The Interplay between Covariation, Temporal, and Mechanism Information in Singular Causation Judgments

Simon Stephan and Michael R. Waldmann

Imagine you are taking a dose of a certain medication that causes stomach cramps as a side effect. Later that day you actually feel an unpleasant pain in your stomach. Did your taking the drug cause your stomach cramps, or did the two events merely co-occur coincidentally and your stomach cramps were actually caused by some alternative cause? Maybe instead of the drug it was your eating a salad for lunch that was spoiled?

The question being asked in the example is a singular causation question because it refers to a potential causal connection between events that actually occurred at a particular time in a particular place. Singular causation can be contrasted with general causation, which refers to causal relations obtaining between event types instead of event tokens. For example, the claim "Taking this type of medication causes stomach cramps" is a generic-level causal claim, whereas the claim that "My having taken this particular medication this morning was the cause of my stomach cramps at noon" is a singular-level causal claim. Singular causation queries are ubiquitous in our lives. Answers to singular causation queries often guide the goals we set and the actions we take to achieve them. For example, finding the cause of a disease in a specific patient guides treatment decisions for this patient. Finding the cause of environmental damage in a specific area may lead to specific interventions designed to reverse the situation.

Despite the prevalence of singular causation queries in our daily lives and their importance in professional disciplines such as medicine or law, answering the question of how we can know that two events were causally connected is far from trivial. The reason for this difficulty is, as has famously been noted by Hume (1748/1975), that causal connections between individual events are not directly perceivable (see also Cheng and Ichien, in this volume).

While the analysis of singular causation has kept philosophers busy for decades (see Beebee, Hitchcock, and Menzies, 2009) (see also Danks, 2017), most studies in psychology on causal cognition have focused on how people learn about general causal relationships (e.g., Cheng, 1997; Griffiths and Tenenbaum, 2005; Liljeholm and Cheng, 2007) (see Cheng and Buehner, 2012; Waldmann, 2017; Waldmann and Hagmayer, 2013, for overviews), how they use their general causation knowledge to make predictions (e.g., Rehder, 2014; Waldmann, 2000), or how they designate among a set of

potential causes a specific one as the cause of that effect (i.e., causal selection) (see, e.g., Cheng and Novick, 1991; Kahneman and Miller, 1986; Kominsky et al., 2015; Samland et al. 2016; Samland and Waldmann, 2016).

We have recently begun to investigate how reasoners use their general causation knowledge to answer singular causation queries (Stephan, Mayrhofer, and Waldmann, 2020; Stephan and Waldmann, 2018, 2022). Building on the power PC model of causal attribution proposed by Cheng and Novick (2005), we have proposed the *generalized* power PC model of singular causation judgments. The model computes the probability that a particular target cause *c* actually caused a target effect *e*. The model combines information about the general strengths of the potential causes of the target effect, which can be induced based on covariational information, with information about temporal relations between causes and effects (Stephan, Mayrhofer, and Waldmann, 2020).

Apart from covariation and temporal properties, an important cue helping in the assessment of singular causation is information about the mechanisms linking causes and effects (see, e.g., Cartwright, 2015, 2017; Danks, 2005). Cartwright (2017) lists the discovery of intermediate steps in a causal chain as one of the crucial indicators of singular causation. Also, psychological studies (e.g., Ahn et al., 1995; Johnson and Keil, 2018) have shown that reasoners are sensitive to mechanism information when making causal judgments (see Johnson and Ahn, 2017, for an overview): when asked to determine the singular cause of an observed effect, reasoners tend to search for information about the status of known causal mechanism variables (Johnson and Keil, 2018).

What has been missing in the literature is a formal computational model explaining why causal mechanism information is helpful for the assessment of singular causation. We have recently (Stephan and Waldmann, 2022) proposed an extension of our generalized power PC model of singular causation judgments that incorporates causal mechanism information and thus provides an answer to this question.

In this chapter, we provide an overview of our latest work. The focus will be on our model, but we will also briefly summarize key empirical findings. Since our model is a generalization of Cheng and Novick's (2005) power PC model of causal attribution, we will start with a brief review of their model (which we henceforth refer to as the standard model) and show what the standard model gets right and where it fails. To foreshadow, a limitation of the standard model is that it solely focuses on the strengths of the potential causes of an effect (and their base rates if the causes are unobserved) and thus neglects the possibility of *causal preemption*. We will then summarize our generalized power PC model of singular causation judgments and demonstrate how it overcomes shortcomings of the standard model. Then, we will show that our model can also be used to provide a formal answer to the question under what conditions causal mechanism information is a valuable cue in the assessment of singular causation. We will end with a discussion of open questions and ideas for future studies.

The role of causal strength information in singular causation judgments—the standard power PC model of causal attribution

The standard power PC model of causal attribution (Cheng and Novick, 2005) is an application of Cheng's (1997) causal power PC theory to situations in which reasoners

want to answer different types of causal attribution queries. For example, reasoners might observe an effect event (*e*) and want to know the probability that different candidate causes actually caused the observed effect. A different query might start with the observed co-occurrence of a potential cause and effect event (*c* and *e*) and ask about the probability that this co-occurrence is causal as opposed to coincidental. According to the standard model, to obtain this probability reasoners need to apply their knowledge about the general causal strength with which *C* generates *E*, as well as the strength of potential alternative causes *A* of the effect.

As has been shown by Cheng (1997), the strength (or power) of a target cause *C*, which is defined as the probability with which the target cause generates the effect independent of alternative causes (*A*), can be induced from observable patterns of covariation between a target cause and an effect given certain background conditions (see Liljeholm and Cheng, 2007; Novick and Cheng, 2004) by applying the following equation:

$$power_c = \frac{\Delta P}{1 - P(e \,|\, \neg c)}.$$

Under the causal Bayes net framework (Glymour, 2001; Gopnik et al., 2004; Pearl, 1988, 2000; Sloman, 2005) causal strength as defined by this equation corresponds to the parameter w_c of the causal arrow connecting *C* to *E* within a common-effect causal model in which the target cause *C* and an alternative cause *A* combine according to a *noisy-OR* gate (Glymour, 2003; Griffiths and Tenenbaum, 2005; Pearl, 1988). A graphical illustration of such a causal model is shown in Figure 5.1a. The additional parameters b_a and b_c represent the base rates of the alternative and the target cause, respectively. The parameter w_a denotes *A*'s causal strength (or power).

To illustrate how the standard power PC model of causal attribution applies knowledge about the causal strengths of the potential causes to attribute causality in a singular case, we will focus on a situation in which a reasoner has observed that *C*, *A*, and *E* have actually occurred in a specific situation (i.e., *C* = 1 or *c*, *A* = 1 or *a*, and *E* = 1 or *e*).
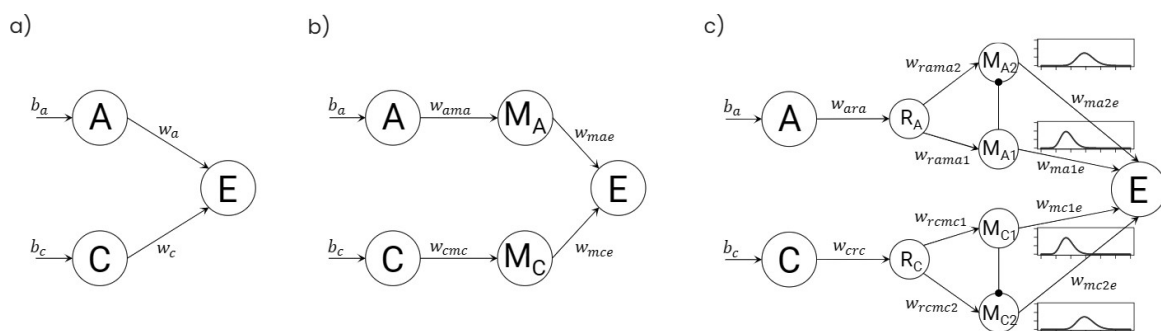

Figure 5.1 Different causal models in which C and A represent root causes of a common target effect E. Note: (a) a common-effect model with two causes, (b) an augmented version of that model with causal mechanism variables, (c) a model in which C and A generate E via different alternative mechanism paths. In this last model, the gray curves represent the causal latencies of the respective causal arrows. The causal link with the round head denotes an inhibitory link.
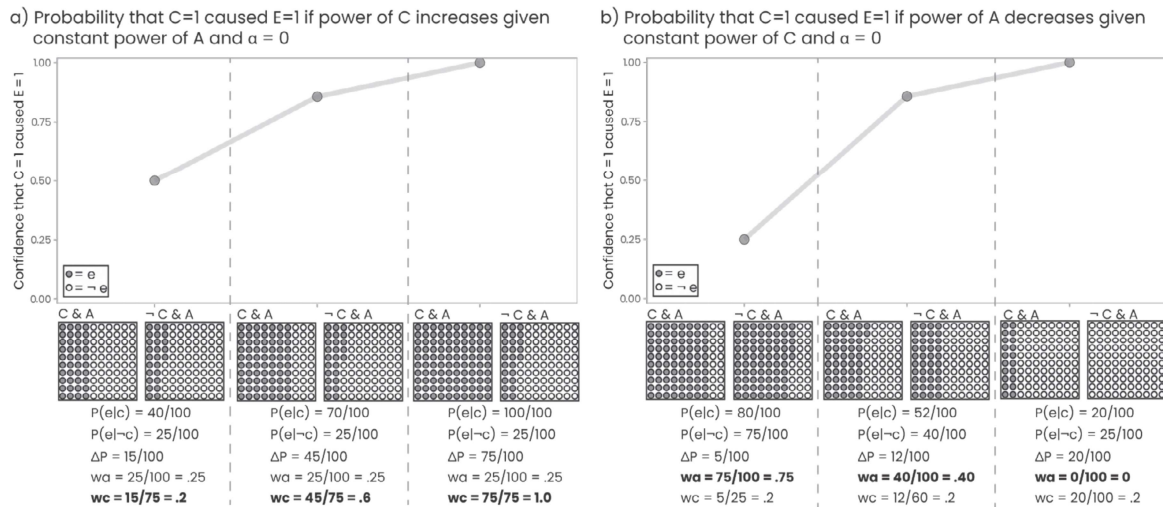
Figure 5.2 Predictions of the standard power PC model of causal attribution for different levels of w_c and w_a.

According to the standard model, the probability that *c* was the singular cause of *e* in this case is given by:

$$P(c \rightarrow e \,|\, c,a,e) = \frac{w_c}{w_c + w_a - w_c \cdot w_a} = \frac{w_c}{P(e\,|\,c,a)}.$$

(1)

The standard model thus predicts that in a situation in which *C*, *A*, and *E* are present, the probability that *c* caused *e* corresponds to *C*'s general causal strength divided by the conditional probability of the effect given the presence of the two potential causes *C* and *A*. Situations in which one or both of the potential causes are unobserved can also be modeled. In this case, the causes' strength parameters need to be multiplied with their corresponding base rate parameters (b_c and b_a).

Equation 1 captures a number of intuitions about singular causation, which are summarized in Figure 5.2. The different pairs of panels below the graphs' x-axes show different contingency tables. In each pair, the right panel displays cases in which only the alternative cause is present while the target cause is absent ($\neg c, a$), and the left panel contains cases in which both potential causes are present (c, a). Cases in which the effect is present (*e*) are depicted as gray circles. $P(c \rightarrow e|c,a,e)$, as defined in Equation 1, can be understood as the relative frequency of gray circles in the left panel that were produced by the target cause *C*. In Figure 5.2a, the causal strength w_c of the target cause *C* increases from the left to the right contingency data set (from $w_c = 0.2$ to $w_c = 0.6$ to $w_c = 1.0$), while the strength w_a of the alternative cause remains constant ($w_a = 0.25$). The graph shows that the probability that *c* caused *e* increases when w_c increases. The model thus captures the intuition that we should be more confident that the target cause actually caused the target effect in a specific situation when this cause is generally more likely to produce the effect.

The right graph shows another plausible way of reasoning about singular causation that is captured by the model. This time the target cause's strength is constant ($w_c = 0.2$) while

the strength of the alternative cause decreases (from $w_a = 0.75$ to $w_a = 0.40$ to $w_a = 0$). The graph shows that $P(c \rightarrow e|c,a,e)$ increases with decreasing strength of the alternative cause. The model captures the intuition that a candidate cause is more likely to be the singular cause of the effect if alternative causes are unlikely to produce the effect. In the case in which an alternative cause has a strength of zero, this event can be ruled out as the effect's singular cause. It therefore must have been the target cause that produced the effect, no matter how weak this cause generally is. This example shows that Equation 1 instantiates a reasoning principle that philosophers have called abductive reasoning, reasoning by elimination, or *Holmesian inference* (Bird, 2005, 2007, 2010). Later, when we address the role of mechanism information, we will show that mechanism information may allow reasoners to exclude alternative causes of a target effect even when these alternative causes have a general causal strength greater than zero.

Causal preemption—the generalized power PC model of singular causation judgments

In Stephan and Waldmann (2018) we have shown that a key problem of the standard model is that it solely focuses on the strengths of the potential causes of the target effect but neglects cases of potential preemption. Equation 1 predicts that the target cause is the singular cause of the target effect whenever the target cause is sufficiently strong to produce the effect. However, in the philosophical literature cases of redundant causation involving possible preemption have been discussed demonstrating that causal sufficiency of a target cause is itself not sufficient to conclude that a cause generated the effect in a singular case (see for example, Hitchcock, 2007; Paul and Hall, 2013). A cause that otherwise is sufficient to generate the effect might have been causally preempted in its efficacy on a singular occasion by an alternative cause that was simultaneously present and also sufficiently strong to generate the effect. A classic scenario in the philosophical literature illustrating the problem of causal preemption involves two rock throwers. Billy and Suzy are perfectly precise rock throwers, which means that neither of them ever fails to hit and destroy glass bottles when throwing stones at them. On a particular occasion, both protagonists end up throwing their rocks toward the same bottle. Both are throwing their rocks with identical force but Suzy manages to throw her rock a little bit earlier than Billy. Consequently, the bottle shatters. Intuitively, it was Suzy's and not Billy's throwing that was the singular cause of the bottle's shattering, even though we know that Billy's throwing was precise and strong enough to do the job. The problem of causal preemption can also occur in cases in which the potential causes are probabilistic, as we will show below.

To handle the problem of causal preemption, we (Stephan and Waldmann, 2018) have proposed the generalized power PC model of singular causation judgments, which is given by the following equation:

$$P(c \rightarrow e|c,a,e) = \frac{w_c - w_c \cdot w_a \cdot \alpha}{w_c + w_a - w_c \cdot w_a} = \frac{w_c \cdot (1 - w_a \cdot \alpha)}{P(e|c,a)}.$$
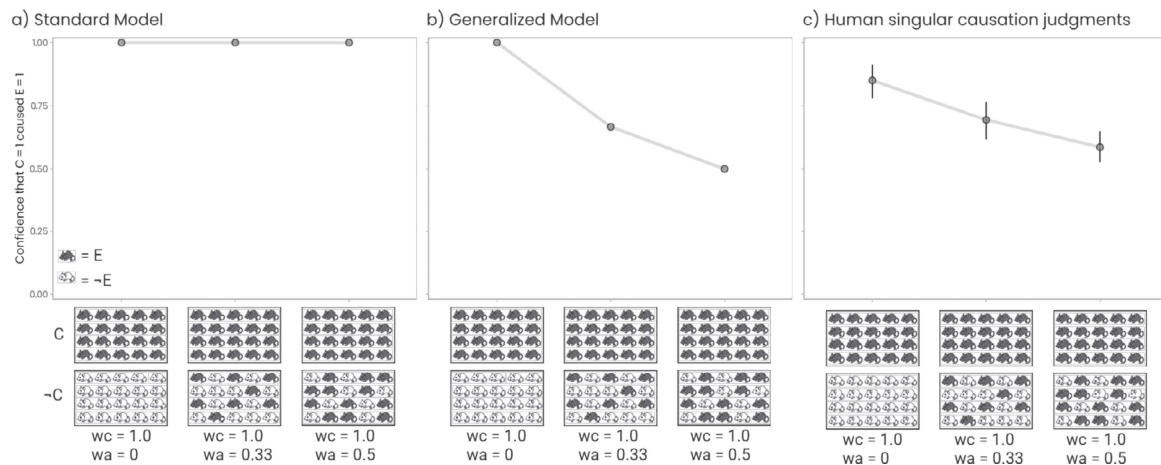
$$(2)$$

Figure 5.3 Model predictions and results of Experiment 1a from Stephan and Waldmann (2018). Results in c) show means and 95% CIs.

We now explain how the generalized model works by comparing Equations 2 and 1. As can be seen, the generalized model extends the numerator of Equation 1 by the product of the potential causes' causal strength and an additional weighting parameter α. The product $w_c \cdot w_a \cdot \alpha$ corresponds to the probability of causal preemption of the target cause c by the alternative cause a: the product of the causes' strength parameters, $w_c \cdot w_a$, is the probability that both causes are simultaneously strong enough to generate the effect. This product is relevant because the possibility of causal preemption only arises on occasions on which both potential causes have sufficient strength to produce the effect.

The product can be neglected in scenarios involving deterministic causes but it needs to be included in cases in which the potential causes have probabilistic strengths. On occasions on which causal preemption may occur (i.e., $w_c \cdot w_a$) it needs to be determined how likely it is that the target cause has indeed been preempted by the alternative cause. This probability is given by the α parameter of Equation 2. For example, an α value of 1.0 means that the target cause is always preempted by the alternative cause on occasions on which both a strong enough to produce the effect. An α value of 0.5 means that the target cause is preempted by its competitor on half of these occasions. Because $w_c \cdot w_a \cdot \alpha$ is the probability of causal preemption of the target cause by its competitor, it must be subtracted from the target cause's strength in the equation's numerator. Situations of "symmetric overdetermination" (see Paul and Hall, 2013) in which both potential causes are sufficiently strong to produce the effect and are acting absolutely synchronously, can also be handled. As no preemption occurs in situations of symmetric overdetermination α takes on a value of zero: the effect is caused by both c and a.

A relevant question is how the preemptive relation between the competing causes (i.e., the α parameter in our model) can be determined. In Stephan and Waldmann (2018) we suggested that this can be done based on *temporal* information about onsets and latencies of the competing causes.[1] We hypothesized that lay people intuitively incorporate assumptions about the causes' temporal relation in their singular causation judgments, even when temporal information is not explicitly mentioned in a situation. To test this

hypothesis, we conducted an experiment (Experiment 1 in Stephan and Waldmann, 2018) in which we employed a standard causal induction paradigm (see, for example, Buehner, Cheng, and Clifford, 2003; Griffiths and Tenenbaum, 2005) as it is typically used in studies on how reasoners infer general causal relationships based on contingency data. We presented our participants with a fictitious scenario in which they were asked to take the perspective of biologists conducting an experiment to learn whether a particular chemical substance causes the expression of a particular gene in mice. The fictitious experiment was presented as a classical randomized control trial (RCT). The results were shown in summary format, similar to the panels shown on the x-axes in Figure 5.3. We tested three contingency data sets, which are shown in Figure 5.3. In each data set, all mice of the treatment group (C) were expressing the gene (i.e., $w_c = 1.0$), while the relative frequency of the control mice ($\neg C$) expressing the gene because of alternative causes varied between 0, 0.33, and 0.5. The test question was a singular causation test query referring to a specific mouse from the treatment group. Subjects were asked to indicate on a slider how confident they were that it was the treatment with the chemical substance that caused this mouse to express the gene (from "very certain that it was not the chemical" to "very certain that it was the chemical").

Since the target cause has a causal strength of $w_c = 1.0$ (i.e., it is always sufficiently strong to cause the effect) in all conditions, the standard model predicts that subjects should be maximally confident that the chemical treatment was the singular cause of the target mouse's expressing the gene, $P(c \rightarrow e|c,e) = 1.0$ (see Figure 5.3a). By contrast, we predicted that subjects would incorporate temporal assumptions about when the alternative causes produced their effects, and that this would lead to lower singular causation ratings with increasing strength of the alternative causes (w_a). We assumed that in a scenario about factors triggering the expression of a gene, it would be plausible to assume that factors determining gene expression in mice (i.e., alternative causes) were already in place and operating before the biologists conducted their study and treated some mice with the chemical substance (the target cause). Under this assumption, participants should conclude that the alternative causes preempted the target cause in all cases in which the target and alternative causes were simultaneously sufficiently strong to produce the effect (given by $w_c \cdot w_a$). We modeled this assumption by setting the α parameter of the generalized model to 1.0. The predictions of our generalized model are shown in Figure 5.3b. We found that subjects' mean singular causation judgments closely traced the predictions of the generalized model (see Figure 5.3c). Different additional control questions further corroborated the hypothesis that subjects incorporated temporal assumptions in their singular causation judgments. For example, we found that subjects assumed that the alternative causes had already exerted their influence before the target cause was introduced. We could also rule out that the observed decrease in the $P(c \rightarrow e|c)$ ratings resulted from increasing uncertainty about the existence of a general causal relationship between C and E. This was a possibility because lower values of Δ_P (which measures the degree of contingency) imply less support for the existence of a general causal link between the factors (Griffiths and Tenenbaum, 2005; Meder, Mayrhofer, and Waldmann, 2014). To address this problem, we used sample sizes for our contingency data that were large enough so that the posterior probability of the existence of a general causal link between C and E,

as computed by the structure induction model (Meder, Mayrhofer, and Waldmann, 2014), was close to 1.0 even for the data set with the lowest contingency. Moreover, we asked subjects how strongly they believed that the chemical can generally cause the expression of the gene (which is a general causal structure query). Subjects reported high and almost identical confidence levels in all conditions.

The role of temporal relations

While our experiments in Stephan and Waldmann (2018) demonstrated that reasoners go beyond general causal strength knowledge and also incorporate temporal information when assessing singular causation relations, an open question was what exactly the relevant temporal factors are and how they need to be combined to determine the probability of causal preemption. In Stephan, Mayrhofer, and Waldmann (2020) we addressed this question by developing a theory of how the size of the model's α parameter might be determined.

Onset times

One relevant temporal factor we identified is the onset time difference between the competing causes. For example, in the philosophical preemption scenario about the two rock throwers our intuition that Billy's throwing was preempted by Suzy's is suggested by the information that Suzy's throwing occurred earlier than Billy's. Formally, if a target effect e could have been caused by either the target cause c or by one potential alternative cause a, the probability that c was preempted by a tends to be higher if a happened earlier than c. The onset time difference of two competing causes can be represented as $\Delta_t = t_a - t_c$, where t_a and t_c denote a's and c's onset times, respectively.

Causal latency

If two potential causes c and a are sufficiently strong to produce e but one occurred earlier, it is still not certain that the earlier cause actually preempted the other. A second temporal dimension that needs to be considered is *causal latency*, that is, the time it takes the causes to generate their effect. In the philosophical preemption scenario about Billy and Suzy, causal latency information is conveyed by saying that both protagonists are throwing their rocks with identical velocity. Thus, in this scenario identical causal latencies are assumed and the causes' preemptive relation depends only on their onset time difference. Generally, however, a potential cause occurring earlier than its competitor can still fail to preempt its competitor if the competing cause acts quicker than the target cause.

The causal latency of a cause C can formally be denoted as $t_{C \rightarrow E}$. Following previous research on general causal structure learning in dynamic contexts (e.g., Bramley et al., 2018) and queuing theory (Shortle et al., 2018), we used gamma distributions to model causal latency. The gamma distribution generalizes the exponential distribution and is defined by two parameters, a shape parameter $\kappa > 0$ and a scale parameter $\theta > 0$. A gamma distribution's expected value is given by $\kappa \cdot \theta$ and its variance by $\kappa \cdot \theta^2$. A cause's

latency can be induced based on multiple observations tracking the onset differences between cause and effect. For example, in a context shielded from the influence of alternative causes, the expected value of a causal latency distribution can be estimated by the average of the observed cause-effect delays.

Having identified and formalized these two temporal dimensions we showed in Stephan, Mayrhofer, and Waldmann (2020) that α, that is, the probability that the target cause c was preempted by the competing cause a if both happen to be sufficiently strong to produce e, corresponds to the probability that the value of the target cause's latency is greater than the sum of the causes' onset time difference and the alternative cause's latency: $\alpha = P(t_{a \to e} + \Delta_t < t_{c \to e} | e, c, a)$. We also showed that $P(t_{a \to e} + \Delta_t < t_{c \to e} | e, c, a)$ can be approximated by a simple Monte Carlo (MC) algorithm that randomly draws causal latency samples from each cause's latency distribution.

An illustration showing how causal latency information determines the probability that c was the singular cause of e according to our generalized power PC model of singular causation judgments, and how causal latency information is integrated with causal strength information, is given in Figure 5.4. The figure shows five common-effect causal structures that differ with respect to the causal latency distributions assigned to C and A. For example, in the first causal structure C's causal latency distribution has a much smaller expected value than A's, and the two distributions almost do not overlap. Neglecting onset time differences ($\Delta_t = 0$), the resulting value of α is 0.001 in this case.

In Figure 5.4 the value of α increases from the first to fifth's structure because the causal latency distribution of A shifts to the left while the causal latency of C shifts to
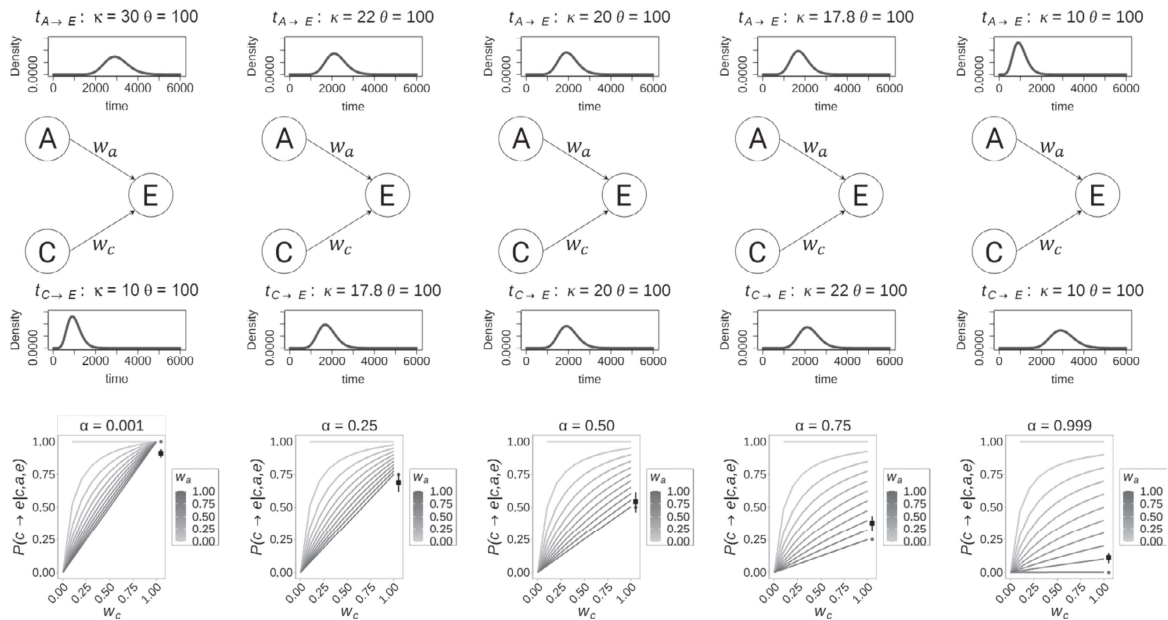

Figure 5.4 Illustration of how the causes' strengths and their causal latencies determine the predictions of the generalized power PC model of singular causation judgments. The black squares shown in the graphs of the last row show participants' mean singular causation ratings (error bars denote 95 percent CIs) as measured in Experiment 2 in Stephan, Mayrhofer, and Waldmann (2020). The gray points close to the means show the model predictions for the tested conditions.

the right. In the third causal structure, for example, both causes' latencies follow the same gamma distribution. In this case, the resulting α value is 0.5 because the chance that a randomly sampled single value from C's causal latency distribution is higher than a randomly sampled single value from A's causal latency distribution is fifty percent when the distributions fully overlap.

The five graphs in the last row of Figure 5.4 show how the value of α influences the predictions for $P(c \rightarrow e|c,a,e)$ made by the generalized power PC model of singular causation judgments for different combinations of causal strength values. In each graph, the causal strength of the target cause C is shown on the x-axis and different strengths of the alternative cause are represented by the different gray lines. Comparing the five graphs, it can be seen that increasing values of α tend to reduce the probability of a singular causal link between c and e. It can also be seen that the extent to which α diminishes $P(c \rightarrow e|c,a,e)$ tends to increase with increasing strengths of the alternative cause A. In the first graph, in which α is close to 0, the predictions of the generalized model correspond to those of the standard model.

In Stephan, Mayrhofer, and Waldmann (2020) we conducted several experiments testing the validity of the generalized model. The influence of causal latency information on singular causation judgments was tested in Experiment 2. A demo video of one of the study conditions can be found at https://osf.io/g6p72/. We presented subjects with a fictitious scenario about the emergency alarm system of a medieval kingdom. The scenario introduced soldiers located on two watchtowers, the Western and the Eastern tower. The task it was to look out for invading barbarians and to send out carrier pigeons to alarm the King's palace in case of an invasion. In two separate learning phases subjects first learned about the flight durations of the carrier pigeons from the two watchtowers by observing each tower multiple times. The flight durations followed a variant of the gamma distribution. The different pairs of latency distributions that were used are shown in Figure 5.4. They were manipulated between subjects. Since we were interested in the role of causal latency information on singular causation judgments in this experiment, we held the causal strengths of the two causes constant at $w_c = w_a = 1.0$. Thus, subjects observed that all carrier pigeons eventually arrived at the palace. In the test phase, subjects were asked to imagine a day on which both watchtowers simultaneously sent out a pigeon and on which an alarm later occurred in the palace. We then asked subjects how strongly they believed that the alarm was caused by the Western [or Eastern] tower. The graphs in Figure 5.4 show that when the potential causes, the pigeons, have causal strengths of $w_c = w_a = 1.0$, the generalized model predicts that subjects' singular causation judgments should corresponds to $P(c \rightarrow e|c,a,e) = 1 - \alpha$ (indicated by the gray points in each graph). The generalized model thus predicts a negative linear trend across the different test cases. The standard model, by contrast, always predicts that $P(c \rightarrow e|c,a,e) = 1.0$ when $w_c = w_a = 1.0$.

The mean singular causation ratings are shown as black squares (error bars denote 95 percent CIs) in Figure 5.4. As can be seen, the singular causation judgments followed the negative trend predicted by the generalized model, although ratings tended to be slightly less extreme than predicted by the model. In this experiment we also found that predictions that were based solely on the expected values of the latency distributions accounted less well for the ratings than predictions based on the full gamma

distributions. This finding indicates that the singular causation ratings were not only sensitive to the expected values of the latency distributions but also to their variance.

This experiment successfully demonstrated that reasoners incorporate and integrate information about causal latency in their singular causation judgments, as predicted by the generalized model. In a further experiment in Stephan, Mayrhofer, and Waldmann (2020) we found that subjects also correctly incorporate onset time difference (Δ_t). Singular causation judgments thus were confirmed to be sensitive to both temporal dimensions considered by the generalized model. In a final study, we then crossed different causal latency distributions with probabilistic causal strengths of the competing causes. Since a crucial prediction of our model is that causal strength and temporal information should interact, this study provided an important further test of the model. The singular causation ratings were again predicted well by the generalized model.

The role of mechanism information

A further cue helping in the assessment of singular causation relations that has been discussed both in philosophy and psychology is knowledge about mechanisms. Cartwright (2017, p. 8) has, for example, included the "presence of expectable intermediate steps (mediator variables)" in a list of factors indicating singular causation. Different psychological studies have shown that people make use of mechanism information when asked to determine the singular cause of an effect (e.g., Ahn et al., 1995; Johnson & Keil, 2018). For example, in a recent study Johnson and Keil (2018) presented their participants with either generic (e.g., "eating polar bear liver causes a person to become dizzy") or singular (e.g., "eating polar bear liver caused Bill to become dizzy") causation statements and asked them to select the type of information that they consider most relevant to determine if the presented statement was true. Subjects could choose between anecdotal, statistical/covariational, and mechanistic information. Johnson and Keil (2018) found that subjects preferred mechanistic information over statistical information when asked to evaluate singular causation claims, while the reverse was found when subjects evaluated general causation statements. It is true that general causal relationships often are discovered without any knowledge about details of the underlying causal mechanisms. For example, the causal link between smoking and lung cancer had been established based on statistical evidence before anyone knew *how* smoking leads to lung cancer. The native people of Guyana knew that dipping the heads of their arrows into *wouralia*, a substance extracted from certain plants that today is known as *curare*, creates a deadly weapon, but nobody knew how the poison worked until French scientist Claude Bernard conducted first experiments in 1856 (see Craver and Darden, 2013). However, even when detailed mechanism knowledge is not available, in most cases people use *abstract* mechanism intuitions to select candidate causes and effects. It is more likely that smoking is viewed as a possible cause of lung cancer than watching TV because the inhalation of smoke makes it a more plausible candidate cause. Moreover, general causal knowledge induced based on covariation information is typically not enough to establish singular causation relations, as a singular co-occurrence of events consistent with known regularities may still just be

coincidental. Information about causal mechanism variables seems to be relevant in this case. We have recently extended our generalized power PC model of singular causation judgments to provide a formal answer to the question of why mechanism information is helpful in this case (Stephan and Waldmann, 2022).

Under the causal Bayes net framework, causal mechanisms are modeled as intermediate nodes within causal chains or networks (see for example, Danks, 2005; Woodward, 2011) (see also Stephan et al., 2021). The original version of our generalized model was applied to the basic common-effect causal model shown in Figure 5.1a with two causes converging on an effect. In Stephan and Waldmann (2022) we applied the model to more complex causal structures that include intermediate mechanism nodes. One such structure containing mechanism nodes is shown in Figure 5.1b. This model can be thought of as an augmented version of the basic common-effect model in which the mechanisms underlying the direct causal arrows connecting C and A with E have been explicitly represented. In this augmented causal model, assuming that the causal Markov assumption holds (or assuming probabilistic modularity as proposed in Woodward, 2003), the original strength parameters w_c and w_a can now be re-represented as the products of single link strengths of the two paths: $w_c = w_{cmc} \cdot w_{mce}$ and $w_a = w_{ama} \cdot w_{mae}$ (see, for example, Waldmann et al., 2008). Similarly, the probability that c is the singular cause of e corresponds to the probability c caused m_C times the probability that m_C caused e: $P(c \rightarrow e | c, m_c, e) = P(c \rightarrow m_C | c, m_c, e) \cdot P(m_C \rightarrow e | c, m_C, e)$. The two equations allow us to see under which conditions the observation that a target cause's mechanism is active ($M_C = 1$ or m_C) increases a reasoner's confidence that the target cause c is the singular cause of the target effect e. Since the strength w_{mce} of the link connecting M_C to E will be higher than the strength w_c of the overall (unelaborated) causal path from C to E for $w_c < 1.0$, the probability that m_C caused e will on average be higher than the probability that c caused e in situations in which M_C is unobserved: $P(m_C \rightarrow e | m_C, e) > P(c \rightarrow e | c, e)$. Observing m_C will then lead to an increase in the probability that c is the singular cause of e if we can be certain that m_C was caused by c. One situation in which we can be certain that c caused m_C is when C is a necessary cause of its mechanism variable M_C. $P(c \rightarrow m_C | c, m_C)$ is 1.0 in this case and, as a result, $P(c \rightarrow e | c, m_C, e)$ reduces to $P(m_C \rightarrow e | c, m_C, e)$.

Learning about the absence instead of the presence of mechanism variables may also help in assessing whether c caused e. For example, if C can cause E only via M_C, and M_C is discovered to be absent in a singular case, c can be ruled out as the singular cause of e. Conditionalizing on situations in which $M_C = 0$, w_c takes on a value of 0 and, as result, $P(c \rightarrow e | c, \neg m_C, e)$ also equals 0. By contrast, if we discover in a situation that the alternative cause's mechanism variable M_A is absent, then we can be certain that c is the singular cause of e. In this case $w_a = 0$ and Equation 2 reduces to

$$P(c \rightarrow e \,|\, c, \neg m_A, e) = \frac{w_c}{w_c} = 1.0.$$

The previous analysis showed that causal mechanism information helps to assess singular causation relations because it allows us to insert more specific values for the causal strength parameters. In Stephan and Waldmann (2022) we showed that causal mechanism information also helps to specify the temporal relations between the

potential causes of a target effect. This is necessary when the causes can alternatively produce their effect via different possible mechanism paths. For example, in the shooting scenario one might consider the different ways via which being hit by a bullet can cause people to die. The bullet might either kill somebody by hurting their heart or their aorta. These different possible mechanism paths are not only associated with different causal strengths, but also with different causal latencies. Bullets hurting the heart not only kill more reliably than bullets destroying the aorta, they also kill quicker.

A causal structure in which the causes C and A generate their common effect E via different possible mechanism paths is shown in Figure 5.1c. The gamma distributions attached to the final causal links connecting the different mechanism nodes to the effect illustrate that the links are assumed to differ in their causal latencies. In particular, C and A each have one fast and one slow causal path. For C the fast causal path is the one via M_{C1} and for A the fast causal path is the one via M_{A1}. These different latency distributions assigned to the different mechanism paths imply that the α value will change depending on which mechanism components of the target and the alternative cause are observed to be active on a singular occasion. For example, if both potential causes activate the mechanism components that have the same causal latency (e.g., $M_{C1} = 1$ and $M_{A1} = 1$), then α would be 0.5. By contrast, in a situation in which C activates M_{C1} and A activates its slower mechanism variable M_{A2}, then α would take on a small value. In the reverse case, α would take on a high value.

An illustration of the scenario we used in the first experiment in Stephan and Waldmann (2022) is shown in Figure 5.5. Subjects learned about a kingdom in which one castle was located in the Southwest and another one in the Southeast, and in which the King's palace was located in the North.

The scenario was about the medical emergency system of the kingdom. Subjects learned that the only healer in the kingdom was living in the King's palace, and that the castles, to call the healer in case of a medical emergency, had to send out carrier pigeons.
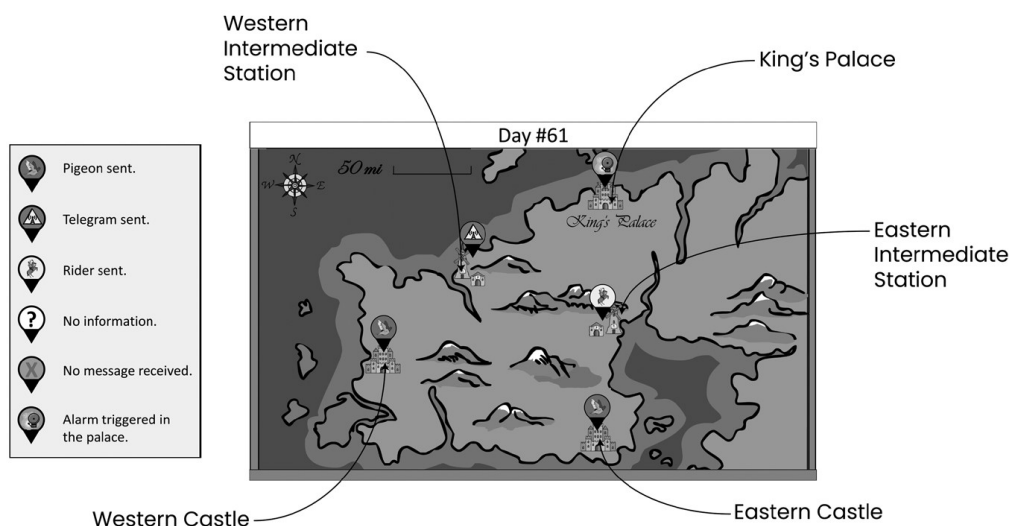

Figure 5.5 Illustration of the scenario used in Experiment 1 in Stephan and Waldmann (2022).
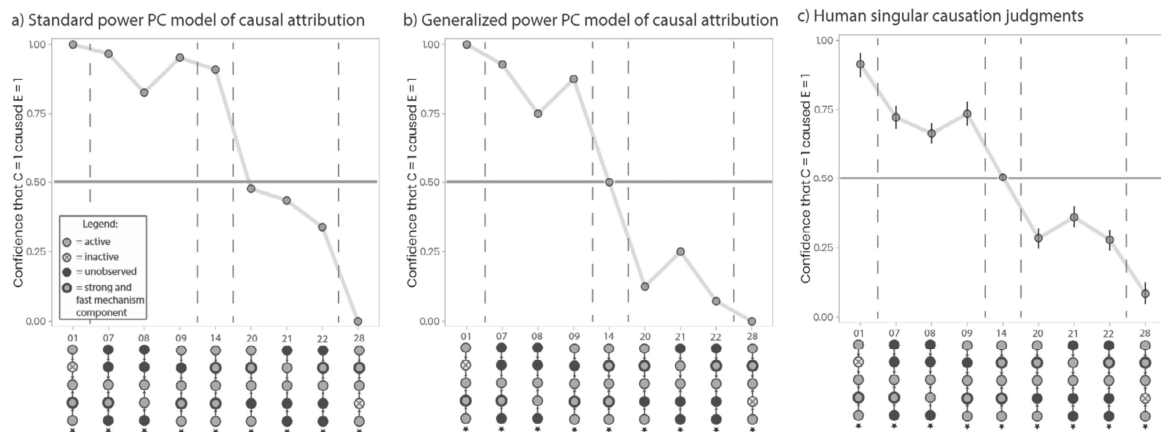
Figure 5.6 Model predictions and results (means and 95% CIs) of a subset of test cases used in Experiment 1 in Stephan and Waldmann (2022). Note. Asterisks in the neuron diagrams refer to the target cause.

Subjects learned that the pigeons could not fly the whole distance to the palace, which is why each castle had installed intermediate stations. To establish different possible mechanism paths, two possibilities were introduced of how incoming emergency signals could be forwarded to the palace at these intermediate stations. One possibility was telegraph towers and the other was pony riders stationed next to the telegraph towers (see Figure 5.5). The pony riders were described as back-ups ensuring that emergency signals could be forwarded if a telegraph tower was blacked out. The routes the pony riders had to take on their way to the palace were described as dangerous and it was pointed out that pony riders might get lost on their way to the palace. We hypothesized that subjects would assume that telegraph towers have a higher causal strength as well as a shorter causal latency than pony riders. Subjects were informed that they would observe what happened on different days and that the different events would be illustrated using different event icons (see Figure 5.5).

Subjects were informed that, for each day, they would be asked to indicate how strongly they believed that the observed alarm in the King's palace had been caused by the Western [Eastern] castle on that day. A demo video can be found at https://osf.io/ycv8u/.

Subjects were then asked to evaluate twenty-eight different singular observations that were compatible with the instructed causal structure. These test cases could be grouped into five different categories. For example, one category consists of all cases in which none of the alternative cause's mechanism components are active. A subset of nine test cases from the five different categories is shown on the x-axes in Figure 5.6 in the form of simplified neuron diagrams. The bottom nodes of these neuron diagrams marked by asterisks represent the target cause to which the singular causation test queries referred. The last four test cases (cases 20, 21, 22, and 28 in the original study) mirror the first four shown test cases. The middle case (case 14 in the original study) is a "symmetric" test case because the same event types are instantiated for the target and the alternative cause. The graphs in Figure 5.6 show the predictions made by the standard model (a) and the generalized model (b). For the first four test cases both models make similar qualitative predictions, but the predictions differ for the remaining

test cases. For the symmetrical test case, for example, the standard model predicts that subjects should be confident that the target cause was the singular cause of the effect. In this test case the target cause activates its strong and fast mechanism path. As we have seen, the standard model predicts high values for $P(c \rightarrow e|c,e)$ whenever the causal strength of the target cause is high. The generalized model, by contrast, predicts that reasoners should be uncertain in this case because the alternative cause also activates its strong and fast mechanism path, and thus has a fifty percent probability of successfully preempting the target cause. A further notable difference between the models is the contrast between the predictions for the first and the last four test cases. Since the last four test case mirror the first ones, the generalized model predicts that singular causation judgments for the last four cases should mirror those for the first four cases. For example, the difference between the test cases 09 and 20 is that target and alternative cause have switched. The generalized model predicts that $P(c \rightarrow e|c,e)$ for test case 20 corresponds to $1 - P(c \rightarrow e|c,e)$ for test case 09. The standard model does not make this prediction. As the standard model solely focusing on causal strength and neglects temporal information, it predicts an advantage of the target cause in all these cases.

The mean singular causation judgments for the nine test cases are shown in Figure 5.5c. We have found that their ratings (including those given for the remaining test cases) were predicted best by the generalized model. By contrast, the standard model as well as further alternative models that we included in our analyses failed to account for subjects' singular causation ratings. In sum, the results of this experiment demonstrate that people use and integrate mechanism information in an elaborate way when answering singular causation queries. The study shows that reasoners seem to understand that mechanism information helps with assessing singular causation relations because it allows us to insert more specific values for the strength and temporal parameters that are relevant for determining whether a target cause actually produced a target effect.

In Stephan and Waldmann (2022) we not only provided a formal account of why mechanism information is helpful for assessing singular causation relations, but also used our model to identify situations in which mechanism information is less useful. One factor diminishing the value of mechanism information is whether a mechanism path can also be activated by alternative causes. Observing the presence of a mechanism variable is more helpful for singular causation judgments when the target cause is the only cause that can activate this mechanism variable. A second factor diminishing the utility of mechanism information concerns the way in which the overall strength of the target cause (w_c) with respect to the target effect is distributed across the different components of its mechanism. We have used our model to show that, given constant levels of $w_c = w_{cmc} \cdot w_{mce}$ (see Figure 5.1b), the observation that the mechanism variable is active in a singular cause ($M_C = 1$) should lead to a stronger increase in the probability that c caused e when w_{mce} is the stronger component of the product $w_c = w_{cmc} \cdot w_{mce}$ than when it is the weaker component. We conducted two studies testing whether lay people understand the factors that reduce the utility of mechanism information. The studies showed that a subset of reasoners have a tentative understanding of the relevance of these factors. However, we also found that many people generally rely on the heuristic that mechanism information is always helpful.

Summary, open questions, and directions for future research

The generalized power PC model of singular causation judgments makes explicit how different types of information can be combined to determine how strongly a reasoner should believe that an observed effect event was actually caused by a potential target cause. Causal strength information is important for the assessment of singular causation relations because causes that reliably produce their effects are more likely to be successful than causes that only rarely generate an effect. Covariational information also allows us to learn about the prevalence and strength of alternative causes. For example, if an effect is known to occur only very rarely in the absence of the target cause, this means that alternative causes are sparse or weak. If alternative causes are sparse, then we should be more confident that an observed co-occurrence of a target cause and a potential effect was actually causal rather than coincidental, even if the target cause is generally relatively weak. A well-known example for this situation is the relation between smoking and lung cancer. Although smoking is a rather weak cause of lung cancer even in chain smokers, the probability that a singular instance of lung cancer in a heavy chain smoker was caused by that person's smoking is relatively high. The reason for this divergence is that lung cancer is extremely rare in people who are not smokers. The generalized power PC model of singular causation judgments provides a formal account of these facts.

Knowledge about the general strengths of causes is not sufficient to determine singular causation relations (see also Woodward, in this volume). Philosophical examples of redundant causation show that even deterministic general causes can fail to cause the effect in a specific situation if they are preempted by competing alternative causes. Philosophical process theories of causality attempt to solve the problem of causal preemption by referring to the transfer of energy or the exchange of conserved quantities between cause and effect (e.g., Dowe, 1992, 2007) (see also Wolff, 2007; Wolff, Barbey, and Hausknecht, 2010; Wolff and Thorstad, 2017). Counterfactual accounts of singular causation (e.g., Halpern and Hitchcock, 2015; Halpern and Pearl, 2005) aim to solve the problem of causal preemption by defining causality as causal dependence of the effect on the cause by counterfactually holding constant suitable causal factors (see also Hitchcock, 2009).

Our model implements a different solution to the problem of preemption. People often will rely on temporal information to determine the preemptive relation between potential causes of an effect. We have identified two types of temporal information that help with determining causal preemption: onset times and causal latencies. Our model makes explicit how these factors can be integrated and combined with causal strength knowledge to assess whether a cause was preempted by a potential alternative cause. Our experiments support the assumptions of the model.

In research on causal reasoning, causal mechanism and covariation-based theories have for a long time been regarded as competing alternative accounts of causal inference. But in fact, both frameworks are compatible with each other, as has been shown, for example, by Danks (2005). Both types of theories traditionally have been applied to explain different tasks of causal inference, though. Covariation theories have mostly been used to explain how people infer general causation relationships (e.g.,

Cheng, 1997; Gopnik et al., 2004; Griffiths and Tenenbaum, 2005; Novick and Cheng, 2004), while mechanism theories have focused on explaining singular causation judgments (e.g., Ahn et al., 1995). According to the causal Bayes net framework, mechanisms can be understood as a sequence of mediating variables within causal networks, and knowledge about them can be acquired based on covariational information. We have shown that mechanism knowledge can constrain inferences and thereby help with determining the singular cause of an effect. What had been missing in the literature is a formal explanation of why mechanism information can be useful to assess singular causation, and when its contribution is limited. In Stephan and Waldmann (2022) we used the generalized power PC model of singular causation judgments to provide a formal analysis along with empirical evidence demonstrating the value of the model.

Despite the success of the generalized model there are still important open empirical and theoretical questions. So far, we have only tested relatively simple scenarios. For example, in all experiments in Stephan, Mayrhofer, and Waldmann (2020) and Stephan and Waldmann (2022), we limited the causal model to two potential causes of a single effect. In everyday life reasoners will only very rarely be in a situation in which the number of potential causes is so small. It would therefore be interesting to assess in future studies how well participants perform in more complex situations. We suspect that subjects' performance will decrease rather quickly with increasing complexity.

Another interesting question is how singular causation judgments are generated in situations in which there are unknown background causes of an effect. Although we studied such a scenario in Stephan and Waldmann (2018), in all experiments in which we specifically tested the role of temporal relations (Stephan, Mayrhofer, and Waldmann, 2020), the status of all relevant causes as present or absent was explicitly mentioned. However, in the real world it is common that effects are not only determined by observed but also by unobserved causes that need to be considered in the judgments. In such situations it appears to be at first sight impossible to conclude confidently that the target cause actually produced the target effect because the possibility of preemption by an unobservable cause cannot be ruled out. However, people can nevertheless make rational inferences given suitable background assumptions. In situations with unobserved background causes, the individual parameter values for b_a and w_a cannot be estimated. However, as has been shown by Cheng (1997), the influence of unobserved background causes can be estimated by the probability of the effect in the absence of the target cause: $P(e|\neg c) = b_a \cdot w_a$ if the background causes are assumed to be generative, independent, and not interacting with the target cause. The more difficult problem is to judge potential preemptive relations between the target and the unobserved alternative causes. How can the value of our model's α parameter be specified in this kind of situation? One possibility is that reasoners in such cases simply rely on certain default values for α. For example, setting α to 0.5 would mean that we remain uncertain about the preemptive relation between the potential causes. As a conservative solution, α could be set to 1.0, which would reflect the assumption that the target cause is always preempted by an alternative cause whenever at least one of the alternative causes is simultaneously present and sufficient for the effect (which is estimated by $P(e|\neg c)$). In

this case, the resulting value for $P(c \rightarrow e|c,e)$ can be thought of as a measure of the lower boundary of a range within which the true probability of c having caused e lies.

Alternatively, reasoners may have in some cases more specific intuitions about preemptive relations (Stephan, Mayrhofer, and Waldmann, 2020). In contexts with unobserved background causes, temporal information can be extracted from the observable background rate of the effect. The background rate can be modeled using exponential functions (cf. Bramley et al., 2018). In situations in which the target cause and the unobserved background causes are all present, the distribution of the effect rate is a mixture in which the gamma distribution that represents the causal latency of the target cause is superimposed on the exponential distribution of the unobserved background causes.

When the unobserved background cause is absent, then only the gamma distribution of the target cause can be observed. Thus, if the rate in which the effect occurs in the presence of the target cause is much higher than in its absence, this is an indicator that the target cause has a very short causal latency. A short causal latency makes it less likely than a long latency that the target cause is preempted by a background cause (see Lagnado and Speekenbrink, 2010, for supporting evidence). One interesting future study thus would be to test if reasoners report higher degrees of confidence that c caused e if they have learned that the presence of the target cause leads to a steep increase in the effect's rate than if they have observed that the effect rate only mildly increases in the presence of the target cause.

Another limitation of the present studies is that we have only tested static test situations, that is, situations in which subjects receive information about the status of events in a specific case rather than having observed the unfolding of events. While we are frequently asked to make judgments about events in the past, we also often find ourselves in situations in which we witness events developing in front of our eyes. In such cases, we will directly experience the relevant temporal relations, which may either be consistent or inconsistent with our expectations. For example, if we have learned in the past that a cause typically generates the effect very quickly, but then observe a long delay between the events, it seems reasonable to be cautious to infer that the two events were causally linked.

Our studies so far focused on binary events in both the model and the scenarios we tested. Research on forensic epidemiology (see Freeman and Zeegers, 2016, for an overview) has often studied cases that involve continuous variables. In this literature various quantitative measures of culpability have been suggested. For example, one could ask whether a plaintiff's disease has actually resulted from a particular quantitative amount of toxic exposure. It would be interesting to use these models as an inspiration for an extension of the generalized model to also capture intuitions about singular causation involving continuous variables.

Finally, there are further cues to singular causation that have been proposed that still need to be integrated into our model. For example, one additional cue suggested by Cartwright (2015; 2017) is information about the presence of "required support factors," by which she means enabling conditions that are necessary for a cause to produce the effect. Another cue currently not considered by our model is knowledge about the absence of preventive or disabling factors. These concepts can be easily

integrated within our framework. Enabling factors can, for example, be represented as additional factors that interact with the target cause (Novick and Cheng, 2004). Disablers can either be explicitly represented as inhibitory causes or can be implicitly encoded as affecting causal strength (Stephan et al., 2021). Yet another useful cue is information about the occurrence of known side effects of a causal process because such side effects may provide diagnostic evidence that an unobserved cause – or one of the mechanism variables – have indeed been present on a singular occasion. For example, the high levels of the element iridium at the KT boundary are considered indirect evidence that the extinction of the dinosaurs was caused by the impact of a massive asteroid[2] (see Woodward, 1990). High levels of iridium are not a causally relevant part of the causal chain via which asteroid impacts lead to the eradication of earthly species. They are, however, a strong indicator that a massive impact had occurred 65 million years ago.

Despite the considerable progress in formalizing singular causation in the past years, more needs to be done to develop an account that also addresses more complex cases. It seems important to transition in future research from laboratory studies to analyses of real-world cases, for example in law or medicine, to obtain a better impression of the strengths and shortcomings of the generalized model and to develop it further.

Notes

1 Another relevant cue is information about *spatial* relations between the causal events. For example, in the preemption scenario about Suzy and Billy, we may learn that only Suzy's rock actually came into spatial contact with the bottle. In many cases, however, only temporal but not spatial relations between causes and effects can be observed (e.g., *smoking → cancer*).
2 We thank James Woodward for this illustrative example.

References

Ahn, W., C.W. Kalish, D.L. Medin, and S.A. Gelman (1995). "The role of covariation versus mechanism information in causal attribution." *Cognition*, *54*(3): 299–352.

Beebee, H., C. Hitchcock, and P. Menzies (2009). *The Oxford Handbook of Causation*. Oxford: Oxford University Press.

Bird, A. (2005). "Abductive knowledge and Holmesian inference." In T.S. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology* (1–31). Oxford: Oxford University Press.

Bird, A. (2007). "Inference to the only explanation." *Philosophy and Phenomenological Research*, *74*(2): 424–32.

Bird, A. (2010). "Eliminative abduction: examples from medicine." *Studies in History and Philosophy of Science Part A*, *41*(4): 345–52.

Bramley, N.R., T. Gerstenberg, R. Mayrhofer, and D.A. Lagnado (2018). "Time in causal structure learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12): 1880—910.

Buehner, M.J., P.W. Cheng, and D. Clifford (2003). "From covariation to causation: a test of the assumption of causal power." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6): 1119–140.

Cartwright, N. (2015). "Single case causes: What is evidence and why." In J. Reiss (ed.), *Philosophy of Science in Practice.* Dordrecht: Springer.

Cartwright, N. (2017). "How to learn about causes in the single case." *CHESS Working Paper no 2017–2004*. https://www.dur.ac.uk/resources/chess/ CHESSK4UWP_2017_04_Cartwright.pdf.

Cheng, P.W. (1997). "From covariation to causation: A causal power theory." *Psychological Review*, *104*(2): 367–405.

Cheng, P.W. and M.J. Buehner (2012). "Causal learning." In K.J. Holyoak (ed.), *The Oxford Handbook of Thinking and Reasoning* (210–33). New York: Oxford University Press.

Cheng, P.W. and L.R. Novick (1991). "Causes versus enabling conditions." *Cognition*, *40*(1–2): 83–120.

Cheng, P.W. and L.R. Novick (2005). "Constraints and nonconstraints in causal learning: Reply to White (2005) and to Luhmann and Ahn (2005)." *Psychological Review*, *112*(3): 694–706.

Craver, C.F. and L. Darden (2013). *In Search of Mechanisms: Discoveries Across the Life Sciences*. Chicago: The University of Chicago Press.

Danks, D. (2005). "The supposed competition between theories of human causal inference." *Philosophical Psychology*, *18*(2): 259–72.

Danks, D. (2017). Singular causation. In M.R. Waldmann (ed.), *The Oxford Handbook of Causal Reasoning* (201–15). New York: Oxford University Press.

Dowe, P. (1992). "Wesley salmon's process theory of causality and the conserved quantity theory." *Philosophy of Science*, *59*(2): 195–216.

Dowe, P. (2007). *Physical Causation*. Cambridge: Cambridge University Press.

Freeman, M. and M.P. Zeegers (2016). *Forensic Epidemiology: Principles and Practice*. New York: Academic Press.

Glymour, C. (2001). *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*. Cambridge, MA: MIT Press.

Glymour, C. (2003). "Learning, Prediction and Causal Bayes Nets." *Trends in Cognitive Sciences*, *7*(1): 43–8.

Gopnik, A., C. Glymour, D.M. Sobel, L.E. Schulz, T. Kushnir, and D. Danks (2004). "A theory of causal learning in children: Causal maps and bayes nets." *Psychological Review*, *111*(1): 3–32.

Griffiths, T.L. and J.B. Tenenbaum (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4): 334–84.

Halpern, J.Y. and C. Hitchcock (2015). "Graded causation and defaults." *The British Journal for the Philosophy of Science*, *66*(2): 413–57.

Halpern, J.Y. and J. Pearl (2005). "Causes and explanations: A structural-model approach. Part I: Causes." *The British Journal for the Philosophy of Science*, *56*(4): 843–87.

Hitchcock, C. (2007). "Prevention, preemption, and the principle of sufficient reason." *The Philosophical Review*, *116*(4): 495–532.

Hitchcock, C. (2009). "Causal modelling." In H. Beebee, C. Hitchcock, and P. Menzies (eds.), *The Oxford Handbook of Causation* (299–314). New York: Oxford University Press.

Hume, D. (1748/1975). *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

Johnson, S.G. and W. Ahn (2017). "Causal mechanisms." In M.R. Waldmann (ed.), *The Oxford Handbook of Causal Reasoning* (127–46). New York: Oxford University Press.

Johnson, S.G. and F.C. Keil (2018). "Statistical and mechanistic information in evaluating causal claims." In T.T. Rogers, M. Rau, X. Zhu, and C.W. Kalish (eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (618–23). Austin, TX: Cognitive Science Society.

Kahneman, D. and D.T. Miller (1986). "Norm theory: Comparing reality to its alternatives." *Psychological Review*, *93*(2): 136–53.

Kominsky, J.F., J. Phillips, T. Gerstenberg, D. Lagnado, and J. Knobe (2015). "Causal superseding." *Cognition*, *137*: 196–209.

Lagnado, D.A. and M. Speekenbrink (2010). "The influence of delays in real-time causal learning." *The Open Psychology Journal*, *3*(1): 184–95.

Liljeholm, M. and P.W. Cheng (2007). "When is a cause the 'same'? Coherent generalization across contexts." *Psychological Science*, *18*(11): 1014–21.

Meder, B., R. Mayrhofer, and M.R. Waldmann (2014). "Structure induction in diagnostic causal reasoning." *Psychological Review*, *121*(3): 277–301.

Novick, L.R. and P.W. Cheng (2004). "Assessing interactive causal influence." *Psychological Review*, *111*(2), 455–85.

Paul, L.A. and E.J. Hall (2013). *Causation: A User's Guide*. New York: Oxford University Press.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

Rehder, B. (2014). "Independence and dependence in human causal reasoning." *Cognitive Psychology*, *72*: 54–107.

Samland, J. and M.R. Waldmann (2016). "How prescriptive norms influence causal inferences." *Cognition*, *156*: 164–76.

Samland, J., M. Josephs, M.R. Waldmann, and H. Rakoczy. (2016). "The role of prescriptive norms and knowledge in children's and adults' causal selection." *Journal of Experimental Psychology: General*, *145*(2): 125.

Shortle, J.F., J.M. Thompson, D. Gross, and C.M. Harris (2018). *Fundamentals of Queueing Theory*. New Jersey: John Wiley & Sons.

Sloman, S. (2005). *Causal Models: How People Think About the World and its Alternatives*. New York: Oxford University Press.

Stephan, S. and M.R. Waldmann (2018). "Preemption in singular causation judgments: A computational model." *Topics in Cognitive Science*, *10*(1): 242–57.

Stephan, S. and M.R. Waldmann (2022). "The role of mechanism knowledge in singular causation judgments." *Cognition*, 2018, 104924.

Stephan, S., R. Mayrhofer, and M.R. Waldmann (2020). "Time and singular causation—a computational model." *Cognitive Science*, *44*(7): e12871.

Stephan, S., K. Tentori, S. Pighin, and M.R. Waldmann (2021). "Interpolating causal mechanisms: The paradox of knowing more." *Journal of Experimental Psychology: General*.

Waldmann, M.R. (2000). "Competition among causes but not effects in predictive and diagnostic learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1): 53–76.

Waldmann, M.R. (ed.). (2017). *The Oxford Handbook of Causal Reasoning*. New York: Oxford University Press.

Waldmann, M.R. and Y. Hagmayer (2013). "Causal reasoning." In D. Reisberg (ed.), *Oxford Handbook of Cognitive Ppsychology* (733–52). New York: Oxford University Press.

Waldmann, M.R., P.W. Cheng, Y. Hagmayer, and A.P. Blaisdell (2008). "Causal learning in rats and humans: a minimal rational model." In N. Chater and M. Oaksford (eds.), *Prospects for Bayesian Cognitive Science* (453–84). Oxford: Oxford University Press.

Wolff, P. (2007). "Representing causation." *Journal of Experimental Psychology: General*, *136*(1): 82–111.

Wolff, P., A.K. Barbey, and M. Hausknecht (2010). "For want of a nail: How absences cause events." *Journal of Experimental Psychology: General*, *139*(2): 191–221.

Wolff, P. and R. Thorstad (2017). "Force dynamics." In M.R. Waldmann (ed.), *The Oxford Handbook of Causal Reasoning* (147–68). New York: Oxford University Press.

Woodward, J. (1990). "Supervenience and singular causal statements." *Royal Institute of Philosophy Supplement*, *27*: 211–46.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, J. (2011). "Mechanisms revisited." *Synthese*, *183*(3): 409–27.