Causal Reasoning, Psychology of

Introductory article

Barbara A Spellman, University of Virginia, Charlottesville, Virginia, USA
David R Mandel, University of Victoria, Victoria, British Columbia, Canada

CONTENTS

Causal reasoning is an important universal human capacity that is useful in explanation, learning, prediction, and control. Causal judgments may rely on the integration of covariation information, pre-existing knowledge about plausible causal mechanisms, and counterfactual reasoning.

INTRODUCTION

The subject of causal reasoning has engaged psychologists of many kinds (e.g. cognitive, social, animal, clinical, developmental), other cognitive scientists (e.g. philosophers, computer scientists, anthropologists), and others outside the cognitive science community (e.g. lawyers). The question that psychologists want to answer is: how do we go from the information that the world provides us in the form of events occurring (seemingly at random sometimes) to our beliefs about what causes what? Sometimes causal judgments are made in formal settings: in the laboratory, scientists try to find out what causes cancer or heart disease; in the legal system, before liability or punishment is imposed, jurors are required to determine who caused the accident or who caused someone's death. But more informally, we all reason about causality daily. Why did I fail this exam (or all of the exams in this course)? Why is my friend unhappy? Why is my computer likely to crash in the next five minutes? To answer these kinds of questions we may rely on repeated observations, pre-existing knowledge, thought experiments, or all of these cues to causality.

THE ROLE OF REPEATED OBSERVATIONS

One kind of information that we use to assess causality is information from repetitions of the same events: watching causes and effects as they repeatedly occur. Sometimes people use the word 'cause' deterministically, so that '*A* causes *B*' means that every time *A* occurs *B* must follow. But people also use the word 'cause' probabilistically, so that '*A* causes *B*' means that *A* increases the chances that *B* will occur. For example, someone might say that a baseball team has a winning record early in the season 'because' they have played most of their games at home. The argument is that playing at home increases the chances of winning; however, it's still possible for the team to lose some games at home and win some away games. Similarly, scientists claim that smoking causes lung cancer because it increases the chance of getting lung cancer. Yet there are people who smoke and don't get lung cancer, and people who get lung cancer without smoking.

In order to determine whether something might be causal when we have repeated observations, we divide our observations into the following categories (see Figure 1): cause present and effect present (cell *A*); cause present and effect absent (cell *B*); cause absent and effect present (cell *C*); cause absent and effect absent (cell *D*). We then use that information to decide whether the cause increases the probability of the effect. We do this by comparing the proportion of times the effect occurs when the cause is present, $A/(A+B)$, with the proportion of times the effect occurs when the cause is absent, $C/(C+D)$. The difference between these proportions is called the contingency (symbolized by Δp – read 'delta *p*'), which is a measure of the strength or effectiveness of a cause. If the proportion of times the effect occurs is greater when the cause is present than when the cause is absent then the difference will be positive, and we speak of a 'generative' or 'facilitative' cause – it makes the effect more likely to happen. If the proportion of times the

		Effect (lung cancer) Present	Effect (lung cancer) Absent	
Cause (smoking)	Present	$A = 10$	$B = 90$	$\frac{A}{A+B} = \frac{10}{100}$
	Absent	$C = 10$	$D = 990$	$\frac{C}{C+D} = \frac{10}{1000}$

$$\Delta p = \frac{A}{A+B} - \frac{C}{C+D} = 9\%$$

Figure 1. We need four kinds of information to understand the relation between smoking and lung cancer. Note that we do not directly compare the number of people who smoke and get lung cancer with the number of people who don't smoke and get lung cancer. Rather, we look at whether smoking increases the chances of getting lung cancer. For people who smoke, the proportion who get it is 10/100 or 10%, whereas for people who don't smoke, the proportion who get it is 10/1000 or 1%.

effect occurs is smaller when the cause is present than when the cause is absent then the difference will be negative, and we speak of a 'preventive' or 'inhibitory' cause – it makes the effect less likely to happen. (Contingencies can range from -1 to $+1$.) Usually people are concerned with generative causes.

Using all the Information

The above computation of Δp suggests that all of the information should be equally important in evaluating the effectiveness of a cause. However, studies have shown that people tend to overweight the information in cell *A* – the 'present–present' cell. This overweighting may be one reason why people believe in superstitions or horoscopes. For example, there are people who believe that if they walk under a ladder they will have bad luck. And, in fact, once or twice when they did walk under a ladder they did have bad luck. However, they may fail to remember or consider the times they walked under a ladder and didn't have bad luck (cell *B*), the (possibly many) times they had bad luck without walking under a ladder (cell *C*), and the very many times they didn't walk under a ladder and didn't have bad luck (cell *D*). All of that information is needed before one can correctly evaluate whether walking under a ladder causes bad luck.

Considering the Base Rate of the Effect

Although the Δp computation captures the idea that a cause is something that changes the probability of an effect, the number that results from the computation may be deceptive when one is trying to evaluate the effectiveness of a particular cause. Besides looking at the contingency, people also consider how often an effect occurs in general (called the 'base rate' of the effect).

For example, suppose you have 100 plants in your garden but only 80 of them have flowers. You buy a special plant food, and soon all 100 have flowers. How effective is this product? It may only work 20 percent of the time (and have happened to work on the flowers that hadn't already bloomed); or it may work 100 percent of the time (but you couldn't tell because some flowers had already bloomed anyway).

It appears that people are sensitive to this problem when judging the effectiveness of a cause. For example, compare two plant foods. One was given to 100 plants where there were initially no flowers, and then 20 bloomed ($20/100 - 0/100 = 0.20$). The other, as above, was given to 100 plants where there were initially 80 flowers, and then all 100 bloomed ($100/100 - 80/100 = 0.20$). Both plant foods have a contingency of 0.20, yet the first worked only on 20 percent of the plants that didn't already have a flower whereas the second worked on 100 percent of the plants that didn't already have a flower.

People judge the second plant food as more effective than the first even though the contingencies are equal. Thus, it seems that people take into account the base rate of the effect and adjust their estimates according to how much influence a cause had above and beyond the influence of other causes.

THE ROLE OF PRE-EXISTING KNOWLEDGE

We have seen how we can use statistical covariation information to assess the relation between a potential cause and effect. However, we don't only use statistical information when making causal judgments; our pre-existing knowledge of the world will influence which statistical information we will pick up and use, and how we will limit and interpret the statistical relations we discover.

The most important limitation to note is that even though finding a contingency means that there is a correlation between the two events, that does not mean that the first event causes the second event. As scientists often say: 'correlation does not prove causation'. For example, every morning, just before dawn, the rooster crows. Then the sun rises. Yet we do not believe that the rooster crowing causes the sun to rise. In many cities, when ice cream sales go up, the murder rate goes up; when ice cream sales go down, the murder rate goes down. Yet we do not believe that eating ice cream causes people to commit murder. Why not? Because we have other knowledge.

Causal Mechanisms

Why don't we believe that the rooster causes the sun to rise even though there is a perfect correlation? One idea is that we don't believe it because we can't conceive of a causal mechanism. How could the noise of a tiny animal affect a powerful celestial object? In fact, belief (or non-belief) in a mechanism can direct, or misdirect, searches for potential causes. For example, in the mid-nineteenth century a physician named Ignaz Semmelweis had a very difficult time convincing physicians to wash their hands after examining a cadaver before turning to deliver a baby (to lessen mortality of the mothers) – a practice that seems obvious and obligatory now – because no one then could imagine a causal mechanism.

So where do our beliefs about causal mechanisms come from? One possibility is that they come from our knowledge of similarities, categories, and other statistical relations. We don't believe that the rooster causes the sun to rise because we know that lions roaring don't cause rain and dogs barking don't cause full moons. There are no statistical relations between one 'kind' of event (animal noise) and the other 'kind' of event (weather), so we never developed the idea that there could be a causal mechanism. On the other hand, these days we are willing to accept that many new ailments can be caused by unseen microorganisms which can be transmitted in many ways (e.g. breathing, touching) because we have noted other similar relations in the past.

Labeling a Cause as a Cause

Given the same combination of events, which gets labeled as a 'cause' of the outcome may differ between situations, individuals, and cultures. For example, suppose a fire breaks out in a nearby warehouse and you are explaining the cause of the fire to a friend. You are likely to mention that there was an arsonist or a stroke of lightning. You are unlikely to mention the presence of combustible material or oxygen, even though both of those are necessary for the fire. Under 'normal' circumstances we just assume that they are present, and so their presence or absence does not covary with the outcome and we do not consider them causes. Now imagine a special furniture factory in which an area is kept free of oxygen so that high-temperature welding can take place. One day there is an oxygen leak, and when the usual welding begins a fire ensues. Under these special circumstances, you would call the oxygen a cause of the fire. Thus, what we point to as being causal is not just something that increases the probability of an effect, but rather something that increases it relative to some background assumption of what is stable or normal.

Which of the many relevant factors a person chooses to pick out from the background and label as a cause may also depend on that individual's beliefs. For example, suppose a young man robs a shop. What caused this behavior? Some people would argue that it was because he was brought up in a bad neighborhood, citing the fact that children brought up in his neighborhood are more likely to go on to commit such a crime than children brought up in better neighborhoods. Other people would argue that it was because he was a 'bad apple', citing the fact that there are many other children brought up in his neighborhood who do not commit such crimes. In such a case, what you label as a 'cause' may influence what you believe is the best treatment for the problem.

Different cultures also tend to pick out different factors as causal. For example, in individualist cultures (such as the United States and Australia) people are more likely to attribute causality for an action to the actor's personality or 'disposition', whereas in collectivist cultures (such as China and India) people are more likely to attribute causality to the situation or circumstances. (*See* **Cultural Differences in Causal Attribution**)

Alternative Causes

Although it seems easy enough to figure out the statistical relationship between one cause and one effect once you know where to look, the world is complicated and it is not always possible to examine one potential causal relation at a time. Sometimes potential causes are independent, so you can evaluate each separately. But often potential causes covary with each other, making it difficult to distinguish the causal contribution of each.

Controlling for alternative causes

When two or more potential causes of an effect act at once, and not independently, we have to control for one cause while evaluating the other. As a simple example, suppose you rush into your favorite coffee shop and assert loudly that drinking coffee must cause lung cancer because people who drink lots of coffee get lung cancer more often than those who do not (a positive Δp). Probably none of the coffee drinkers there would be alarmed: they would point out to you that perhaps people who drink more coffee also smoke more, so although it may look as if coffee drinking causes lung cancer, it is really smoking that is doing the causal work. Here people see an alternative causal mechanism to explain the lung cancer: smoking. To evaluate whether coffee drinking causes lung cancer while controlling for smoking you need to (1) consider all people who don't smoke and ask whether coffee drinking increases their probability of getting lung cancer, and (2) consider all people who do smoke and ask whether coffee drinking increases their probability of getting lung cancer. If the answer is negative in both cases, then coffee drinking is not a cause of lung cancer: it is only because it covaries with smoking that it seems to raise the probability of lung cancer.

When evaluating whether something is a cause of an effect, it is important to control for alternative causes. Obviously, it is never possible to know for certain that one has considered all potential alternative causes, but controlling for known alternative causes is a technique intentionally used by psychologists and other scientists to improve scientific reasoning.

However, controlling for alternative causes is difficult without a theory of what those alternative causes might be. In the coffee example, people don't believe there is a way in which coffee drinking could cause lung cancer, so they seek an alternative causal mechanism. But in the case of ice cream sales and murder rates, it might be plausible to think that increased sugar consumption causes violence, and leave it at that. A mechanism is not necessarily correct just because it is plausible. When temperature is controlled for, there is no correlation between ice cream sales and murders; they only seem related because hot weather leads to more of both. Many experiments have shown that people do control for known alternative causes when judging causal effectiveness. If experimenters tell people about an alternative cause, or if they have a reason to believe that some alternative factor (e.g. smoking) might be causal, then they will think of controlling for that cause. However, in experiments where the alternative factor (like temperature) is not so obvious, people are less good at controlling for it.

Thus, pre-existing knowledge of a causal mechanism may affect what information people acquire from the environment, what they control for, and, ultimately, what they will judge as the causes of other events.

Discounting

Although people do control for known alternative causes, the presence of an alternative cause may lead to a misjudgment of causal strength – known as 'discounting'. Discounting occurs when someone learns about two causes at the same time, and the judgment of the strength of the first cause is affected by the strength of the second cause. For example, suppose you have allergies and your doctor prescribes two medications *A* and *B*. Sometimes you take one, sometimes the other, and sometimes both. Medication *A* works ($\Delta p = 0.33$), but medication *B* doesn't work ($\Delta p = 0$). You tell the doctor that *A* is fairly effective in relieving your allergies. Now consider what would have happened had the doctor prescribed medications *A* and *C* instead. *A* still works ($\Delta p = 0.33$), but *C* works even better ($\Delta p = 0.67$). In your report to the doctor, you are likely to judge *A* as being less effective when the alternative cause is strong (*C*) than when the alternative cause is weak (*B*) – even though *A*'s effectiveness is the same.

Experiments have shown that discounting occurs even in simple cases when the two causes are independent. Discounting might result from a strategic decision or a belief that once we have found a good cause of an effect we need not invest in reliably assessing other potential causes.

JUDGING CAUSALITY IN SINGLE INSTANCES

We have considered how people make causal judgments when they have information about many instances of the cause and effect occurring. But how do we make causal judgments for events that occur only once (e.g. an accident, a crime, or a big promotion)? One theory is that we use counterfactual reasoning to make these judgments.

Counterfactual Reasoning

In many situations, people look back on a past episode and wonder what might have happened if some change had occurred leading up to its conclusion. For instance, suppose you decided to drive home from work by a scenic route one day because it was particularly beautiful outside, and along the way your car was hit by a reckless driver. Many, if not most, people in this situation would replay the episode in their mind in such a way that the accident is somehow 'undone'. For instance, you might imagine that if only you had taken your usual route home, or if only you had left a few seconds later, the accident would have been avoided. These imaginings of how the past might have been different involve counterfactual (or contrary-to-fact) thinking because the mentally replayed episode differs in some respects from the real episode.

Some philosophers and psychologists have proposed that counterfactual thinking plays an important role in causal reasoning. To explore the possible causes of a particular outcome, a person may mentally change one of the events preceding the outcome (an 'antecedent') and observe whether the outcome still occurs in the mental replay. If it is easy to imagine that the outcome would also be undone, then the antecedent is likely to be seen as one of its necessary causes. If, however, the outcome still seems inevitable, then the antecedent is unlikely to be seen as causally relevant.

Political scientists, historians, and legal scholars sometimes run counterfactual thought experiments to examine the causal implications of a particular change to a complex system. For example, some historians have considered what might have happened if Archduke Franz Ferdinand had not been assassinated in Sarejevo; and some have concluded that if that event had not happened, then the First World War would not have happened either. Although counterfactual thought experiments can be informative, it is usually impossible to verify whether the causal inferences drawn from them are justified – precisely because history cannot literally be replayed.

Similarities and Differences Between Causal and Counterfactual Reasoning

Counterfactual and causal reasoning sometimes focus on different events. For instance, in the car accident scenario, if people are asked directly about the cause of the accident they tend to identify the reckless driver, whereas if they are asked to generate counterfactuals that would undo the accident they tend to mention the route taken home. The counterfactual thoughts that come to mind may represent one's after-the-fact understanding of how the bad outcome could have easily been prevented. Thus, counterfactual thoughts often focus on behaviors that individuals can control. On the other hand, causal explanations often focus on antecedents that our knowledge of the world indicates would covary with the outcome over a set of similar cases. Thus, people say that the reckless driver was the cause of the accident because they realize that reckless driving is predictive of car accidents in general.

Even though the two forms of reasoning sometimes diverge, there is nevertheless a strong interplay between counterfactual and causal reasoning. If one cannot imagine that an antecedent might have been different, then it is unlikely that the antecedent will be identified as a cause. This is why, under normal conditions, oxygen makes a poor causal explanation for fire, even though it is a necessary condition. Similarly, if one cannot imagine an outcome having been different, then it is questionable whether it could be causally explained other than by recourse to concepts such as fate or destiny, which by definition, emphasize the immutability and inevitability of past episodes. (*See* **Counterfactual Thinking**)

DEVELOPMENT OF CAUSAL REASONING

Causal reasoning is necessary for human survival and, not surprisingly, the ability to perform such reasoning develops early. However, it is difficult to study causal reasoning in infants, because researchers cannot ask them direct questions about their judgments. Instead, researchers often use a technique called the 'habituation paradigm'. This technique takes advantage of the fact that when infants see the same events repeatedly (e.g. pictures on a video monitor, animations, objects moving in real life), they gradually get bored and will look at the events for shorter durations. If, however, they see a new or different event, they will 'dishabituate' and look for longer. Researchers can infer what infants count as 'the same thing' or 'a different thing' using this technique.

When using the habituation paradigm to study causal reasoning, researchers may show infants videotapes of collision events. In the 'causal launching' event, object *A* moves across the screen and hits stationary object *B*. When *A* strikes *B*, *A* stops, and *B* immediately begins to move with the same speed and direction as *A* had previously. This event looks natural to an adult, as if *A*'s collision with *B* caused *B* to move. Researchers can then modify these events. For example, a delay may be introduced, so that after *A* hits *B*, *B* remains in place for a second or two before it begins moving. Or, *B* may begin moving before being hit by *A*. With these modifications, adults will claim that it doesn't look as if *A* caused *B* to move. Do infants treat these modified events as the same as, or different from, the natural-looking causal launching event? Research has shown that by the age of about seven months, infants do perceive a difference between causal launching events and noncausal events. However, they perceive the difference only if the objects involved in the events are simple; for more complicated objects infants may have to be 10 months old to make the distinction.

If it takes infants longer to understand causality regarding more complicated objects, what about more complicated events? Often in life, we don't just say that an outcome was caused by the action that immediately preceded it; rather, we look back in time to see what caused that particular action. For example, when Mum sees milk spilled on the floor and yells at Little Sister to wipe it up because it fell from her tilted glass, if Little Sister says 'Big Brother pushed me' then he, rather than she, is seen as the cause and gets Mum's wrath. It is considered a sign of sophisticated reasoning in adults to be able to look back into the past for causes; this ability also develops over time in infants: at 10 months old they don't look back at earlier causes, whereas at 15 months old they do.

Thus, we see that causal reasoning starts to develop early, but not all at once. Rather, both the complexity of the objects involved in the events and the complexity of the relations between events affect causal understanding. (*See* **Causal Perception, Development of**)

CONCLUSION

Causal reasoning is a pervasive and important form of thinking that begins at an early age. People engage in causal reasoning in order to explain past outcomes, to achieve control over their natural and social environment in the present, and to forecast, plan and prepare for the future. However, reasoning about causality is complicated. It is complicated, in part, because causal judgments depend on multiple cues, such as covariation, spatial and temporal contiguity, and our beliefs about what is normal. It is also complicated because information about such cues may be obtained in a variety of ways, such as by observing new cause–effect sequences, recalling knowledge about the world, and mentally imagining counterfactuals. Moreover, the answer to the question 'What is the cause?' may depend on one's choice of causal background – which may be affected by motivation, knowledge, and culture. Despite the complexity of the concept, the power of such knowledge, when it is accurate, is formidable. Without the ability to understand causality and use causal knowledge, both our internal mental world and the external physical world in which we live would be radically different.

Further Reading

Cheng PW and Wu M (1999) Why causation need not follow from statistical association: boundary conditions for the evaluation of generative and preventive causal powers. *Psychological Science* **10**: 92–97.

Cohen LB, Rundell LJ, Spellman BA and Cashon CH (1999) Infants' perception of causal chains. *Psychological Science* **10**: 412–418.

Hart HL and Honoré AM (1985) *Causation in the Law*, 2nd edn. Oxford, UK: Oxford University Press. [First edition published in 1959.]

Hilton DJ (ed.) (1988) *Contemporary Science and Natural Explanation: Commonsense Conceptions of Causality*. New York, NY: New York University Press.

Mandel DR and Lehman DR (1996) Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology* **71**: 450–463.

Mandel DR and Lehman DR (1998) Integration of contingency information in judgments of cause, covariation, and probability. *Journal of Experimental Psychology: General* **127**: 269–285.

McGill AL (1989) Context effects in judgments of causation. *Journal of Personality and Social Psychology* **57**: 189–200.

Spellman BA (1997) Crediting causality. *Journal of Experimental Psychology: General* **126**: 323–348.

Spellman BA and Mandel DR (1999) When possibility informs reality: counterfactual thinking as a cue to causality. *Current Directions in Psychological Science* **8**: 120–123.

Spellman BA, Price CM and Logan J (2001) How two causes are different from one: the use of (un)conditional information in Simpson's paradox. *Memory and Cognition* **29**: 193–208.

Sperber D, Premack D and Premack AJ (eds) (1995) *Causal Cognition: A Multidisciplinary Debate*. New York, NY: Oxford University Press. [Symposia of the Fyssen Foundation.]

White P (1990) Ideas about causation in philosophy and psychology. *Psychological Bulletin* **108**: 3–18.