Title

Causal scope and causal strength:The number of potential effects of a cause influences causal strength estimates

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors

Stephan, Simon
Waldmann, Michael R.

Publication Date

2020

Peer reviewed

Causal scope and causal strength: The number of potential effects of a cause influences causal strength estimates

Simon Stephan (sstepha1@gwdg.de)
Michael R. Waldmann (michael.waldmann@bio.uni-goettingen.de)
Department of Psychology, University of Göttingen,
Gosslerstr. 14, 37073 Göttingen, Germany

Abstract

Causal scope, the number of different effects a cause can produce, is a salient feature of causes. In the present research, we address the question whether reasoners use causal scope as a diagnostic cue to infer the strengths of individual causal links. In three experiments, we manipulated the number of effects of a cause, and asked subjects to assess the causal strengths of single causal links. The results document a clear influence of causal scope on perceived link strength. In particular, subjects tended to display a "dilution" effect. They perceived a causal link to be weaker if that link belonged to a cause that is capable of producing additional effects. This dilution effect can be explained by a dispositional notion of causality according to which a cause possesses a certain amount of causal power or capacity that it distributes across its different causal pathways.

Keywords: causality; causal strength; causal structure; scope; causal reasoning

Knowledge about the strengths of causal relationships is important for predictions, diagnoses, the selection of effective interventions, and causal attributions. In the literature on causal reasoning, causal strength is often understood as the probability with which a cause generates its effect (e.g., Cheng, 1997; Cheng & Lu, 2017; Griffiths & Tenenbaum, 2005). According to the causal Bayes net framework, for example, the normative strategy a reasoner should apply to infer the strength of a causal link is to assess the contingency (ΔP) between the potential cause and effect factor. If a target cause combines linearly with other causes to produce an effect, it is assumed that ΔP provides a direct measure of causal strength, whereas when different causes of an effect combine according to a *noisy-OR* gate (Pearl, 2000), the correct measure of causal strength is assumed to be given by a scaled version of ΔP that Cheng (1997) has called causal power: $\Delta P/(1-P(e|\neg c)$. Various studies have shown that reasoners' causal strength inferences are influenced by contingency information (see Waldmann, 2017).

In the present research, we investigated whether reasoners' inferences about the strengths of causal links are also influenced by cues other than the statistical association between a target cause and its effect. One such cue on which we focus in this paper is illustrated by the following example: Imagine two different food supplements, A and B, that both are known to be able to improve certain physiological parameters. Supplement A possesses the capacity to cause (1) improved muscle strength, (2) improved joint flexibility, and (3) improved bone density. By contrast, supplement B only possesses the capacity to cause improved bone density. Based on this information about the number of different effects the two supplements can produce, is there a systematic difference in their capacity to lead to improved bone density?

The number of different effects a cause can generate can be called the cause's *causal scope*. In the present research, we investigated whether reasoners regard causal scope as a diagnostic indicator for the strengths of cause-effect links. Although causal scope is a salient feature of causes, the question which role it plays in reasoners' strength inferences has largely been neglected in the literature (but see Johnson, Johnston, Toig, & Keil, 2014; Sussman & Oppenheimer, 2020).

According to causal theories belonging to the dependency framework (see Waldmann & Mayrhofer, 2016, for an overview), such as Bayes net theory, the answer to the question raised above is clear. The strength of the link between a cause and an effect cannot be determined unless something is known about their statistical association. Structural information about the number of effects of a cause should not affect strength estimates under this framework unless specific additional assumptions are made about the mechanisms linking causes and effects. If only provided with the structure of the causal model, the theory does not predict systematic strength differences in the two cases.

Another class of theories allowing for different predictions is the disposition framework of causality (see Kistler & Gnassounou, 2007; Mumford & Anjum, 2011). According to the dispositional view of causality, causes bring about their effects in virtue of an intrinsic causal property, its causal capacity, which explains the observed statistical regularities. For example, the causal relation between ingesting aspirin and experiencing pain relief is explained by dispositional accounts by saying that aspirin possesses an intrinsic causal disposition to bring about pain relief that may typically be dormant but can become manifest under the right conditions (e.g., when ingested by a human body).

One hypothesis that can be derived from a dispositional notion of causality is that reasoners may assume that causes possess a limited amount of causal capacity that is distributed across the different causal paths linked to the cause. This hypothesis leads to the prediction of a dilution effect according to which reasoners will have a tendency to perceive a causal link as weaker if it belongs to a cause with broad causal scope than if it belongs to a cause with narrow causal scope, because in the former case the causal capacity would be perceived to be distributed across the cause's different causal paths. The opposite prediction is also conceivable, however. If reasoners assume that causes have a certain amount of causal capacity, they may use the number of effects that a cause can produce as a diagnostic cue for the amount of causal capacity this cause possesses. A causal link that is part of a cause with broad causal scope might then be perceived to be stronger than a link that is part of a cause with narrow causal scope.

We report three experiments testing the influence of causal scope on causal strength inferences. To foreshadow our results, we found clear evidence for an influence of causal scope on causal strength judgments. In line with the first hypothesis of a dilution effect, Experiment 1 revealed that reasoners seem to assume that the strength of individual links in a causal structure decreases with the number of effects a cause can produce. Experiments 2 and 3 tested the robustness of this "dilution" effect. Experiment 2 compared conditions in which causes that differ in causal scope were either evaluated jointly or separately. The goal was to see whether the causal scope effect found in Experiment 1 can only be observed in situations in which reasoners can directly compare broad-scope with narrow-scope causes or whether the effect can also be found in contexts in which only one target cause is presented. Following the causal strength definition of causal Bayes nets, we used a probabilistic phrasing of the causal strength test query in Experiments 1 and 2. In Experiment 3, we replicated the the dilution effect with a causal strength test query formulated in a way that might be more natural given the effect variables that we used in our experimental scenario. The experimental materials and data for all experiments can be accessed via `https://osf.io/mjswc/`.

Experiment 1

The goal of Experiment 1 was to test whether we would see an influence of causal scope on causal strength judgments. We used a fictitious scenario in which subjects were asked to compare two causes that differed in their causal scope. The multiple-effects cause was described as being capable of producing three effects, while the single-effect cause was described to be able to produce only one of the effects of the multiple-effects cause. The crucial manipulation was whether the target causal link whose strength subjects were asked to estimate was generated by a multiple-effects or a single-effect cause. Additionally, we varied whether the multiple-effects cause was capable of producing three effects belonging to the same physiological domain (the musculoskeletal system) or to different physiological domains.

Methods

Participants One hundred and sixty subjects ($M_{age} = 35.00$, $SD_{age} = 13.18$, age range 18 to 73 years, 108 female, 49 male, three persons selected the option "other" for the gender question) who were recruited via Prolific (`www.prolific.co`) participated in this online study and provided valid data. The inclusion criteria were a minimum age of 18 years, English as native language, and an approval rate concerning participation in previous studies of 90 percent. Subjects were asked to participate only via laptop or desktop computer and not via smartphone or tablet, because we wanted to minimize the chances that subjects take part who are in environments (e.g., public places, subway) that might distract them. Subjects received a monetary compensation for their successful participation.

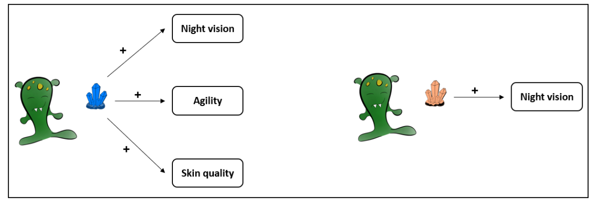

Figure 1: Example of the stimuli presented in Experiment 1.

Design, Materials, and Procedure The study had a 2 (target cause to which the strength query referred: multiple-effects vs. single-effect cause; within subject) $\times$ 2 (physiological domain of effects of multiple-effect cause: same vs. different; between subjects) mixed design. Additional control factors that were counterbalanced between subjects will be introduced below.

Subjects were presented with a fictitious scenario about aliens from a remote planet whose diet involves two different types of crystals (blue vs. red). It was described that the aliens forage for these crystals because these crystals were capable of causing positive physiological effects. One type of crystal was described as having capacity to produce several beneficial effects in the aliens' body, while the the other type of crystal was described has having the capacity to only produce a single positive effect. An example of the graphical illustrations shown to subjects is depicted in Fig. 1. An example of the scenario descriptions that were presented together with the graphical illustrations is:

> A remote planet of our galaxy, called Morgania, is inhabited by tiny aliens, called Morganians. The regular diet of Morganians consists of plants of different types. From time to time, however, Morganians also seek and swallow particular crystals because these crystals can lead to positive effects. There are two different types of crystals, blue and red. Blue crystals are very valuable because they possess the capacity to produce multiple positive effects. Blue crystals can lead to improved night vision, improved agility, and improved skin quality. Red crystals are less valuable because they only possess the capacity to produce a single positive effect. Red crystals can only lead to improved night vision.

This vignette was shown in the condition in which the three effects of the multiple-effects cause belonged to three different domains. In the condition in which the effects of the multiple-effects cause belonged to the same physiological domain, subjects learned that the multiple-effects cause could cause (1) improved muscle strength, (2) improved bone density, and (3) improved joint flexibility. Whether red or blue crystals were presented as the the multiple-effects cause was counterbalanced between subjects. The single-effect cause was described as having the capacity to produce one of the three effects of the multiple-effects cause.

After subjects had learned about the two causes, they answered a causal-strength question for each cause. The formulation of the causal strength question we used in this study assumed a probabilistic interpretation of causal strength (Cheng, 1997): Subjects were asked to indicate the probability with which eating the particular crystal would lead to the target effect. For example, subjects were asked: "We now would like to get to know your intuition about the causal strength with which eating blue crystals causes improved night vision. To express your intuition about the causal strength, please answer the following question: What do you think is the probability with which eating red [blue] crystals leads to improved night vision?". Responses were given on an eleven-point rating scale whose endpoints were labeled "It never leads to improved night vision" and "It always leads to improved night vision" (the midpoint of the scale was labeled "50:50"). Which of the three effects was selected as the target effect was counterbalanced between subjects. The order in which the strength questions were asked was also counterbalanced between subjects. After subjects had given their causal strengths ratings, we also asked them on a separate screen to provide a short explanations for their ratings. Subjects provided their explanations as open-ended responses. Subjects finally provided demographic data and were debriefed.

Results and Discussion

Fig. 2 shows the results. As can be seen in the figure, subjects' causal strength judgments were sensitive to our manipulation of causal scope. Particularly, subjects tended to infer a higher causal strength for the target link when that link belonged to the single-effect cause ($M = 0.81$, $SD = 0.22$ and $M = 0.80$, $SD = 0.29$ for "same domain" and "different domains" condition, respectively) than when the link belonged to the multiple-effects cause ($M = 0.68$, $SD = 0.25$ and $M = 0.69$, $SD = 0.26$ for "same domain" and "different domains" condition, respectively). At the same time, Fig. 2 shows that this effect was not influenced by our domain manipulation. The same pattern was found irrespective of whether the three effects belonged to the same domain or to different domains. A 2 (target cause to which the strength query referred: multiple-effects vs. single-effect cause; within subject) $\times$ 2 (physiological domain of effects of multiple-effect cause: same vs. different; between subjects) mixed ANOVA confirmed the pattern shown in Fig. 2. We only found a significant main effect for the target cause factor, $F(1,158) = 18.95$, $p < .001$, which fell into the range between a small and medium effect size $f = 0.24$.

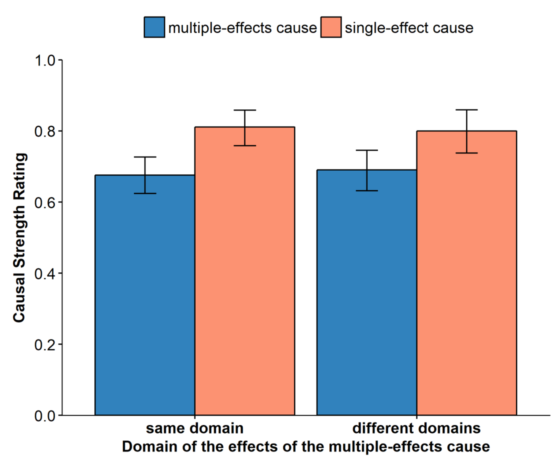

Figure 2: Results (means and 95% CIs) of Experiment 1.

To learn more about the nature of the observed causal scope effect, we next analyzed subjects' individual ratings as well as their open-ended explanations. It was found that of the 160 subjects we tested 91 (57%) gave different causal strength ratings for the two causes, while 69 (43%) subjects did not differentiate between them. Of the 91 subjects who rated the strengths of the two causes differently, the ratings of a majority of 74 subjects (81%) were in line with a dilution effect. These participants gave lower causal strength ratings for the multiple-effects cause. The ratings of 17 of these 91 participants (19%) showed the opposite effect, suggesting that some reasoners seem to use the number of effects a cause can produce as a diagnostic indicator of its overall causal capacity.

Subjects' open-ended responses were grouped into four different categories by two independent raters. The open-ended responses as well as the coding can be accessed via `https://osf.io/ye6rs/`. Responses classified as Category 1 were those that implied a dilution effect. Category 2 responses described the reverse effect according to which the number of effects a cause can generate is diagnostic for its overall causal capacity. Subjects who answered that a causal strengths cannot be inferred based on the provided information alone were assigned to Category 3. Subjects whose explanations were unclear or referred to completely different factors were assigned to Category 4. In an initial round, both raters made identical classifications in 94% of the cases (i.e., for 151 out of the 160 responses). The raters then met to discuss and decide about the remaining nine cases. The coding of subjects' explanations yielded the following results: it was found that 51 explanations implied a dilution effect (Category 1). For example, a prototypical response given by one subject who was assigned to this category was "Because the red crystals cause multiple effects, it seems reasonable that those effects would be more diluted than the effect of the blue crystal, which only causes improved muscle strength. I realize this is not necessarily true, but that is what my intuition says." Interestingly, it turned out that the responses that implied a dilution effect could be sub-categorized further into two different sub-categories. The largest sub-group of participants tended to provide answers like the one just mentioned, in which it was described that the capacity of the multiple-effects cause is spread evenly across its attached causal pathways. Another smaller sub-group of subjects, however, described a different mechanism that also implies a dilution effect. These subjects assumed that the multiple-effects cause would have an overall lower causal strength on the target effect because it would be capable of producing only one of its effects at a time, which in turn would reduce the probability that it exerted an influence on the target effect in the particular test case subjects were asked to evaluate. For example, a prototypical answer given by subjects in this sub-category was "The blue crystals, when eaten, have a one in three chance of producing clearer skin because there are two other possible outcomes. The red crystals only produce one outcome, which is clearer skin, and so are effective in doing so 100% of the time.".

Seven participants offered explanations that described an opposite, strengthening effect (Category 2). For example, one participant assigned to this category wrote "I feel like the crystals with more benefits would be stronger and give a stronger effect for each benefit". 47 subjects were assigned to Category 3. A prototypical response given by a participant in this category was "The information stated that they both could improve joint flexibility but it did not state how much. There was no known evidence to explain why this was possible. They both made a claim and I thought there was a 50% chance of both working from the information." The remaining 55 responses were coded as Category 4. For example, one subject in this category wrote "You use infra red light for night vision hence my assumption that red crystal will improve you night vision".

The results of this study show that at a substantial number of reasoners seems to treat the number of effects a cause can generate as a cue for the causal strength of individual causal links. More specifically, the results of the experiment document that most reasoners seem to assume dilution effect, as they seem to have a tendency to perceive the number of links in a causal structure to be inversely related to the strength of individual links. This effect is in line with a mental representation of causal strength according to which a cause possesses a fixed amount of causal capacity that is distributed across its causal pathways. The more pathways a cause serves, the weaker each causal influence becomes. However, we also found a reverse, strengthening effect in a small sub group of participants, which indicates that some reasoners seem to treat a cause's causal scope as a diagnostic indicator for its overall causal capacity. Finally, we did not find evidence that the observed effect of causal scope on causal strength inference is moderated by effect domain diversity.

Experiment 2

One noteworthy aspect of the design of Experiment 1 is that subjects were presented with both types of causes and were asked to give causal strength estimates for both of them. A question that arises is whether the observed effect only occurs when reasoners directly think about both types of causes or whether it can also be found if only one type of cause is considered at a time. The joint presentation format used in Experiment 1 might have encouraged subjects to focus on relative causal strength. Thus, the assumption that a cause has a fixed amount of power that is spread across its different causal paths may be particularly salient when a context encourages a direct comparison between causes with different causal scope. Another aspect was that subjects were asked to provide strength estimates for both causes, which may have encouraged them to think about a distinction between the two. The goal of Experiment 2 was to address these issues and test the robustness of the causal scope effect. We decided to contrast a condition in which subjects were jointly presented with a single-effect and the multiple-effects cause with one in which they only learned about either the single-effect or the multiple-effects cause. Moreover, we decided to let subjects rate the causal strength of only one of the causes. Since we did not observe an effect in Experiment 1 of whether the three effects of the multiple-effects cause belonged to the same domain or not, we dropped this factor in Experiment 2.

Methods

Participants Seven hundred and twenty subjects (M_{age} = 35.61, SD_{age} = 12.41, age range 18 to 79 years, 416 female, 299 male, four subjects selected the option "other" for the gender question, one subject did not want to provide gender information) who were recruited via Prolific (`www.prolific.co`) participated in this online study and provided valid data. The rationale behind this large sample size was that we assumed that the difference in perceived causal strength between the multiple-effects and the single-effect cause might be smaller if subjects learn about only one of the two causes than if they simultaneously learn about both of them. We wanted to be sure to detect a potentially small ordinal interaction effect of $f = 0.11$ with more than 80 percent power. The applied inclusion criteria were the same as in Experiment 1. Prolific workers who took part in Experiment 1 were excluded from participation. Subjects received a monetary compensation for their participation.

Design, Materials, and Procedure The study had a 2 (presentation format: joint presentation of single-effect and multiple-effects cause vs. presentation of only one of the two types of causes) $\times$ 2 (target cause: single-effect vs. multiple-effects cause) between-subjects design.

The experimental materials and procedure were largely identical to those in Experiment 1 except that half of the subjects (in the separate evaluation condition) learned about only one crystal that the aliens were described to forage for. We also changed some parts of the scenario description. In Experiment 1, the multiple-effects cause was described to be "more valuable" because it could lead to multiple positive effects, while the single-effect cause was described as "less valuable" because it could only lead to a single positive effect. While this evaluative description was intended to highlight merely the difference in causal scope, it might have influenced subjects' assumptions about causal strength. A cause that is described as "valuable" may be perceived to be stronger than a cause that is described as "less valuable". This phrasing may have attenuated the observed dilution effect. To have neutral formulations in Experiment 2, we therefore left out the "[...] are very valuable because they [...]" and the "[...] are less valuable because they only [...]" parts of the respective sentences.

After subjects had learned about the cause(s), they were shown the causal strength query, which either referred to the multiple-effects or the single-effect cause. Like in Experiment 1, the target effect was manipulated between subjects (improved "night vision" vs. "muscle strength" vs. "bone density"). After subjects had given their strength-rating, they reported some demographic data and were debriefed.

Results and Discussion

The results are summarized in Fig. 3. The left two bars show the mean causal strength ratings for the conditions in which subjects learned about both causes, whereas the right two bars show the mean strength ratings for the conditions in which either the multiple-effects or the single-effect cause were presented. First, it can be seen that the ratings show that we successfully replicated the dilution effect found in Experiment 1. Subjects who were presented with both the multiple-effects and the single-effect cause tended to give higher causal strength ratings for the target link when that link belonged to the single-effect cause than when it belonged to the multiple-effects cause ($M = 0.77$, $SD = 0.21$ vs. $M = 0.60$, $SD = 0.27$). A similar pattern was observed for the condition in which each participant learned about only one of the two causes ($M = 0.67$, $SD = 0.22$ vs. $M = 0.60$, $SD = 0.22$). However, Fig. 3 shows that the observed dilution effect was weaker in this case. If subjects only learned about one of the two causes, the differences in perceived causal strength for the target link tended to become smaller. It can also be seen that this reduction in the perceived causal-strength difference was driven by an attenuation of causal strength ratings given for the single-effect cause, whereas the strength ratings for the multiple-effects cause were almost identical to those given in the condition in which subjects were shown both causes.

A 2 (presentation format: joint presentation of single-effect and multiple-effects cause vs. presentation of only one of the two types of causes) $\times$ 2 (target cause: single-effect vs. multiple-effects cause) between-subjects ANOVA confirmed the descriptive pattern displayed in Fig. 3. The analysis yielded a significant main effect of "target cause", $F(1,716) = 47.16$, $p < .001$, $f = 0.26$, confirming that subjects overall tended to give higher strength ratings for the single-effect cause. Planned comparisons confirmed that the observed strength difference between the multiple-effects and the single-effect cause was significant for both presentation formats, $t(716) = 6.99$, $p < .001$, $d = 0.71$ and $t(716) = 2.72$, $p < .01$, $d = 0.3$. We also found a significant main effect of "presentation format", $F(1,716) = 9.52$, $p < .01$, $f = 0.12$: strength ratings were overall higher in the condition in which subjects were presented with both causes. As Fig. 3 shows, this main effect is driven by the difference in the causal strength ratings given for the single-effect cause. Finally, we found a significant interaction effect between "presentation format" and "target cause", $F(1,716) = 9.52$, $p < .01$, $f = 0.11$, confirming that the observed difference in strength ratings was indeed smaller in the condition in which subjects were presented with only one type of cause.

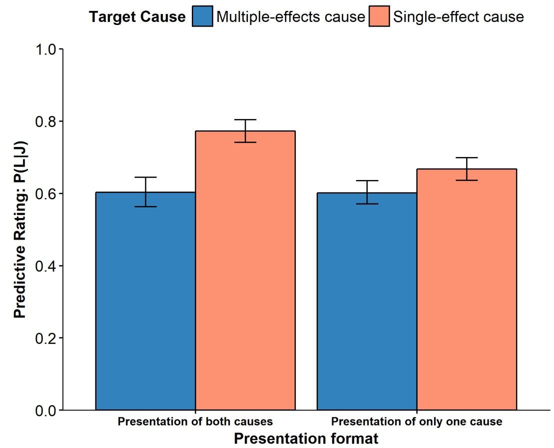

Figure 3: Results (means and 95% CIs) of Experiment 2.

In sum, Experiment 2 replicated the dilution effect we found in Experiment 1. Subjects generally tended to attribute lower causal strength to an individual causal link if that link belongs to a cause that possesses the capacity to produce further additional effects. We also found that the magnitude of this effect depends on whether subjects are in a context in which differences in causal scope are particularly salient or not. When subjects observe different types of causes, the impact of causal scope on perceived causal strength is more pronounced than when they only see one type of cause.

Experiment 3

In Experiments 1 and 2 we used a formulation for the causal power test queries that was based on the causal Bayes net interpretation of causal power according to which causal power represents the probability with which a cause generates an effect. This probabilistic reading of causal power appears to be natural if causes and effects represent binary events. However, in the scenario we used in Experiments 1 and 2, it appears to be more natural to think of the described effect factors (e.g., muscle strength) as continuous rather than binary variables and to also conceptualize causal power as a quantity exerting a gradual influence on its effects. The primary goal of Experiment 3 therefore was to replicate the dilution effect with a causal strength test query that conceptualizes the involved effects as continuous variables. Our hypothesis was that we would see a larger effect of causal scope in this case. Like in Experiment 1, we also again manipulated effect domain diversity (i.e., whether the effects produced by the multiple-effects cause belonged to the same or different physiological domains) because we wanted to see whether we would see a domain diversity effect with the new causal strength test query.

Methods

Participants One hundred and twenty subjects ($M_{age} = 32.94$, $SD_{age} = 12.81$, age range 18 to 69 years, 74 female, 46 male) who were recruited via Prolific (`www.prolific.co`) participated in this online study and provided valid data. The applied inclusion criteria were the same as in the previous experiments. Prolific workers who participated in Experiments 1 and 2 were excluded from participation. Subjects received a monetary compensation for their successful participation.

Design, Materials, and Procedure The study had a 2 (target cause: multiple-effects cause vs. single-effect cause) × 2 (physiological domain of effects of multiple-effect cause: same vs. different) between-subjects design.

We used the same cover story about the aliens and the crystals as in the previous experiments. The scenario description introducing the two types of causes was identical to the one used in the "joined presentation" condition of Experiment 2. The three effects that were introduced in the condition in which the multiple-effects cause was described to generate effects belonging to different physiological domains were "night vision", "joint flexibility", and "alertness". In the "same effect domain" condition, the three effects were "muscle strength", "bone density", and "joint flexibility". Whether the blue or the red crystal was introduced as the multiple-effects cause was counterbalanced between subjects.

After subjects had read the information about the two types of causes and the effects each cause is able to produce they were presented with the causal strength test query. The query either referred to the multiple-effects or to the single-effect cause. Which of the three effects of the multiple-effects cause was selected as the target cause was counterbalanced between subjects. Unlike in the previous studies, the causal strength query subjects were asked conceptualized the effects as continuous variables. For example, in the condition in which the target effect was "night vision" the question read: "We now would like to get to know your intuition about the causal strength with which eating blue crystals causes improved night vision. To express your intuition about the causal strength, please answer the following question: If an alien eats a blue [red] crystal, how much do you think will its night vision improve?" Another difference from the previous studies was that subjects this time provided their causal strength ratings on a continuous slider instead of a discrete rating scale whose endpoints were labeled "not at all" and "maximally".

Results and Discussion

The theoretically relevant results are summarized in Fig. 4 and in Table 1. As can be seen there, we successfully replicated the dilution effect with the new causal strength test query. A 2 (target cause: multiple-effects cause vs. single-effect cause) × 2 (physiological domain of effects of multiple-effect cause: same vs. different) yielded a significant main effect for the target cause factor, $F(1,116) = 20.97$, $p < .001$. Further, as we hypothesized, the observed dilution effect was stronger with our new continuous causal strength query than with the probabilistic causal strength query we had used in Experiments 1 and 2. While the effects size of the dilution effect was found to be $f = 0.24$ in Experiment 1, we this time measured an effect size that was almost twice as large, $f = 0.43$. Planned directed contrasts further confirmed that the dilution effect was significant in both the "different effect domains" ($t(116) = 3.68$, $p < .001$, $d = 1.24$) and the "same effect domain" condition ($t(116) = 2.80$, $p < .01$, $d = 0.65$). Figure 4 also shows that, unexpectedly, ratings were overall higher in the "different effect domains" condition than in the "same effect domain" condition, $F(1,116) = 5.58$, $p < .02$, $f = 0.22$. However, like in Experiment 1, the effect domain manipulation had no influence on the magnitude of the dilution effect ($F(1,116) = 0.39$, $p = .53$, $f = 0.06$ for the interaction effect).

Table 1: Results of Experiment 3

	different effect domains		same effect domain	
	multiple-effects cause	single-effect cause	multiple-effects cause	single-effect cause
Mdn	0.49	0.78	0.39	0.68
M	0.57	0.79	0.49	0.66
SD	0.25	0.18	0.24	0.26
95% CI	[0.48; 0.66]	[0.72; 0.86]	[0.40; 0.58]	[0.61; 0.71]

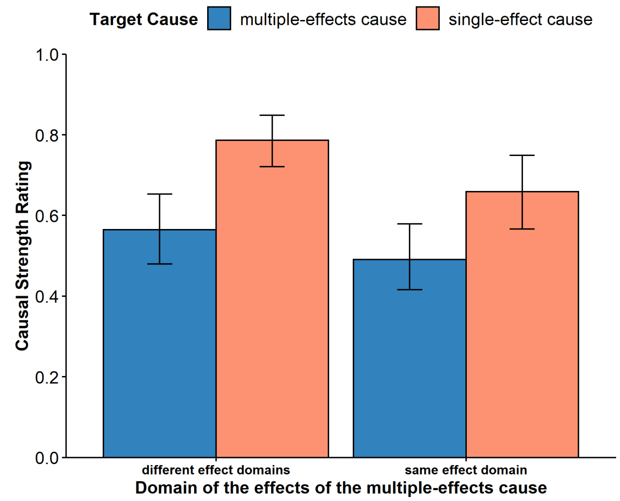

Figure 4: Results (means and 95% CIs) of Experiment 3.

To see whether the unexpected main effect of "effect domain diversity" was driven by a particular target effect subjects had evaluated, we also looked at the ratings subjects made for the different physiological effects the crystals were described to be able to generate. These ratings are summarized in Fig. 5. As can be seen there, the main effect of "effect domain diversity" was not driven by a particular target effect. Rather, ratings tended to be higher for each of the three different target effects in the "different effect domains" condition.

General Discussion

The experiments presented in this paper show that reasoners use causal scope as a cue for causal strength. More specifically, we have found a dilution effect in our three experiments. Individual causal links tend to be perceived as weaker if these links belong to causes that can generate further effects. This dilution effect can be explained by the hypothesis that reasoners adopt a dispositional view of causality and

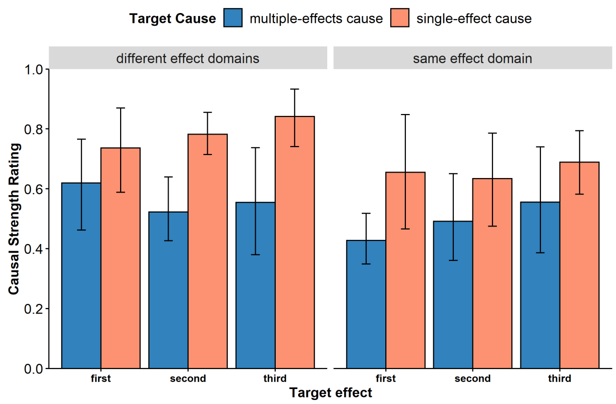

Figure 5: Results (means and 95% *CI*s) of Experiment 3 showing the causal strength ratings subjects made in the different target effect conditions.

assume that a cause possesses a particular amount of causal capacity that it distributes across its causal paths.

However, we also found in Experiment 1 that a minority of subjects (about 20%) tended to indicate an effect in the opposite direction (strengthening), suggesting that some reasoners seem to treat the number of effects a cause can generate as a diagnostic indicator for its overall causal capacity. In future studies, it would be interesting investigate what the factors are that determine whether reasoners tend to adopt one view (leading to a dilution effect) or the other (leading to a strengthening effect). Interestingly, Sussman and Oppenheimer (2020) have found in a recent set of studies that one factor that seems to play a role in this respect, which we did not test in the present studies, is whether the effects a multiple-effects cause can produce are positive or negative. In line with the results we reported here, Sussman and Oppenheimer (2020) found that positive effects tended to elicit a dilution effect. If the multiple-effects cause was described to lead to negative effects (e.g., skin irritation), however, the reverse effect was found. In this case, subjects tended to give higher causal strength ratings for the multiple-effects than for the single-effect cause. However, a noteworthy aspect of their studies was that subjects tended to assume additional generative causal links among the effects of the multiple-effects cause (i.e., subjects had a tendency to infer that the effects of the multiple-effects cause would mutually amplify each other). Furthermore, this tendency was particularly pronounced in the condition in which the causes were described to lead to negative effects. It would therefore be interesting to test in a future study whether the strengthening effect observed when the different causes produce negative effects can be replicated with scenarios in which it is implausible that the effects of the multiple-effects cause are themselves causally connected to each other. A further idea for future studies would be to test cases in which positive and negative effects are combined. For example, in addition to their desired primary effects almost all highly effective medicines also tend to lead undesired side effects. Our prediction in this case would be that reasoners will tend to perceive the link between the cause and the target effect as stronger if that cause generates further undesired side effects.

One obvious limitation in our studies is that we have so far only tested the effect in one particular domain, physiology. A straightforward question that we plan to address in future studies therefore is whether the effect also occurs in other domains. One interesting domain that we plan to test are artifacts.

Finally, we have so far only contrasted two levels of causal scope (one effect versus three effects). An interesting question for future studies is whether the negative relation we observed between causal scope and causal strength is monotonous or not. We have mentioned in the beginning that a possible hypothesis that can be derived from the dispositional view is that the number of effects a cause can generate may be treated as a signal by reasoners for the amount of causal capacity a cause possesses. Moreover, we have found in Experiment 1 that about one fifth of our participants responded in that way. It would be interesting to test whether the dilution effect can be attenuated or even reversed by drastically increasing the number of effects.

Acknowledgments This work was supported by the Deutsche Forschungsgemeinschaft (DFG) Grant WA 621/24-1.

References

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

Cheng, P. W., & Lu, H. (2017). Causal invariance as an essential constraint for creating a causal representation of the world: Generalizing the invariance of causal power. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 65–84). New York: Oxford University Press.

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

Johnson, S., Johnston, A., Toig, A., & Keil, F. (2014). Explanatory scope informs causal strength inferences. In P. Bello, M. Guarini, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2453–1558).

Kistler, M., & Gnassounou, B. (2007). *Dispositions and causal powers*. Aldershot, UK: Ashgate.

Mumford, S., & Anjum, R. L. (2011). *Getting causes from powers*. Oxford University Press.

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

Sussman, A. B., & Oppenheimer, D. M. (2020). The effect of effects on effectiveness: A boon-bane asymmetry. *Cognition*, *199*, 104240.

Waldmann, M. R. (Ed.). (2017). *The Oxford handbook of causal reasoning*. New York: Oxford University Press.

Waldmann, M. R., & Mayrhofer, R. (2016). Hybrid causal representations. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 85–127). New York: Academic Press.