Title
The induction of hidden causes: Causal mediation and violations of independent causal influence

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN
1069-7977

Authors
Carroll, Christopher
Cheng, Patricia

Publication Date
2010

Peer reviewed

The Induction of Hidden Causes: Causal Mediation and Violations of Independent Causal Influence

Christopher D. Carroll (cdcarroll@ucla.edu)
Department of Psychology, UCLA
Los Angeles, CA 90095 USA

Patricia W. Cheng (cheng@lifesci.ucla.edu)
Department of Psychology, UCLA
Los Angeles, CA 90095 USA

Abstract

In order to explain the apparent violation of a causal assumption, people often posit hidden causes. The assumption of independent causal influence states that the power of a cause to produce or prevent an effect is independent of other causes. Some preventers violate independent causal influence; we conducted an experiment to test whether people posit a hidden mediating cause to explain these preventers. The results indicated that participants are more likely to posit a hidden mediator when the preventer violates independent causal influence.

Keywords: causal reasoning; causal inference; prevention; hidden causes; unobserved causes

Introduction

Although people often reason about simple cause and effect, they typically assume that such causal relationships are embedded in complex causal structures with hidden causes. So while people know that aspirin prevents headaches, they also believe that this relationship is mediated by some complex biological mechanism involving hidden causes. In many circumstances, the hidden causes are inconsequential. Knowing how aspirin prevents headaches is less important than knowing that it does so. Indeed, people often reason appropriately with only shallow causal knowledge (e.g., Keil, 2003). However, hidden causes may be important in other circumstances.

In particular, hidden causes may be important when the observed causes violate causal assumptions such as the Markov assumption in causal Bayesian network models (Pearl, 2000; Spirtes, Glymour, & Scheines, 1993/2000) or the assumptions for inferring causal power (Cheng, 1997; Novick & Cheng, 2004). Inferences about hidden causes have been demonstrated in a number of studies where some causal assumption is violated. Children appeal to hidden causes in order to explain probabilistic causation, and this may reflect an assumption that causation is deterministic (Schulz & Sommerville, 2006; see also Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008). Similarly, both adults (Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Hagmayer & Waldmann, 2007; Luhmann & Ahn, 2007) and infants (Saxe, Tenenbaum, & Carey, 2005) posit hidden causes when there is an unexplained effect, presumably reflecting the assumption that every effect has a cause (Kant, 1781/1965). Finally, people infer a hidden contextual cause when the causal power of the observed cause interacts with its context (Liljeholm & Cheng, 2007; Rottman & Ahn, 2009).

In this paper, we focus on the assumption of *independent causal influence* (Cheng, 1997; Novick & Cheng, 2004). Independent causal influence requires that the power of one cause to produce or prevent the effect is constant: it does not change with context or with the occurrence or non-occurrence of other causes. According to independent causal influence, if aspirin prevents headaches caused by colds, then it will also prevent headaches caused by dehydration, stress, and so on. We investigate a specific violation of independent causal influence that arises in prevention.

Preventive scope is the range of circumstances across which a preventer works (Carroll & Cheng, 2009). A *broad preventer* stops the effect no matter what the cause, but a *narrow preventer* only stops the effect when the effect is produced by a certain *targeted cause*. Aspirin and nasal spray illustrate the difference between broad and narrow prevention. As a broad preventer, aspirin prevents headaches of all kinds (e.g., headaches caused by colds and headaches caused by stress). As a narrow preventer, sinus spray only prevents headaches caused by colds; it would not prevent a headache caused by stress.

Narrow prevention violates the assumption of independent causal influence because the power of the preventer depends on which cause is producing the effect *e*: a narrow preventer prevents *e* when it is brought about by the targeted cause *c*, but it does not prevent *e* otherwise. However, it is possible to reconcile narrow prevention and the assumption of independent causal influence by positing a certain type of hidden cause: a hidden *mediator*. Suppose that *c* produces *e* indirectly through a mediator and that the narrow preventer prevents the mediator rather than preventing *e* directly (see Figure 1). Once the mediator is included in the explanation, none of the causal relationships violate independent causal influence: *c* and the preventer independently influence the mediator, and the mediator and other causes independently influence *e*. As long as other causes of *e* produce *e* via mechanisms other than the mediator, the preventer will only stop *e* when it is being produced by *c*. Thus, narrow prevention would only appear to violate the assumption of independent causal influence because there is an unobserved mediator.

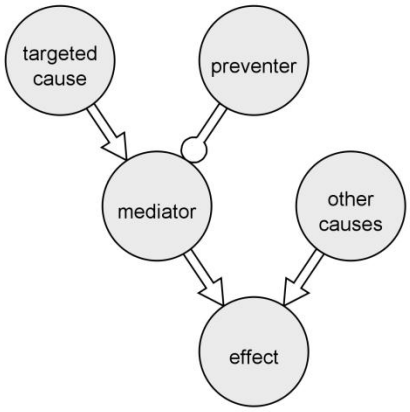

Figure 1: Mediation as an explanation of narrow prevention. Generative and preventive causation are denoted by arrows and modified arrows terminating in a circle, respectively.

Previous research (Carroll & Cheng, 2009) demonstrated that people distinguish between narrow and broad prevention, but the relationship between preventive scope and the inference of hidden mediation is less clear. Although participants in the previous research identified the mediation shown in Figure 1 as an explanation for narrow prevention, they only did so in a two-alternative forced-choice procedure. This shows that participants preferred the explanation in Figure 1 to the available alternative (an explanation where the preventer directly stopped the effect), but it is impossible to tell whether participants would have endorsed this explanation outside of this particular forced-choice question. Participants may have endorsed causal mediation as the better – but potentially unappealing – explanation of the two available choices. Furthermore, the experiment previewed the choices before showing the participants any data, and this may have biased participants towards interpreting the data with one of the provided explanations. Whether causal mediation is a favored explanation for narrow prevention more generally remains to be seen. Moreover, previous research did not clarify the relationship between preventive scope and the assumption of independent causal influence.

To assess whether the causal mediation explanation of narrow prevention is appealing more generally, we tested whether people endorse causal mediation after encountering a narrow preventer.

Method

Participants were asked to imagine themselves as researchers at a medical research company. They were directed to investigate how two fruit products from the rain forest - pane fruit and asmine juice - influence whether someone will have a headache. In all conditions, participants were shown some clinical trials where pane fruit caused headaches and asmine juice prevented headaches. We manipulated whether asmine juice was a narrow or broad preventer. After viewing the data, participants reported whether they expected asmine juice to prevent headaches under various circumstances. Finally, the participants were given a series of statements and were asked to endorse or reject each statement. One of these statements presented the mediation explanation.

Participants

Forty undergraduates at the University of California, Los Angeles (UCLA) participated to obtain course credit in a psychology course. Participants were assigned to the narrow (n = 20) or broad (n = 20) prevention condition.

Materials

The data presented in the narrow and broad prevention conditions are shown in Table 1. The critical difference between the conditions can be seen by comparing the effect of asmine juice on headaches attributed to the background cause. In the broad prevention condition, drinking asmine juice reduced the number of headaches even when pane fruit was not consumed. This can be seen by comparing the number of headaches when people neither ate pane fruit nor drank asmine juice to the number of headaches when people drank asmine juice but did not eat pane fruit (see the top half of Table 1). In the narrow prevention condition, it did not do so.

Table 1: The frequency of headaches (the effect) as a function of pane fruit (cause), asmine juice (preventer), and prevention condition. F = pane fruit, J = asmine juice.

Observed Causes	Broad prevention	Narrow prevention
none	10 out of 50	10 out of 50
J	5 out of 50	10 out of 50
F	40 out of 50	40 out of 50
J, F	20 out of 50	20 out of 50

As shown in Figure 2, the data were presented in displays containing cartoon faces. Each cartoon face represented a person in the clinical trial, and the type of cartoon face (happy face or sad face) indicated whether the person had a headache.

Procedure

Participants were randomly assigned to the broad or narrow prevention conditions and then given the following cover story:

Imagine that you work for a drug company that develops headache medications. You have heard rumors about a certain area in a rainforest where many of the fruits influence whether someone has a headache (either by causing a headache or preventing it).

The drug company has asked you to investigate these claims.

You decided to run clinical trials to assess the influence of pane fruit and asmine juice. You recruited

volunteers and randomly divided them into groups. Each group was assigned a specific treatment (e.g., eating pane fruit but not drinking asmine juice).

The results of each trial are summarized by tables of cartoon faces, and you can tell whether someone had a headache by looking at the cartoon face.

Participants were shown data in a display similar to Figure 2, and were given a print-out of the data to reference while answering subsequent questions. The instructions emphasized that the results had been replicated in much larger studies so that any differences in the frequency of the effect were reliable.

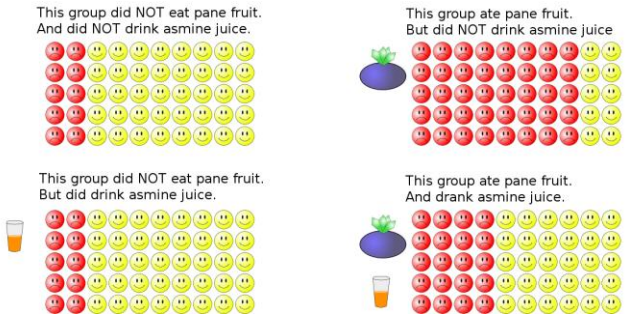

Figure 2: The data shown in the narrow prevention condition. Shaded frowning faces indicate people with headaches, and the lighter smiling faces indicate people without headaches.

Then, participants answered a series of counterfactual questions designed to measure beliefs about how pane fruit and asmine juice influence headaches. Each question asked the participants to imagine a group of people with certain characteristics and to predict whether consuming one of the food products would lead to more, fewer, or the same number of headaches in that group. For example, the *fruit* counterfactual - designed to assess the influence of pane fruit among a group of people who have not drunk asmine juice – asked following question:

Imagine that you go to a small town in the United States. If you brought PANE FRUIT to the town and everyone ate it, do you think that MORE people would have headaches, FEWER people would have headaches, or the SAME NUMBER of people would have headaches?

The other questions assessed when participants believed that asmine juice would prevent headaches. The *juice|fruit* question asked participants to predict the effect of asmine juice among people who live in a town near the rainforest and frequently consume pane fruit. The *juice|no fruit* question measured the influence of asmine juice among people living in a small town in America (who presumably have not eaten pane fruit). The *juice|withdrawal* question measured the influence of asmine juice among a group of people who have headaches for a specific reason other than eating pane fruit: they have stopped drinking coffee and are experiencing caffeine withdrawal.

Finally, participants were shown a series of statements about pane fruit and asmine juice. The statements were shown one at a time, and participants were asked to endorse whichever statements they agreed with. Table 2 lists these statements in the order that they were presented. Endorsement of the *mediation* statement provided the critical measure of whether participants inferred a hidden mediator. It should be noted that the mediation statement is compatible with broad prevention as well as narrow prevention: a broad preventer might destroy the substance in addition to directly preventing the effect when it is produced by other mechanisms.

Table 2: Participants were asked to indicate whether they agreed or disagreed with the following statements

Type	Statement
prevents	Asmine juice can sometimes prevent or relieve headaches.
develop drug	Your company may be able to turn asmine juice into a drug like aspirin, selling it widely as a headache treatment.
mediation	Pane fruit produces a RARE substance that causes headaches, and asmine juice destroys THAT substance.
combination	There is something special about the combination of asmine juice and pane fruit that prevents headaches.

Results

For the counterfactual questions, participants indicated whether there would be more, fewer, or the same number of headaches after consuming one of the food products. To analyze these responses, we coded responses of "more" as 1, "fewer" as -1, and "same number" as 0.

As expected, most participants predicted that pane fruit causes headaches. For the pane fruit counterfactual, the mean response was .90 (SD = 0.45) in the broad prevention condition and .95 (SD = 0.22) in the narrow prevention condition. The difference between these experimental conditions was not significant, $t(38) = 0.45$, $p = .66$.

On the other hand, the predicted influence of asmine juice depended on the experimental condition and the specific counterfactual (see Figure 3). Participants in both conditions believed that asmine juice would prevent headaches among groups of people that had eaten pane fruit (juice|fruit counterfactual). However, there were noticable differences between the conditions for the other counterfactuals. When participants were shown broad prevention, they believed that asmine juice would prevent headaches in every counterfactual. In contrast, when participants were shown narrow prevention, they were much less likely to believe

that asmine juice would prevent headaches when the headaches were produced by either an unknown (juice|no fruit) or a known non-targeted (juice|withdrawal) cause.[1]

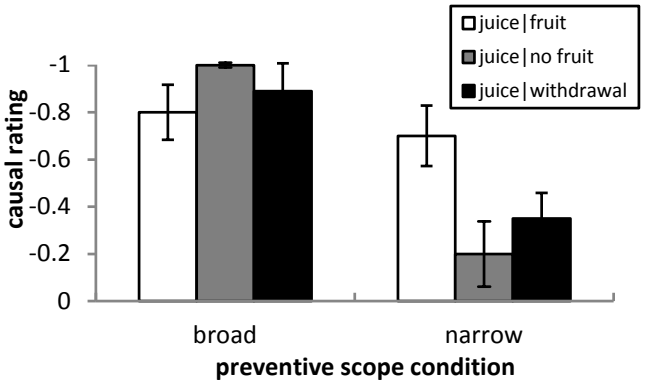

Figure 3: Prevention ratings for the counterfactual questions involving asmine juice. Error bars show the standard errors.

To confirm these patterns, we conducted an ANOVA with prevention condition (broad or narrow prevention) and prevention counterfactual (juice|fruit, juice|no fruit, or juice|withdrawal) as the independent variables. The ANOVA found a main effect of prevention condition, $F(1,38) = 17.08$, $p < .001$, and an interaction between prevention condition and prevention counterfactual, $F(2, 76) = 6.28$, $p < .01$. To investigate the source of the interaction, we conducted a separate ANOVA in each prevention condition. These ANOVAs confirmed that there was a non-significant effect of counterfactual question under broad prevention, $F(2, 38) = 1.85$, $p = .17$, and a significant effect of counterfactual question under narrow prevention, $F(2, 38) = 6.34$, $p < .01$.

The percentages of participants endorsing the statements are shown in Table 3. Participants in both conditions were very likely to report that asmine juice sometimes prevents headaches, but participants in the broad prevention condition were more likely to do so, $p < .05$ by Fisher's exact test. This difference might reflect the failure of some participants in the narrow prevention condition to notice that the preventer prevents the effect. Participants in the broad prevention condition were much more likely to believe that asmine juice could be developed into a headache drug and widely marketed, $\chi^2(1, N = 40) = 20.42$, $p < .001$. Participants in the narrow prevention condition were much more likely to believe that pane fruit and asmine juice might produce and prevent headaches via a rare shared mediator, $\chi^2(1, N = 40) = 4.91$, $p < .05$. Neither of the experimental conditions led many participants to suggest that the combination of pane fruit and asmine juice prevented headaches, and the difference between the conditions was not statistically significant, $p = .41$ by Fisher's exact test.

Table 3: Percentages of participants in each condition who agreed with the statements.

Question	Broad prevention	Narrow prevention
prevents	100%	75%
develop drug	95%	25%
mediation	35%	70%
combination	10%	25%

Discussion

This experiment demonstrates that people will endorse causal mediation in order to explain narrow prevention. The results also confirm that people distinguish between narrow and broad prevention, using preventive scope to guide generalization. Broad prevention was generalized irrespective of context, but narrow prevention was only generalized when the effect was produced by the targeted cause. A narrow preventer was not expected to stop the effect when the effect was produced by an unknown cause or a cause other than the targeted cause.

These findings contribute to a growing body of evidence showing that causal assumptions play a central role in the induction of hidden causes. Models of causal inference that make minimalistic assumptions (e.g., Pearl, 2000; Spirtes, Glymour, & Scheines, 1993/2000) may fail to explain these findings (see Griffiths & Tenenbaum, 2009 for other situations where minimalistic assumptions prove inadequate). In fact, since broad and narrow prevention imply the same conditional independencies among the observable variables,[2] Pearl's (2000) causal Bayesian network model represents them with the same causal graphs, treating a causal graph with mediation and a causal graph without mediation as equivalent.

Why is independent causal influence so important that its preservation warrants positing a hidden cause? The power PC theory (Cheng, 1997; Novick & Cheng, 2004) uses the assumption of independent causal influence as a defeasible default assumption to justify the inference of causal power. Without this assumption, the causal power of a candidate cause with respect to an effect is indeterminate even if the usual prerequisites for causal inference (e.g., "no

[1] Average causal ratings that are close to zero might be produced by (1) a roughly even mixture of "fewer" and "more" responses, or (2) many "same number" responses. Few participants reported that asmine juice causes headaches ($n = 2$); the answers close to zero were driven primarily by "same number" responses.

[2] Although the preventer is independent of the effect conditional on the cause being absent in narrow prevention but not broad prevention, causal Bayesian network models do not recognize this distinction. Causal Bayesian network models consider the conditional independencies of variables, not the conditional independencies at certain levels of variables. Since the preventer and effect are dependent for *some* values of the cause in both broad and narrow prevention, conditionalizing on the cause does not render them independent in the sense that causal Bayes nets consider when constructing a causal graph.

confounding") are satisfied. A difference in the probability of the effect in the presence of the cause and in its absence, for example, could be entirely due to the interaction between the cause and the context. If so, then there would be no reason to expect the cause to produce the effect in a different context. Indeed, without independent causal influence, causal power is bound to specific contexts, and causes will combine in unpredictable ways from one context to the next. This would render generalization unjustified. The assumption of independent causal influence jumpstarts the inference of causal power and supports generalization to transfer contexts via a context-independent causal power.[3]

Although people view causal mediation as a viable explanation for narrow prevention, the reason for this inference is less clear. There are at least two possibilities. First, people may posit causal mediation liberally, but only endorse causal mediation of a certain form. If so, participants in the broad prevention condition might be equally comfortable with causal mediation except that they prefer explanations where the mediator is common rather than rare. If this is the case, then people use the assumption of independent causal influence to infer the *form* of causal mediation. That is, the violation or non-violation of independent causal influence would determine whether people expect the mediator to be shared between different causes of the effect.

Alternately, people may posit mediation only when causal assumptions are violated. Since causal relationships can be decomposed almost indefinitely, this represents a reasonable strategy to minimize the complexity of causal explanations while maintaining useful assumptions. Broad prevention, which does not violate independent causal influence, can be explained and predicted without causal mediation. Therefore, positing causal mediation provides little practical benefit. For narrow prevention, however, the representation of mediation provides more tangible benefits: it allows people to generalize more accurately. If people can identify the mediator, they can infer whether the preventer will stop other causes from producing the effect. Thus, the violation of the assumption of independent causal influence serves as a criterion for revising one's causal explanation to achieve more accurate predictions.

In summary, narrow and broad prevention differ in whether they respect the assumption of independent causal influence. In narrow prevention, which violates independent causal influence, people view causal mediation as a plausible explanation. By positing causal mediation, people preserve the assumption of independent causal influence.

[3] The assumption of independent causal influence can be replaced, without changing the predictions regarding generalization, by the assumption that the causal factors in the background that interact with the targeted cause occur with the same probability across contexts (Cheng, 2000). Since our dependent measures do not allow differentiation between these assumptions, we treat them as equivalent for our purposes.

Acknowledgments

The preparation of this article was supported by AFOSR FA 9550-08-1-0489. The authors wish to thank Hannah Har for her assistance with data collection.

References

Carroll, C. D., & Cheng, P. W. (2009). Preventive scope in causation. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 833-838). Austin, TX: Cognitive Science Society.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104,* 367-405.

Cheng, P.W. (2000). Causality in the mind: Estimating contextual and conjunctive causal power. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 227-253). Cambridge, MA: MIT Press.

Gopnik, A, Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111 (1),* 1-31.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction, *Cognitive Psychology, 51,* 334-384.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116 (4),* 661-716

Hagmayer, Y., & Waldmann, M. R. (2007). Inferences about unobserved causes in human contingency learning. *Quarterly Journal of Experimental Psychology, 60 (3),* 330-355.

Kant, I. (1965). *Critique of pure reason.* London: Macmillian. (Original work published 1781).

Keil, F. C. (2003). Folkscience: coarse interpretations of a complex reality. *Trends in Cognitive Sciences, 7 (8),* 368-373.

Liljeholm, M., & Cheng, P. W. (2007). When is a cause the "same"? *Psychological Science, 18 (11),* 1014-1021.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115 (4)* , 955-984.

Luhmann, C. & Ahn, W. (2007). BUCKLE: A model of unobserved cause learning. *Psychological Review, 92 (3),* 657-677.

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review, 111,* 455-485.

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* Cambridge, England: Cambridge University Press.

Rottman, B. M. & Ahn, W. (2009). Causal inference when observed and unobserved causes interact. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 1477-1482). Austin, TX: Cognitive Science Society.

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: inferences about hidden causes by 10- and 12-

month-old infants. *Psychological Science, 16 (12),* 995-1001.
Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and children's inferences about unobserved causes. *Child Development, 77 (8),* 427-442.
Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction and search* (2nd ed.). Cambridge, MA: MIT Press. (Original work published 1993).