. Here causal inferences are likely because conditions of operativity and sufficiency (Trabasso & van den Broek, 1985) are met. Subjects later learned that the closet did not contain volatile materials. In the second group, a causal reference condition, subjects received the same initial information but did not receive the later correction. The third condition was a noncausal reference group, which presented the volatile materials within the context of the account, but as products in a store across the street from the fire location, so the information did not meet conditions conducive to involvement in causal inferences, such as operativity and sufficiency (Trabasso & van den Broek, 1985). The presence of volatile materials in a store across the street is not sufficient, in the context of the reports, to account for the warehouse fire; one would need to make many inferences beyond the story context (e.g., that someone moved the materials into the warehouse, that someone intended this, etc.) to establish these materials in a causal role. For this group also, the series of reports did not include a correction.

¹The presentation of the experiments reported here has been abridged. Complete descriptions of all experimental methods and results are available in Johnson and Seifert (1993).

Method. Sixty-one University of Michigan undergraduates received a booklet of reports and were instructed to read through it at their own pace, and that they would be asked to recall the information later. The materials were modified versions of a series of reports used by Wilkes and Leatherbarrow (1988), describing the investigation of a warehouse fire. The series consisted of 13 individual messages, each 2-4 sentences long (see Appendix A). For the discredited causal reference group, the fifth message stated that the closet contained volatile materials (oil paint and pressurized gas cylinders), and the correction appeared as the eleventh message. For the causal reference group, Message 5 mentioned the volatile materials stored in the closet, and in Message 11, a police investigator stated that several firefighters had been released from the hospital. In the noncausal reference group, Message 3 stated that the fire department was called by the manager of a store that sold oil paint and pressurized gas cylinders, located across from the warehouse. In this group, Message 5 stated that the closet in the warehouse was empty, and Message 11 was the same as in the causal reference group. Thus, the noncausal reference group saw the volatile materials mentioned within the context of the story, but as a peripheral and not integral fact.

After completing the reading, each subject took a free recall test, where subjects wrote a free recall of the reports' contents and were asked what was responsible for the fire. Following a ten minute distractor task, a memory questionnaire adapted from Wilkes and Leatherbarrow (1988) was presented. This included ten questions on facts directly presented in the messages, ten other questions requiring the subjects to make inferences about the event (such as, "What was the possible cause of the toxic fumes?" and "What could have caused the explosions?"), and two final questions assessing whether subjects were aware of any corrections or contradictions in the series. For the noncausal reference group, the memory questionnaire contained an additional question about the contents of the store across the street, appearing after the two contradiction questions.

Results and Discussion. A coder who was blind to the experimental conditions scored the responses. Responses consistent with the discredited information -- believing that the warehouse contained carelessly stored volatile materials that caused or contributed to the fire -- encompassed references using key words from the discredited message (e.g., oil, paint, gas(es), cans or cylinders), mentions of the closet itself without indications that it was empty, and attributions of carelessness or negligence. The mean number of inferences consistent with a volatile materials theme is shown in Table 1.

Table 1. Number of inferences consistent with a volatile materials theme, by group.

	Causal Reference	Discrediting Causal	Noncausal Reference
Volatile materials inferences	3.88	3.40	1.13

A post-hoc comparison showed a significant difference between the two causal reference groups, which heard about those materials, and the group that did not, $t(57) = 5.07, p < .001$, but no significant difference between the discredited causal reference and causal reference groups. Overall, 64.5% of the subjects made at least one direct and uncontroverted reference to the volatile materials. However, when directly asked what was responsible for the fire, only 13.1% of the subjects attributed it to the volatile materials.² In the discredited causal reference group, 80% of the subjects recalled the correction, and in the noncausal reference condition, 80% of the subjects correctly recalled the incidental information about the store's contents. The groups did not differ in summary recall or in fact recall ($F < 1.2$ and $F < 1$, respectively).

The results of Experiment 1 provide evidence that discredited information influences inferences because it affords causal structure for the account, and not simply because its key concepts remain activated and are therefore easily retrievable. In the noncausal reference group, 80% of the subjects recalled the information about the store's contents in response to a direct question, which suggests that it was available for retrieval and use. However, these subjects failed to apply it in making further inferences. In contrast, in the groups mentioning the volatile materials within a causal context, subjects showed a substantial number of volatile-materials inferences. The continued influence that discredited information has on inferences appears to depend on the role the information plays within the context of the account; that is, information that affords causal structure will become involved in further inferences about causes within the account, whether the information is discredited later or not. Thus, the continued influence effect is not simply due to the information being mentioned and therefore more available in memory for use in inferences.

However, one might expect that when a correction provides a causal alternative, rather than simply negating earlier information, subjects would show less influence from the discredited information. When information is negated, one is left with many aspects of the account that are unaccounted for. However, if one

²The rest of the responses attributed the fire to such causes as faulty wiring and arson, or were left blank.

also hears an alternate cause, one has a concretely stated alternative that can play the same causal role as the discredited information. Experiment 2 will examine whether a new causal alternative that accompanies a negation can successfully mitigate the influence of the discredited cause.

Experiment 2: Causal Alternative

Experiment 2 used modified versions of the warehouse fire scenario used by Wilkes and Leatherbarrow (1988). For the mention-control group, the implied theme of the account was that the fire was an accidental result of the presence of the volatile materials. In the "negation" group, subjects learned about the volatile materials and then had this information discredited, leaving an implied theme involving stationery products stored in the warehouse (as in the previous experiment). In the "alternative" group, the implied theme, after the discrediting, was that the fire occurred due to arson. Finally, in the arson control group, subjects received information suggesting arson, but did not hear mention of the volatile materials. If subjects are able to discredit information when the correction presents an alternative instead of just a negation, one would expect the alternative group to show fewer references to the volatile materials than the negation group would, and potentially as many arson inferences as in the arson control group. Subjects could create and report a coherent account of the event by relying on the information about the arson materials, and thus would not use the information about the volatile materials in their inferences.

Method. Eighty-one University of Michigan undergraduates participated in the same procedure used in the previous experiment. The materials were essentially the same as those used in the previous experiments, except for the critical messages that defined the different conditions (see Table 2).

For the three groups listed above, Message 6 stated that volatile materials had been carelessly stored in a warehouse closet. In the arson control group, Message 6 stated that the warehouse closet was empty. In all conditions, Message 13 either stated that the previous message regarding the closet's contents was incorrect and the closet was empty, or that two firefighters had been released from the hospital. Message 14 gave either a list of supplies normally stocked by the warehouse, or indications that gasoline-soaked rags (arson materials) had been found in suspicious circumstances. The memory tests used did not differ from those in the previous experiment, except that the memory questionnaire contained several additional questions, on topics like motives of the owner and the cause of a previous fire on the premises, to assess inferences consistent with the different themes used here.

Table 2. Message manipulations by group.

Group	Message 13	Message 14
Mention control	firefighters	warehouse supplies
Negation	empty closet	warehouse supplies
Alternative	empty closet	arson materials

Results and Discussion. The responses to inference questions were scored as consistent with either a "volatile materials" theme or an "arson" theme. The "volatile materials" theme was coded as in the previous experiment; the arson theme encompassed responses that referred to key words from the message presenting the arson materials (e.g., gasoline, rags, drums or barrels), arson intent, or motives for arson. The groups were significantly different on both the negligence and the arson themes; $F(3, 76) = 15.02, p < .0001$ for volatile materials; $F(3, 76) = 27.98, p < .0001$ for arson.

Both the correction groups showed high levels of recall of the correction, with 95.2% of the negation group and 75.0% of the alternative group recalling it. The groups did not differ in summary recall or in fact recall ($F < 1.7$ and $F < 2.61$, respectively). The results show that when considering the overall pattern of inferences, subjects showed less influence from the discredited information when they received a correction that provided an alternate causal explanation, as compared to those subjects receiving only a negating correction. Subjects in the negation group made as many volatile-materials inferences as a control group for whom the information on volatile materials was not discredited. However, subjects in the alternative group showed fewer volatile-materials inferences, and made as many arson inferences as those in the arson control group, where the reports never mentioned the volatile materials. The results suggest that a correction accompanied by a causal alternative works better than one entailing only negation, at least in cases where the alternative does a good job of filling the causal role the discredited information would have filled.

However, an analysis of the number of uncontroverted references to the volatile materials themselves, in either memory test showed a significant difference between the arson control group and the three other groups that had heard about the volatile materials ($t(76) = 6.80, p < .0001$ in a post-hoc analysis). Thus, a causal alternative does not necessarily prevent all influence from the discredited information, because subjects made several direct references to the volatile materials if they had seen those materials mentioned, regardless of the type, or existence, of a correction.

Table 3. Number of inferences consistent with a volatile materials theme, arson theme, and direct references to volatile materials, by group.

Theme	Group			
	Mention	Negation	Alternative	Arson
Volatile materials	4.32	4.14	2.60	2.10
Arson materials	1.47	1.33	4.00	4.33
References to volatile materials	3.47	3.71	3.25	0.29

General Discussion

The results found here provide evidence that one must consider the causal structure information affords, rather than its mere presence, in explaining the continued influence of discredited information. In Experiment 1, when the reports presented the volatile materials in a way that was causally sufficient to account for the fire's starting, subjects showed many inferences involving them, whether a correction occurred or not. However, when the volatile materials were simply mentioned in the story context, subjects recalled the volatile materials information, but made very few inferences referring to those materials. Thus, it appears that information will influence inferences despite discrediting based on the causal structure it affords, and not just because mentioning the information makes retrieval and use more likely.

Further, in Experiment 2, subjects made fewer inferences involving the discredited information when the correction also presented a causal alternative to it. This suggests that one may involve discredited information in inferences when it is the only material that affords causal structure to the account, and without it, one would be unable to construct as good a representation of the event. This is consistent with other proposals that subjects use discredited information when they do not have other "relevant" information available (Anderson, et al, 1985; Wegner, Coulton, & Wenzlaff, 1985), but argues that the mere presence of plausible information does not as readily lead to influence, thereby specifying "relevance" in terms of the causal role in which the discredited information was presented.

The results here may reflect idiosyncrasies of the particular text used; however, there is much evidence for similar effects using other types of materials in widely varying informational contexts. Others have

found that subjects use later information about a scenario to reconstruct earlier memories of it (Hertel, 1982; Spiro, 1980). Subjects have been found to involve discredited information in making judgments about task ability in others and oneself (Ross, et al., 1975), social theories (Anderson, et al., 1980), person impressions (Schul & Burnstein, 1985; Wyer & Budesheim, 1987; Wyer & Unverzagt, 1985), and acceptability of a work contract (Golding, Fowler, Long, & Latta, 1990).

These results have several implications for understanding accounts containing corrections. Asserting information can result in its propagation through later inferences despite even direct, immediate corrections. This in turn suggests that one cannot completely discredit information by merely addressing the literal content expressed earlier, so an effective correction must go further if it is to result in a complete discrediting. Further, the more likely a piece of information is to become involved in causal inferences, the more likely it is to lead to continued influence upon discrediting. Thus, a central and important piece of information in an account would be harder to eradicate from memory if incorrect, but would also be the information most important to correct accurately. Lastly, because having a causal alternative associated with the correction of previous information decreases, although does not eliminate, influence from discredited information, corrections that suggest alternatives may improve the likelihood of accurate correction of information in memory.

References

- Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, 39(1-6), 1037-1049.
- Carretta, T. R., & Moreland, R. L. (1983). The direct and indirect effects of inadmissible evidence. *Journal of Applied Social Psychology*, 13(4), 291-309.
- Golding, J. M., Fowler, S. B., Long, D. L., & Latta, H. (1990). Instructions to disregard potentially useful information: The effects of pragmatics on evaluative judgments and recall. *Journal of Memory and Language*, 29(2), 212-227.
- Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Abex.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hertel, P. T. (1982). Remembering reactions and facts: The influence of subsequent information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(6), 513-529.
- Johnson, H. M., & Seifert, C. M. (1992). Memory and discredited information: Can you forget I ever said that? *Proceedings of the Fourteenth Annual Cognitive Science Society Conference*, Bloomington, IN.
- Johnson, H. M., & Seifert, C. M. (1993). The continued influence of discredited information: Can you forget you ever heard this? Unpublished manuscript. University of Michigan, Ann Arbor, MI.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 162-182.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32(5), 880-892.
- Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality & Social Psychology*, 49(4), 894-903.
- Spiro, R. J. (1980). Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior*, 19, 84-95.
- Trabasso, T., & van den Broek, P. W. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612-630.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232.
- van den Broek, P. (1990). Causal inferences and the comprehension of narrative texts. In Graesser, A. C., & Bower, G. (Eds.), *The Psychology of Learning and Motivation (Vol. 25): Inferences and Text Comprehension*. San Diego: Academic Press, Inc.
- Wegner, D. M., Coulton, G. F., & Wenzlaff, R. (1985). The transparency of denial: Briefing in the debriefing paradigm. *Journal of Personality & Social Psychology*, 49(2), 338-345.
- Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *Quarterly Journal of Experimental Psychology*, 40A, 361-387.
- Wyer, R. S., & Budesheim, T. L. (1987). Person memory and judgments: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology*, 53(1), 14-29.

Wyer, R. S., & Unverzagt, W. H. (1985). Effects of instructions to disregard information on its subsequent recall and use in making judgments. *Journal of Personality & Social Psychology*, 48(3), 533-549.

Appendix A

Materials adapted from Wilkes and Leatherbarrow (1988).

Message 1: Jan. 25th 8:58 p.m. Alarm call received from premises of a wholesale stationery warehouse. Premises consist of offices, display room, and storage hall.

Message 2: A serious fire was reported in the storage hall, already out of control and requiring instant response. Fire engine dispatched at 9:00 p.m.

Message 3: The alarm was raised by the night security guard, who had smelled smoke and gone to investigate.

Message 4: Jan. 26th 4:00 a.m. Attending fire captain suggests that the fire was started by a short circuit in the wiring of a closet off the main storage hall. Police now investigating.

Message 5 (mention and discrediting groups): 4:30 a.m. Message received from Police Investigator Lucas saying that they have reports that cans of oil paint and pressurized gas cylinders had been present in the closet before the fire.

Message 5 (control): 4:30 a.m. Message received from Police Investigator Lucas saying that they have reports that the closet was empty before the fire.

Message 6: Firefighters attending the scene report thick, oily smoke and sheets of flames hampering their efforts, and an intense heat that made the fire particularly difficult to bring under control.

Message 7: It has been learned that a number of explosions occurred during the blaze, which endangered firefighters in the vicinity. No fatalities were reported.

Message 8: Two firefighters are reported to have been taken to the hospital as a result of breathing toxic fumes that built up in the area in which they were working.

Message 9: A small fire had been discovered on the same premises, six months previously. It had been successfully tackled by the workers themselves.

Message 10: 10:00 a.m. The owner of the affected premises estimates that total damage will amount to hundreds of thousands of dollars, although the premises were insured.

Message 11 (control): 10:40 a.m. A second message received from Police Investigator Lucas regarding the investigation into the fire. It stated that the two firefighters taken to the hospital had been released.

Message 11 (discrediting groups): 10:40 a.m. A second message received from Police Investigator Lucas regarding the investigation into the fire. It stated that the closet reportedly containing cans of paint and gas cylinders had actually been empty before the fire.

Message 12: The shipping supervisor has disclosed that the storage hall contained bales of paper; mailing and legal-size envelopes; scissors, pencils, and other school supplies; and a large number of photo-copying machines.

Message 13: 11:30 a.m. Attending fire captain reports that the fire is now out and that the storage hall has been completely gutted.

New message added in Experiment 1

Message 3 (noncausal reference group): The alarm was raised by the manager of a store located across the street, which sells cans of oil paint and pressurized gas cylinders. He had smelled smoke and become concerned.

New messages added in Experiment 2

Message 5: The fire marshal had recorded several fire code violations on the premises at a surprise inspection two months earlier.

Message 7: The display room was reported to contain display cases, catalogs, and the sales staffs' desks. It was only staffed from 11 a.m. to 2 p.m., due to diminishing sales.

Message 14: (alternative and arson control groups) 11:08 a.m. Firefighters have found evidence of gasoline-soaked rags near where the bales of paper had been stored in the storage hall, as well as several emptied steel drums of suspicious nature. The owner denies any knowledge of these materials.