Cognitive Psychology 140 (2023) 101540

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

The perceived dilution of causal strength

Simon Stephan [a,*], Neele Engelmann [b], Michael R. Waldmann [a]

[a] Department of Psychology, University of Göttingen, Gosslerstrasse 14, 37073 Göttingen, Germany
[b] Center for Law, Behaviour, and Cognition, Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany

ARTICLE INFO

Keywords:
Causal strength
Causal structure
Dilution
Causal reasoning
Causal Bayes nets
Causal scope

ABSTRACT

Dependency theories of causal reasoning, such as causal Bayes net accounts, postulate that the strengths of individual causal links are independent of the causal structure in which they are embedded; they are inferred from dependency information, such as statistical regularities. We propose a psychological account that postulates that reasoners' concept of causality is richer. It predicts a systematic influence of causal structure knowledge on causal strength intuitions. Our view incorporates the notion held by dispositional theories that causes produce effects in virtue of an underlying causal capacity. Going beyond existing normative dispositional theories, however, we argue that reasoners' concept of causality involves the idea that continuous causes spread their capacity across their different causal pathways, analogous to fluids running through pipe systems. Such a representation leads to the prediction of a structure-dependent *dilution* of causal strength: the more links are served by a cause, the weaker individual links are expected to be. A series of experiments corroborate the theory. For continuous causes with continuous effects, but not in causal structures with genuinely binary variables that can only be present or absent, reasoners tend to think that link strength decreases with the number of links served by a cause. The effect reflects a default notion reasoners have about causality, but it is moderated by assumptions about the amount of causal capacity causes are assumed to possess, and by mechanism knowledge about how a cause generates its effect(s). We discuss the theoretical and empirical implications of our findings.

1. Introduction

Imagine two different chemical substances, C_1 and C_2, that, when ingested over a longer period of time, are known to cause the degeneration of red blood cells (E_3) in the human body. Also, while chemical C_1 is known to have two additional physiological effects (E_1 and E_2), and therefore can be called a *common cause* of three effects, chemical C_2 only causes a degeneration of red blood cells and nothing else. An abstract causal structure of this scenario is shown in Fig. 1. Based on this causal structure information, if you consider the causal strength with which chemicals C_1 and C_2 lead to the degeneration of red blood cells (E_3), would you intuitively think that one of them has a stronger causal influence on the degeneration of red blood cells? If so, does chemical C_1, which in addition to the degeneration of red blood cells (E_3) also influences two other effects (E_1 and E_2), seem to have a stronger degenerative effect on red blood cells than C_2, or does chemical C_2, which only causes the degeneration of red blood cells (E_3) seem stronger?

Our mental representations of causal relationships (see Sloman, 2005; Waldmann, 2017, for overviews) involve different dimensions. One of these dimensions is causal structure, which encodes what variables are causally connected. If we observe a reliable correlation between two variables A and B, it is known that this correlation may obtain because A causes B ($A \rightarrow B$),

* Corresponding author.
E-mail address: simon.stephan@psych.uni-goettingen.de (S. Stephan).

https://doi.org/10.1016/j.cogpsych.2022.101540
Received 11 January 2022; Received in revised form 30 August 2022; Accepted 3 December 2022
Available online 15 December 2022

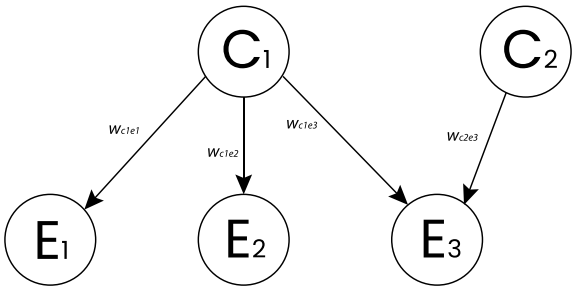

Fig. 1. A causal model representation of two causes, C_1 and C_2, differing in the number of effects they produce. *Note:*. C_1 is a common cause of three effects, E_1, E_2, and E_3, while C_1 is a cause of only E_3. The parameters w_i represent the causal strength of the respective causal arrows.

because B causes A ($B \rightarrow A$), or because both are downstream effects of an underlying common cause C ($A \leftarrow C \rightarrow B$), or an even more complex causal network. Causal structure knowledge is fundamental for carrying out successful interventions. For instance, if A and B are only indirectly connected via a common cause C (e.g., both are symptoms of a disease), an intervention changing the status of C is an effective means to change the status of B, whereas an intervention changing A is not (see, e.g., Waldmann & Hagmayer, 2005; see also Rottman & Hastie, 2014, for an overview).

A second dimension is the strength with which causes influence their effects, that is, the degree to which a cause changes the value of an effect. Knowledge about the strength of causal links is important to infer the most probable cause of an observed effect (e.g., the most probable cause of a patient's symptoms) (Meder, Mayrhofer, & Waldmann, 2014; Stephan, Mayrhofer, & Waldmann, 2020; Stephan & Waldmann, 2018), to predict possible future events (e.g., the probability of remission after a certain therapy), or to select the most effective of different alternative interventions to achieve a desired goal (e.g., introducing flexible working time vs. increasing salary to boost job satisfaction).

A prominent class of philosophical theories of causality adopts the so-called dependency view of causality (see, e.g., Beebee, Hitchcock, & Menzies, 2009; Hall, 2004; Paul & Hall, 2013; Waldmann & Mayrhofer, 2016, for overviews). According to the dependency view of causality, a factor C is the cause of another factor E if E's status depends on C. In the case of actual causation, which concerns causal connections between specific events (e.g., "John's drunk driving caused the accident at 5th Avenue this morning"), dependency is typically defined as *counterfactual* dependency of the effect on the cause (e.g., Lewis, 1973; Paul, 2009): An event c constitutes the cause of another event e if it is true that e would not have happened if c had not happened (see also Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021, for a recent psychological theory). In the case of general causation (e.g., "Drunk driving causes car accidents"), dependency is often defined as a *statistical* dependency (e.g., Eells, 1991; Williamson, 2009); causes are assumed to raise or lower the probability of their effects (under appropriate conditions).

One of the most influential theories belonging to the class of dependency theories is causal Bayes net theory. The theory was originally developed by computer scientists (Pearl, 1988, 2000; Spirtes, Glymour, & Scheines, 1993) as a normative theory of how to extract causal structure from statistical data. It has later inspired cognitive scientists in their search for psychological theories of human causal learning and reasoning (Glymour, 2001; Gopnik et al., 2004; Griffiths and Tenenbaum, 2005; Rehder and Burnett, 2005; Sloman, 2005; Waldmann, 1996; see Rottman, 2017 and Rottman and Hastie, 2014, for overviews).

According to causal Bayes net theory, causal structures can be represented as acyclic graphs in which cause and effect variables are nodes connected by arrows pointing from the causes to the effects. An example of a causal Bayes net is shown in Fig. 1. The nodes and arrows in Fig. 1 capture the qualitative causal structure of the depicted variables, while the strengths of the different causal connections are expressed by the parameters w_i attached to them.

Causal strength, in the case of binary variables that can be present or absent or on or off, is defined by causal Bayes net theory as the probability with which a cause variable generates or prevents its effect variable in the hypothetical absence of alternative causes. An important assumption of causal Bayes net theory is that the strength with which a cause generates its effects needs to be inferred based on statistical data about the contingency between the putative cause and effect variables. For example, if we want to infer the strength with which C_1 causes E_3, w_{c1e3}, in the causal network shown in Fig. 1, causal Bayes net theory prescribes we should gather information about the statistical association between C_1 and E_3 (see, e.g., Cheng, 1997; Griffiths & Tenenbaum, 2005; Meder, Mayrhofer, & Waldmann, 2014; Stephan et al. 2020; see also Cheng & Lu, 2017a). In the case of continuous variables (see, e.g., Park, McGillivray, Bye, & Cheng, 2022), a causal Bayes net can be understood as a collection of regression models and causal strength can be measured by regression coefficients (see Heckerman, 1998; Rottman & Hastie, 2016).

Taking causal Bayes net theory as a model for how lay people think about causality, it is predicted that in the absence of data about the statistical relation between a target cause and effect people should stay agnostic about the target cause's strength (Griffiths & Tenenbaum, 2005; Meder et al., 2014; Stephan & Waldmann, 2018), or rely on causal strength priors. The causal strength priors that have been discussed in the literature are postulated to be independent of causal structure (see, e.g., Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008; Mayrhofer & Waldmann, 2016). This independence of causal strength and causal structure directly follows from the "modularity assumption" of causal Bayes net theory. In causal Bayes net theory, the strength of each causal link is assumed to be independent of other causal links within the causal network. This is a consequence of the Markov condition, which assumes that

the state of a variable only depends on its parents and children but is independent of other variables conditional on the parents (see Rottman & Hastie, 2014, for an overview).

In the present article, we do not argue against causal Bayes nets as a normative theory of how causal structure and strength knowledge can be extracted from data. We argue against the idea that the dependency view of causality adopted by classic causal Bayes net theory paints a complete picture of how (lay) people tend to think about causality (see also Sloman & Lagnado, 2015). In particular, we argue that people's representation of causality is richer than is postulated by dependency accounts. As was illustrated by our initial example about causes C_1 and C_2, we here ask whether reasoners' intuition about the strength of causal links is influenced by *causal structure* knowledge. We ask how reasoners' knowledge about the number of effects a cause influences, which can be called its *causal scope* (see also Johnson, Johnston, Toig, & Keil, 2014; Sussman & Oppenheimer, 2020), influences reasoners' causal strength intuitions.

In Fig. 1, causal scope differences for common cause models are illustrated by variables C_1 and C_2. C_1 here represents a common cause with multiple effects, E_1, E_2, and E_3. Compared to C_2 that only causes a single effect E_3, C_1 can thus be called a broad-scope cause. The empirical question in this article is whether people who know that C_2 causes only E_3 while C_1 is known to also cause effects other than E_3 form different intuitions about the strength with which C_1 and C_2 causally influence E_3.

Thus, one way to summarize the work we present here is to say that we propose a psychological theory predicting that reasoners have certain prior expectations about the strength of individual causal links that are influenced by what they know about the causal structure in which these links are embedded. These prior expectations may be incorporated into a dependency model of causal reasoning, but we argue that the notions leading to these prior expectations cannot be derived from a dependency view of causal reasoning alone.

To foreshadow the results of our experiments, we found a clear effect of causal scope on perceived causal strength. Specifically, our studies document that reasoners perceive a scope-dependent *dilution of causal strength*: We found that reasoners tend to expect individual causal links of a cause to be weaker the more effects this cause influences. Also, we discovered that causal strength dilution seems to be a default assumption that reasoners make: we found that the effect not only occurs in concrete scenarios, but also when an abstract, content-free causal structure is presented, either in the form of a causal graph (like in Fig. 1) or verbally without any graphical representation. Our theory also specifies boundary conditions for the strength dilution effect. It predicts that the strength dilution effect occurs only if reasoners represent the instructed variables to be continuous, but not when they represent them as being genuinely binary variables that can only be either present or absent. Our studies confirm this prediction. Also, our theory predicts that the effect depends on assumptions about the nature of the causal process via which a cause influences its effect(s). Strength dilution is predicted to occur when reasoners assume that causes exert a "directed" influence, that is, when they think that the causal process unfolds along effect-specific paths. It does not occur, by contrast, when reasoners assume a "radial unfolding" of a causal process, that is, a process analogous to sound waves that spread out in all directions at once. Our studies also confirm these predictions.

2. Why causal structure might influence causal strength intuitions

Dependency accounts, such as causal Bayes net theory, do not predict influences of causal scope on perceived causal strength. Due to the modularity assumption entailed by the Markov condition, they view each causal relation as independent of other causal relations within a network.

A different framework that better captures the intuitions underlying the predicted dilution effect are dispositional theories of causality. According to dispositional theories, observable statistical regularities between variables are merely secondary; they are regarded as the result of interactions between objects endowed with certain powers or dispositions, which have also been called causal "capacities" (Cartwright, 1989, 1999; Cartwright & Pemberton, 2012; Mumford & Anjum, 2011) (see also Cheng, 1997). Although different capacity theories give different exact definitions of what a power, disposition, or capacity is, the shared basic notion is that entities possess powers that can initiate changes in the world. Powers, dispositions, or capacities may be dormant most of the time but can, when placed in the right context, become manifest and produce the observable events that may then be represented as variables in a causal network (Cartwright, 1989, 1999; Cartwright & Pemberton, 2012). Consider the swallowing of an Aspirin that leads to the feeling of pain relief. While dependency theories would say that taking Aspirin, the cause variable, causes pain relief, the effect variable, *because* the probability of experiencing pain relief is higher when an Aspirin was taken (under the right conditions), dispositional theories view this covariation as arising from the interaction between two entities endowed with capacities, Aspirin and human bodies. The statistical regularity or probabilistic dependency between takings of Aspirin and episodes of pain relief is a result of the repeated operation of these fundamental causal powers.

A different psychological theory of causal learning, which can also be considered a variant of a hybrid theory, that postulates that reasoners believe in (unobservable) causal powers that give rise to observable statistical regularities is Cheng's causal power theory (causal power PC theory in short) (Cheng, 1997; Cheng & Lu, 2017b; Cheng & Novick, 2005; Liljeholm & Cheng, 2007; Lu et al., 2008; Novick & Cheng, 2004; Park et al., 2022). The overall goal of the theory is to reconcile the Humean view that causal connections between events are absent in our sensory input with the Kantian view that we have the a priori idea that causal connections exist in the world. The power PC theory provides a normative computational account showing how the unobservable power of a target cause can be induced based on observable covariations. Thus, the observed causal dependencies are viewed as empirical indicators of an underlying non-observable power, which according to Cheng (1997) is defined as the probability with which a cause event generates or prevents an effect in the hypothetical absence of alternative causes. Since power cannot be directly observed, background assumption about independence and the functional form that underlies the combination of observed and

unobserved causes (e.g., noisy or; linear) etc. have to be made to translate observed covariations into power estimates. Although the theory is inspired by more general accounts of power or capacity (e.g., Cartwright, 1989), its focus is on the probabilistic strength of cause events with respect to their effects. Here the theory contrasts with dispositional theories, which study (correct or incorrect) assumptions about the properties of causal objects (e.g., their chemical composition) that participate in causal relations. One important methodological difference to our approach is that we do not focus on the relationship between covariation and power, but directly study how people conceive of causal dispositions in the absence of learning input. Finally, and maybe most importantly, we study in the present project common cause models. Power PC theory has so far focused on causal models with single effects and single or multiple causes. It can therefore only be speculated what the theory would predict regarding perceived dilution effects.

We will here propose a psychological hybrid theory of causal reasoning that combines causal structure representations with specific notions about causal capacities. Our theory predicts a strength dilution effect. Generally, according to hybrid theories, reasoners' representation of causality is a hybrid of interacting concepts from different (often competing) theoretical frameworks rather than a maximally coherent unitary theory (see, e.g., Waldmann & Mayrhofer, 2016, who argue for a hybrid account of causal reasoning). Our idea here is that reasoners' representation of causal structure resembles causal Bayes nets, whereas how they think of the causal mechanisms within these networks is guided by dispositional notions.

Initial empirical evidence for the psychological validity of hybrid causal representations comes from studies by Mayrhofer and Waldmann (2015), who investigated the interplay between dependency and dispositional notions of causality (see also Waldmann & Mayrhofer, 2016, for an overview). Mayrhofer and Waldmann (2015) in their experiments varied which of the cause and effect variables of a common cause network represented active causal *agents* and which passive causal *patients*. Agents and patients are dispositional concepts from Talmy's (1988) semantic *force dynamics* theory of causation (see also Wolff, 2007; Wolff, Barbey, & Hausknecht, 2010). Essentially, force dynamics describe how different entities such as agents and patients interact with respect to force. Patients are entities with a natural tendency towards inactivity that are acted upon by an agent exerting an *opposing force*. The nature of the result of an agent-patient interaction depends on whether and to what extent the agent's force overcomes the patient's force. Dispositional concepts, such as agents and patients (Talmy, 1988; Wolff, 2007; see also Wolff, 2017, and White, 2017, for overviews), should not influence reasoners' causal inferences if they were solely based on dependency notions of causality. However, Mayrhofer and Waldmann (2015) observed a pronounced effect of their agent-patient manipulation on subjects' causal judgments. They observed stronger Markov violations when the cause variable involved an agent than when the effect variable involved agents (see also Davis & Rehder, 2020; Rehder, 2014; Rottman & Hastie, 2016).

In the present article, we argue that reasoners have a representation of causality in which causal structure knowledge and causal capacity notions interact. This interaction of structure knowledge and dispositional notions explains how structure influences the perceived strength of causal links in a causal network.

2.1. Dispositional mechanism schemas

A key assumption that we make is that people rely on what we call "dispositional mechanism schemas". People rarely know or understand the exact mechanisms via which causes produce their effects. For example, we do not know precisely the mechanisms underlying driving, bike riding, or cooking, but we have partial knowledge about the relevance of gas, pedals, or frying pans. Research on what laypeople typically know about mechanisms has shown that their knowledge is often vague, rudimentary, and contains erroneous assumptions (Chater, 2018; Rozenblit & Keil, 2002; Sloman & Fernbach, 2017). Thus, a key assumption that we make is that reasoners rely on certain default dispositional mechanism schemas when detailed information about causal mechanisms or statistical data about the association between a cause and its effect(s) are not available. Occasionally, domain knowledge about specific mechanisms may be available, especially in expert reasoning. This knowledge may then overwrite the default representation that, according to our theory, should lead to a causal strength dilution effect.

We propose that reasoners' default representation of causality combines causal structure knowledge with a dispositional mechanism scheme according to which most causes possess a (1) certain limited amount of a (2) continuous causal capacity that is (3) distributed across their causal pathways, disposing their effects towards particular values, and (4) leads to a perceived decrease of causal strength proportional to the number of effects a cause with a limited amount of capacity serves.

A schematic illustration of how this hybrid representation of causal structure and dispositional notions leads to a dilution of causal strength is shown in Fig. 2. The left part of the figure shows a (normative) Bayes net representation of a situation in which one cause influences three effects and another cause influences only one effect. The right part shows an example of how we think people tend to represent such situations. The large cylinders represent the two causes. Their amounts of causal capacity is illustrated by the blue substance inside, which is identical in both cases. In the left case, this substance is distributed via three separate tubes to three smaller cylinders, representing the effect variables. In the right case, the substance flows through a single tube to the single small cylinder. The distribution of the substance to three effect variables in the left case results in lower levels in each of the three effect variables compared to the level in the single effect in the right case.

We speculate that this kind of representation may originate from physical analogies that people experience in their daily lives, and then apply to situations in which they lack detailed knowledge about a causal mechanism. The most obvious analogy is the distribution of a liquid through a tube system. Another example often presented in physics classes is the insertion of additional bulbs to a battery-powered electrical circuit where the brightness of all bulbs decreases with the number of bulbs connected to the battery.[1] An analogous psychological example is the case of someone who is focusing on a single or on multiple tasks at once. A typical experience in this situation is that the performance tends to decrease with an increasing number of tasks.

[1] Note, however, that this is true only in the case of a series circuit.

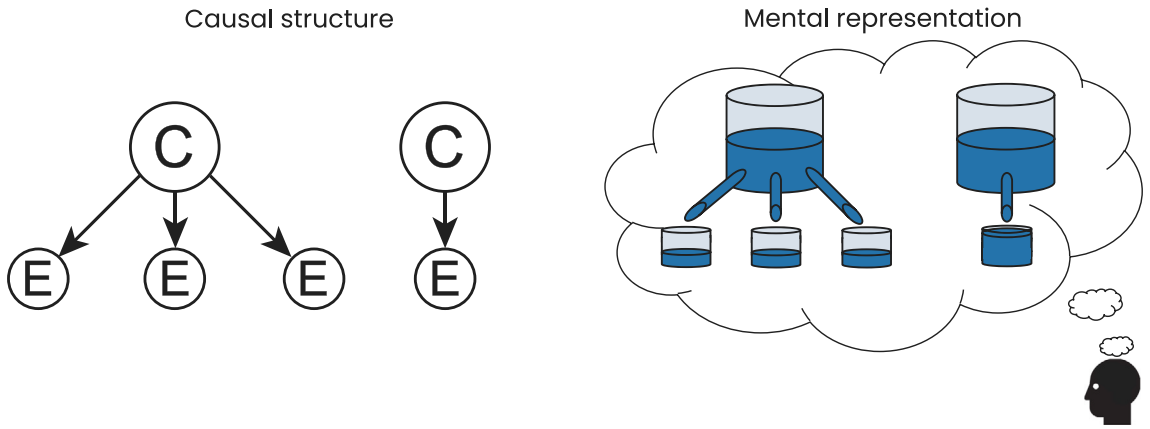

Fig. 2. Illustration of how a combination of causal structure and capacity representations leads to perceived strength dilution.

2.2. Boundary conditions of the strength dilution effect

We mentioned above that Fig. 2 captures how people *tend* to think about the interplay between causal structure and causal strength. We also said that the particular dispositional mechanism scheme that leads to perceived strength dilution is applied by people if they are lacking more detailed information about a given causal system. This leads to the question what the different factors are that we expect to influence the degree to which reasoners perceive causal strength dilution.

We assume that one premise for perceived strength dilution is that the variables of the causal system are perceived to be continuous. We thus expect that people tend to not assume causal strength dilution in the case of *genuinely* binary variables, in which the causal variables just vary between presence and absence. As an example, imagine a genuinely binary switch that can either be off or on. The switch is connected to three genuinely binary LED lights that also can only be off or on; they cannot vary in brightness. We expect that people tend to not assume a dilution effect in such genuinely binary systems: We assume that reasoners do not think of genuinely binary variables as entities that possess a continuous causal capacity, or volume, that is distributable. Note, however, that binary variables can also vary between two levels of a continuous variable (e.g., high vs. low blood pressure). As an example, consider the variable blood pressure and imagine that it is presented as being either "high" or "low". We would call such a binary variable an "artificially dichotomized" variable, because it is actually a continuous variable in the real world (see, for example, Rehder, 2014, for experimental materials that use such dichotomized variables); blood pressure can take on infinitely many values. In the case of such "dichotomized" variables we also would expect a dilution effect because their true continuous nature is often apparent. In fact, when going through materials of causal learning experiments in the literature, most binary variables actually fall into this category of artificially "dichotomized" variables. Cases in which the presence of a cause does not elicit a continuity assumption seem rare.

We think that continuous variables are necessary for the occurrence of perceived strength dilution but not sufficient. People may sometimes also invoke a different partial mechanism schema about how a cause exerts its influence in a given case. Such rudimentary mechanism notions may also moderate the degree to which reasoners perceive causal strength dilution. Our hypothesis is that perceived strength dilution should be strong if a reasoner assumes that a cause spreads its influence in a *directed* way, as in the case of a fluid distributed through different tubes. We may call this notion "channel influence". Another possible propagation mechanism that reasoners may sometimes assume to be at play is what we call "radial influence". Imagine a person lost in the forest shouting for help. In this case, we have the intuition that the sound waves propagate in all directions at once, independent of the number of people who might be listening to the message.[2] Another example are waves caused by a stone thrown into a calm lake. The degree to which the waves will swirl around tiny toy ships arranged around the point of impact is independent of the number of present toy ships. As a further example, consider a radial sprinkler on a lawn. The number of flower pots placed around the sprinkler does not influence the amount of water that each individual plant receives. The mechanism schemas may also sometimes be erroneous. People may think that internet access is radially organized when in fact the likelihood of a breakdown may go up the more users try to access the internet. Abstract graphs representing causal systems in which causes differ with respect to causal scope and either exert their influence radially or via directed channels are shown in Fig. 3.

Apart from the importance of the continuousness of the causal variables and the assumption of a channel transmission, we additionally assume that causal strength dilution requires the notion that a cause's causal capacity is limited. Thus, in cases in which a reasoner believes that a cause possesses sufficient capacity to fully influence a given number of effects, the strength dilution effect should be mitigated.

[2] We thank an anonymous reviewer for this example.

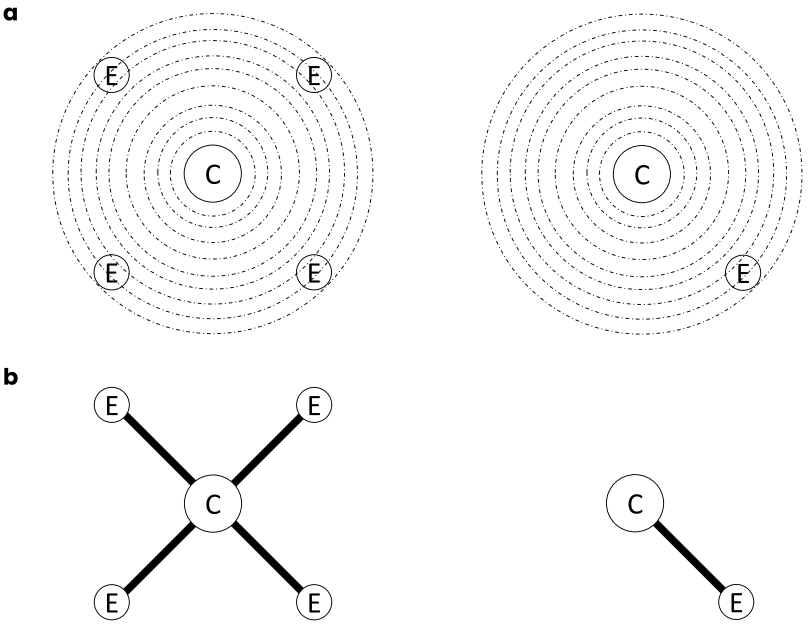

Fig. 3. Illustration of radial and channel transmission. *Note.* (a) illustrates a case in which a cause *C* exerts its influence radially. (b) illustrates a case in which a cause *C* exerts directed influence via channels. The left graphs show causes that influence four effects and the right graphs show causes that influence a single effect.

Finally, the magnitude of perceived strength dilution may depend on the "dispositional roles" that reasoners ascribe to the cause and effect variables of a causal system. Causal processes are in most cases initiated by cause events that play an active role (i.e., are agents). On the effect side, the process terminates and exerts its influence on a target that is often perceived as a passive receiver (i.e., as a patient) that simply takes in what is sent to it. But there are cases in which the effect side may be conceived as playing a more active role (see the reader condition in Mayrhofer & Waldmann, 2015). For example, radios may be equipped with mechanisms that strengthen or boost the signal sent to them. In cases in which effect variables are assumed to be able to actively compensate for a signal loss, perceived strength dilution should also be mitigated.

We next review a number of previous studies that provided initial evidence for an interplay between causal structure knowledge and causal strength intuitions.

3. Previous evidence for causal strength dilution

3.1. Sussman and Oppenheimer (2020)

Sussman and Oppenheimer (2020) report a series of experiments in which subjects were presented with two causes that differed in scope, a common cause of multiple effects, and a single-effect cause influencing one of the effects of the common cause (as in C_1 and C_2 in our opening example; see Fig. 1). Subjects were asked to rate which of the two causes they expect to lead to a faster (Experiment 1) or to a more pronounced change in the target effect (Experiments 2 to 4). Sussman and Oppenheimer presented an account according to which people would tend to perceive single-effect causes (which they called narrow-scope causes) to be stronger than common-causes with multiple effects (which they called broad-scope causes). As is predicted by our theory, Sussman and Oppenheimer speculated that people might think that narrow-scope causes have a higher causal strength because their "power" is distributed to fewer effects: "If we believe that the power of a cause must be divided among its possible effects, a narrow scope cause will lead to a stronger outcome [..]" (p. 2). Contrary to our theory, however, Sussman and Oppenheimer also predicted an influence of outcome valence: According to their account, whether or not strength dilution occurs, depends on whether a cause leads to good or bad effects or, as Sussman and Oppenheimer call them, to boons or banes. They hypothesized that people assume narrow-scope causes to be stronger only in the case of positive effects (boons). In the case of negative effects (banes), by contrast, people are assumed to "believe that a broad scope cause will be more effective. This could be the case if a greater number of distinct outcomes generated by a single cause serves as evidence that the cause is more powerful, which may lead an evaluator to infer that the cause is stronger even on a per-effect basis". (p. 2). The experiments seem to support the prediction.

One problem of Sussman and Oppenheimer's test materials was that it did not preclude participants from applying more specific background knowledge. It could thus be the case that the strength dilution effect they observed in their boon conditions was driven by specific prior knowledge rather than by the more abstract notion that causes possess a certain amount of causal power that they distribute across pathways. For example, one test scenario described two (fictitious) medicines, Saldomorphan and Rydorphan, that

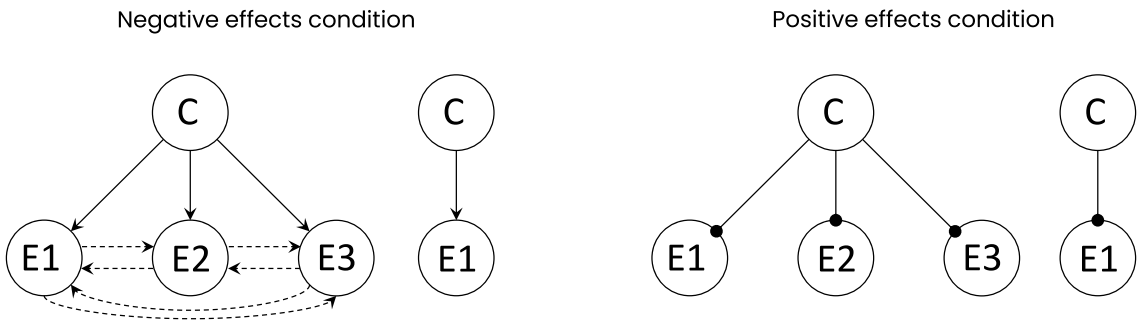

Fig. 4. Causal model illustration of the boon and bane conditions in Sussman and Oppenheimer (2020). *Note.* Links with arrow heads indicate generative causal links and links with round heads indicate preventive causal links. Solid links indicate instructed causal links and dashed links indicate (generative) causal links inferred and reported by participants.

are used to treat eye irritation. One of the medicines was described as a cure of eye irritation, while the other was described as a cure of eye irritation, sneezing, coughing, and a sore throat. One possible assumption that provides an alternative explanation of why subjects tended to say that the medicine that only cures eye irritation is more effective might be that more vs. less specialized medicines can reasonably be expected to differ in their effectiveness. Medicines are usually created with a certain goal in mind. The same problem exists in other scenarios that Sussman and Oppenheimer used (e.g., shampoo, shaving cream, mattress).

Thus, we think that it might not have been the outcome valence per se that led to the reversal of the effect in the negative effects conditions. Rather, we believe that the reversal resulted from additional assumptions of the participants that were prompted by the characteristics of the study materials. A further finding supporting this hypothesis was that subjects in the negative outcome conditions tended to believe that the effects (e.g., painful symptoms caused by a chemical substance) would mutually *amplify each other*. For example, subjects reported that with a shaving cream that causes ingrown hair and skin irritation ingrown hair may also worsen skin irritation. That is, subjects' causal structure representations of the test scenarios with negative effects, but not of those with positive effects, seem to have contained causal arrows between the effects. Moreover, the idea of mutual amplification among the effects in the negative outcome conditions could have arisen because outcome valence (good vs. bad effects) and type of causal process (preventive vs. generative) were confounded in the studies: causes leading to positive outcomes were *preventive* causes that made *a single* existing aversive state disappear, whereas causes leading to negative effects were *generative* causes that produced *initially absent* negative states. A causal model illustration of the different conditions is shown in Fig. 4. In the positive effects condition in which the causes were described to make aversive states disappear, only one aversive state was described to be present (e.g., only $E1$ in Fig. 4) in the test case, which makes mutual causal influence among the effects impossible. Thus, Sussman and Oppenheimer's finding that people expect higher effect intensities for causes leading to negative effects does not imply that people do not perceive a dilution of causal strength in such cases. To test whether reasoners actually do not perceive strength dilution when causes lead to negative effects, outcome valence and type of causal process need to be deconfounded, and any tendency to infer additional causal links between the effects must be kept constant between the two conditions. In a supplementary study (Supplementary Study 1 on the repository site) we deconfounded the two factors using a 2 (causal process: generative vs. preventive) $\times$ 2 (effect valence: positive vs. negative) study design. We also used fictitious materials for which mutual causal connections between effects were implausible. Our results did not replicate the moderation by outcome valence effect. Rather, our results documented a clear causal strength dilution effect in all conditions.

3.2. Studies by White

In a series of studies, White (1997, 1998, 1999, 2000) reported what he called a "dissipation effect" (see White, 2000, for an overview). The effect was found in experiments in which subjects were asked to judge the influences of perturbations in food webs. Food webs represent the feeding relationships between different species of an ecosystem. A perturbation in a food web is defined as a "sudden and significant change to a part of it" (White, 2000, P. 606). An example is the sudden extinction of a specific type of plant that constitutes the food source of a certain herbivore, which in turn is the food source for a certain carnivore. Contrary to what is often observed in real food webs, namely that perturbations tend to interrupt an ecosystem's equilibrium only for a relatively short period of time, White found that lay people tend to think that perturbations have long-lasting effects. Moreover, he found that lay people seem to think that the most drastic changes resulting from a perturbation are experienced by species close to the perturbation point, while species farther away in the web tend to be less affected. In other words, lay people seem to assume that the effects of perturbations dissipate or weaken, while spreading out from the perturbation point across the food web.

The dissipation effect studied by White seems to share many characteristics with the causal strength dilution effect that we observed in our studies. For example, subjects tended to judge the effects of perturbations to be weaker for species that are located at the branches of a food web. Yet, to which extent the effect reflects a more general notion that reasoners have about causality or resulted from specific assumptions about food web dynamics is unclear, as the study only focused on food webs. Some findings indicate scenario-specific assumptions. One such effect is what White called a "terminal effect": subjects tended to judge perturbation

effects to be *stronger* for species located at the endpoints of a food web, that is, for species without a "posterior species" as a food source. According to White, the explanation for this terminal strength effect might be the intuition that "species that have no posterior species lack this source of resistance and are correspondingly more susceptible to the influence of the perturbation". (White, 2000, p. 622).

In sum, previous studies provided possible evidence for the existence of a causal strength dilution effect. However, these studies left open the question whether the findings reflected general assumptions about causality, or rather resulted from domain or scenario-specific background knowledge. We next report a series of studies in which we tested our theory of causal strength dilution in better controlled experiments.

4. Overview of experiments

4.1. Main experiments

In our experiments we used either abstract or artificial scenarios to preclude participants from relying on specific background knowledge. The goal of our first experiments, Experiments 1a, 1b, and 1c, was to establish the existence of a dilution effect and to demonstrate that it results from default assumptions about causality that reasoners make in the absence of additional information. Experiment 1a probed the effect in a scenario about aliens on a foreign planet that forage on different crystals that have physiological effects in the alien body. The only theoretically relevant factor that was manipulated in this study was causal structure (common cause vs. single-effect cause). Experiment 1b then studied perceived causal strength dilution using abstract and content-free materials. Subjects received a causal model illustration similar to the one shown on the left in Fig. 2, together with a brief description of the two presented causal structures. Experiment 1c tested the causal strength dilution effect without showing subjects any graphical causal models. Subjects received a short description of an abstract common cause of three effects and of a single-effect cause, and then rated the causes' strengths. Experiment 2 then provides a test of the postulated boundary conditions: Using abstract materials, we manipulated the different factors that according to our theory should increase or attenuate the degree to which reasoners assume causal strength dilution.

All experiments were conducted as online experiments, and all subjects were recruited via the online platform Prolific (www.prolific.co). The inclusion criteria for all studies were a minimum age of 18 years, English as native language, and an approval rate concerning participation in previous Prolific studies of 90 percent. Also, subjects were asked to participate only via laptop or desktop computer and not via smartphone or tablet because we wanted to minimize the chances of distraction during the study. Subjects received a monetary compensation for their participation that depended on the duration of the study. Subjects who did not respond to all test questions, failed to confirm to have paid attention during the study, or did not pass the comprehension tests were excluded from the data sets prior to any analyses. Participants who took part in our pilot studies or other experiments of this project were also excluded from participation in the experiments.

All experimental materials, demo videos of each study, data, and analyses scripts, including those of pilot and supplementary studies, are provided in an OSF repository (https://osf.io/nrv7p/) whose content can be displayed via a GitHub page at https://simonstephan31.github.io/The-Dilution-of-Causal-Strength/index.html.

4.2. Supplementary and pilot experiments

We also conducted a number of pilot studies to pretest some our materials or to estimate parameters that helped us determine adequate sample sizes for our main experiments. Moreover, we conducted five supplementary studies (Supplementary Studies 1, 2, 3, 4, and 4b on the repository site) in which we further established the robustness of the strength dilution effect. A separate summary of these supplementary studies is made available on the repository site. Commented analysis scripts, data, and further information about these additional studies are also provided on the repository site.

5. Experiment 1a

In Experiment 1a our goal was to establish the existence of a causal strength dilution effect. We used a fictitious scenario about a common cause of three effects and a single-effect cause that causes one of the effects of the common cause, and contrasted subjects' causal strength ratings for the two causes.

5.1. Methods

5.1.1. Participants

One hundred and twenty subjects (M_{age} = 32.94, SD_{age} = 12.81, age range 18 to 69 years, 74 female, 46 male) participated in this online study and provided valid data. With this sample size, a directed t-test analyzing the mean difference between the strength ratings for the single-effect and the three-effects common cause can be expected to detect an effect size of d = 0.45 with eighty percent probability.

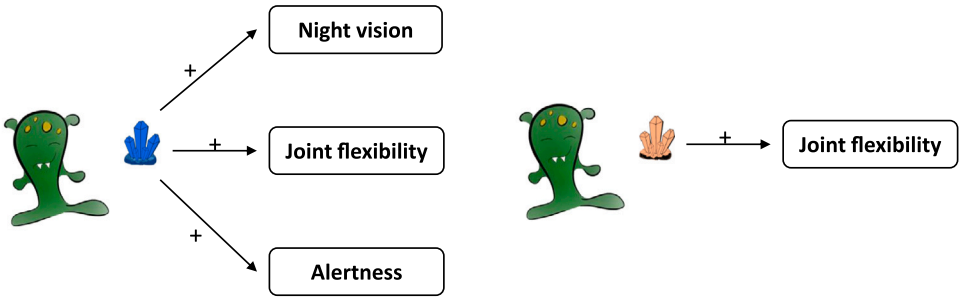

Fig. 5. Example of the scenario illustrations subjects were shown in Experiment 1a.

5.1.2. Design, materials, and procedure

The study had a 2 (target cause: common cause of three effects vs. single-effect cause) × 2 domain (same vs. different physiological domain) between-subjects design. The latter factor was included as a control factor that should not play a role according to our theory.

Subjects were presented with a fictitious scenario about aliens whose diet involves two different types of crystals (blue and red). An example of the illustrations that subjects were presented along with the scenario description is shown in Fig. 5. An example of the scenario description 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 possess the capacity to produce multiple positive effects. Blue crystals can lead to improved night vision, improved joint flexibility, and improved alertness.
>
> Red crystals possess the capacity to produce a single positive effect. Red crystals can lead to improved night vision.

The three effects that were introduced in the condition in which the common cause was described to generate effects belonging to different physiological domains were (1) night vision, (2) joint flexibility, and (3) alertness. In the same effect domain condition all three effects belonged to the musculoskeletal system. The three effects were (1) muscle strength, (2) bone density, and (3) joint flexibility. Whether the blue or the red crystal was introduced as the common cause was counterbalanced between subjects.

The test query either referred to the common cause or to the single-effect cause. Which of the three effects of the common cause was selected as the target cause was counterbalanced between subjects. For example, in the condition in which the target effect was "improved agility" the question read: "We now would like to get to know your intuition about the causal strength with which eating red crystals causes improved agility. 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?". Subjects provided their causal strength ratings on a continuous slider whose endpoints were labeled "not at all" and "maximally".

Subjects then provided demographic information, were given the opportunity to report any technical errors they might have encountered, and then finished the study reading a short debriefing screen.

5.2. Results and discussion

The results are shown in Fig. 6. It can be seen there that we observed a clear causal strength dilution effect: Subjects gave higher causal strength ratings for the single-effect cause than for the common cause. The results also show that the magnitude of the effect was uninfluenced by our effect domain manipulation. The difference plots included in the graphs show the mean difference between subjects' ratings. In the same effect domain condition, the dilution effect was $M_{diff} = M_{sec} - M_{cc} = 0.659 - 0.491 = 0.17$, 95% CI [0.05; 0.29], $d = 0.68$, 95% CI [0.16; 1.20]; planned contrast: $t(116) = 2.80$, $p_{one-sided} = .003$. In the different effect domains condition, the dilution effect was $M_{diff} = M_{sec} - M_{cc} = 0.786 - 0.565 = 0.22$, 95% CI [0.10; 0.34], $d = 1.02$, 95% CI [0.48; 1.56]; planned contrast: $t(116) = 3.68$, $p_{one-sided} < .001$. The estimated difference between the dilution effects in the two domain conditions was $\Delta_{diff} = 0.053$, 95% CI [−0.12; 0.22]. Averaging over the domain factor, the estimated overall size of the dilution effect was $d = 0.82$, 95% CI [0.45, 1.20].

We also conducted a factorial between-subjects ANOVA that included all the counterbalancing factors. In addition to our theoretically relevant effect for the target cause factor (strength rating "single-effect cause" vs. "common cause"), $F(1, 96) = 22.88$, $p < .001$, $\eta^2_{ges} = 0.192$, this analysis yielded two further significant main effects. One was obtained for the domain factor ("effects from same domain" vs. "effects from different domains"), $F(1, 96) = 6.09$, $p = .015$, $\eta^2_{ges} = 0.060$. Subjects gave overall higher ratings when the effects of the common cause belonged to different physiological domains. The other significant main effect was obtained for the target cause color factor ("red" vs. "blue"), $F(1, 96) = 14.89$, $p < .001$, $\eta^2_{ges} = 0.134$. Subjects tended to give overall

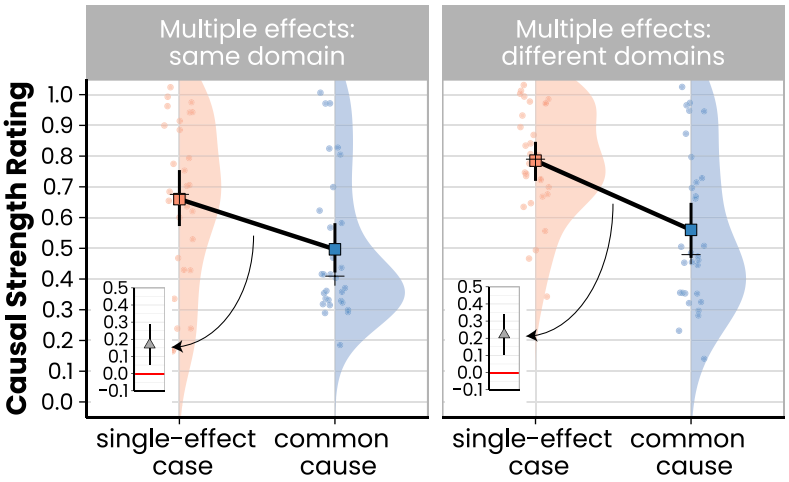

Fig. 6. Subjects' causal strength ratings in Experiment 1a. *Note.* Squares denote means and "+" denote medians. All error bars represent 95% CIs. Jittered dots show subjects' individual ratings, and density plots show their distribution. Difference plots show estimated mean differences.

higher strength ratings when the target cause was red. Importantly, however, there were no significant interaction effects, which suggests that the magnitude of the strength dilution effect is not moderated by any of the other experimental factors.

In sum, Experiment 1a provides evidence for a causal strength dilution effect and suggests that strength dilution is a default notion about the interplay between causal structure and strength that reasoners have. However, a potential problem of Experiment 1a is that the effect may have been triggered by the specific scenario we tested. Although the scenario described unfamiliar situations to avoid an influence of background knowledge, it contained elements of a medical treatment scenario. Subjects might have thought that the two crystals resemble two different drugs developed to target either an isolated symptom or a number of symptoms. Subjects may have reasoned that a more specialized drug developed to target an isolated symptom must be more effective than a less specialized drug with a somewhat broader application. We decided to address this problem in Experiment 1b.

6. Experiment 1b

The goal of Experiment 1b was to make sure that the strength dilution effect observed in Experiment 1a does not depend on specific features of the experimental scenario. Since we assume that the effect results from a default assumption reasoners have about causality in the absence of more specific information, Experiment 1a probed the effect with concrete albeit unfamiliar materials. In the present study we decided to show subjects only abstract, content-free causal graphs similar to those shown in the left part of Fig. 2. If our hypothesis that causal strength dilution is a default assumption is correct, we should observe the effect even with these abstract, content-free materials.

6.1. Methods

6.1.1. Participants

One hundred and twenty subjects (M_{age} = 30.01 years, SD_{age} = 9.48 years, age range 18 to 63 years, 46 female, 31 male, 1 non-binary) recruited via prolific.co participated in this online study and provided valid data. The data of a pilot study with N = 81 for this experiment are provided at the repository site. We found a large dilution effect in this pilot study of d = 1.07, 95% CI [0.74, 1.40], but still decided to increase the estimation precision by recruiting a larger sample for the main study. With a sample of N = 120, a one-sided t-test testing the mean change between subjects' causal strength ratings for the single-effect and the multiple-effects cause can be expected to detect a dilution effect of d = 0.23 with a probability of 80%.

6.1.2. Design, materials, and procedure

We contrasted within subject an abstract common cause with three effects with an abstract single-effect cause. Additional control factors that were counterbalanced between subjects will be introduced below. A print version and a demo video of the study can be accessed at the repository site.

Subjects were shown two abstract causal models displayed next to each other on the same screen along with a brief description. The left–right position of the structures on the screen was counterbalanced between subjects. The cause presented on the left side was labeled C_1 and the cause on the right side C_2. An example of the causal model descriptions that subjects read is:

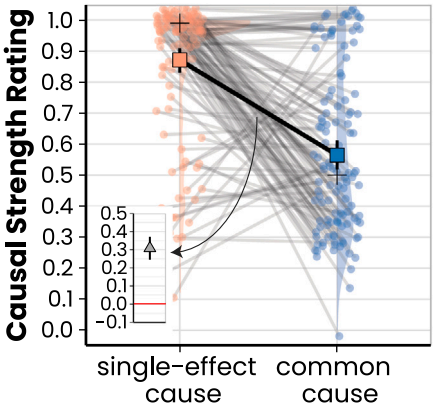

Fig. 7. Subjects' causal strength ratings in Experiment 1b. *Note.* Squares denote means and "+" denote medians. All error bars represent 95% CIs. Jittered dots and lines show subjects' individual ratings, and density plots show their distribution. The difference plot shows the estimated mean change.

> Below you see the abstract diagrams of two different and independent causal structures. Each structure depicts the causal relation between different factors labeled with letters. These abstract factors could be anything, for the present purposes it is only important that you understand how they are causally related, which is indicated by the causal arrows.
>
> As you can see, the two causal structures differ with respect to the number of effects they can influence.
>
> In the first causal structure, cause factor C1 has a causal influence on factor E1, factor E2, and factor E3.
>
> In the second causal structure, cause factor C2 has a causal influence only on factor E1.

Below the causal model information on the same screen, subjects were asked to express their intuition about the strengths with which the two causes influence the target effect. The causal strength test queries read "How strongly does cause factor C1 [C2] causally influence factor E1 [E2, E3]?". The target effect was counterbalanced between subjects. Whether subjects were first asked to rate the strength for the common or the single-effect cause was counterbalanced between subjects. Subjects provided their strength ratings on continuous sliders whose endpoints were labeled: "very weakly" and "very strongly". After subjects had given their strength ratings, we asked them on a separate screen to briefly explain their ratings. Subjects then provided demographic data and finished the study reading a short debriefing screen.

6.2. Results

6.2.1. Subjects' causal strength ratings

Subjects' causal strength ratings are shown in Fig. 7, where it can be seen that subjects tended to report a pronounced causal strength dilution effect also with our abstract and content-free test material: the mean change between subjects' ratings for the single-effect cause and the common cause was $M_{diff} = 0.87 - 0.56 = 0.31$, 95% CI $[0.24, 0.37]$. This corresponds to a large standardized effect of $d = 1.26$, 95% CI $[0.91, 1.60]$. Furthermore, a mixed ANOVA including all between-subjects counterbalancing factors and "type of cause" (single-effect vs. common cause) as within-subject factor yielded only a significant main effect for "type of cause", $F(1, 108) = 92.09$, $p < .001$, $\eta^2_{ges} = .31$.

6.2.2. Subjects' explanations

For both the pilot and the main study we analyzed the explanations that subjects wrote to see to which extent they align with our theory about perceived strength dilution. A supplementary file with the full list of subjects' explanations is provided at the repository site. We indeed found that many descriptions clearly expressed the impression of a dilution of causal strength. For example, one subject wrote "C1 only affected one outcome, and therefore was very strongly causally linked to E1; however C2 affected 3 outcomes and, by intuition I decided that C2 shared its causal strength equally between E1, E2, and E3 resulting in a strength 1/3 of that of C1". Another wrote: "C1 alone with E1 has a very strong effect because all of C1's strength is focused on E1 but with C2, its strength is shared among E1, 2 and 3 so the effect is not so much as C1's influence", and another "C2 had influence over 1 factor whereas C1 had over 3 factors. C2 could use all strength into one but C1 had to influence 3", and another "I would assume that something that affects only one thing (C1 affecting E2), would have a stronger effect than that which influences 3 things at the same time due to it being more concentrated", and yet another "In the first diagram, all of C1's strength was focused on E1. In the second diagram, C2's strength was split in 3, which would mean E1 would only get a 3rd of C2's strength". However, there were also ambiguous explanations (e.g., "C1 directly causes E1, not the case with C2"), and also some explanations that were in line with causal Bayes theory. For example, one subject wrote: "There was not enough information provided to accurately gauge the strength", and another wrote "Both C factors have the ability to influence the occurrence of E1. Have no data that would allow me to conclude that they influence it differently between them"..

6.3. Discussion

In sum, Experiment 1b replicated the causal strength dilution effect with abstract causal models, which corroborates our hypothesis that the assumptions leading to expected causal strength dilution belong to the core assumptions about causality. This hypothesis was further corroborated by subjects' written explanations. A potential problem of Experiment 1b, however, was that the presented abstract causal models might have prompted subjects to think of a physical tube system, as the causal arrows in a graph resemble tubes or channels. We therefore decided to run an additional study in which we omitted any graphical representations of the instructed abstract causal structures.

7. Experiment 1c

Experiment 1c probed the causal strength dilution effect with purely verbal, abstract information to rule out that subjects causal strength ratings in Experiments 1a and 1b were driven by the graphical illustrations of the causal structures.

7.1. Participants

One hundred and twenty subjects (M_{age} = 29.82 years, SD_{age} = 9.19 years, age range 18 to 59 years, 79 female, 39 male, 3 non-binary) recruited via prolific.co participated in this online study and provided valid data. We decided to use the same sample size as in Experiment 1b, N = 120.

7.2. Design, materials, and procedure

We contrasted within subject an abstract common cause with three effects with an abstract single-effect cause. Additional control factors that were counterbalanced between subjects will be introduced below. A demo video and a print version of the study can be accessed at the repository site.

Subjects received the following description of causal structures:

> Please imagine two different and independent causes, called C1 and C2 that differ with respect to the number of effects they can influence. The causes and the effects could be anything, for the present purposes it is only important that you understand how they are causally related, which we describe next.
>
> Cause C1 has a causal influence on effects E1, E2, and E3.
>
> Cause C2 has a causal influence only on effect E1 [E2] [E3].
>
> We would now like to get to know your intuition about the causal strengths with which the two causes C1 and C2 influence effect E3. Please use the sliders below to indicate what you think about their strengths.

Whether C1 or C2 was the common cause, was counterbalanced between subjects. After the description of the causal structures, subjects were asked to provide causal strength ratings for both of them. The target effect mentioned in the test queries was the one shared by the two causes. Whether the target effect was E1, E2, or E3 was counterbalanced between subjects. The test questions read: "How strongly does C1 [C2] causally influence E1 [E2, E3]?". Strength ratings were provided on a continuous slider with the endpoints labeled: "very weakly" and "very strongly". The order of the test queries (single-effect cause first vs. common cause first) was also counterbalanced between subjects. As in the previous study, subjects were also asked to give a brief explanation of their strength ratings because we wanted to see how they represent the causal model in the absence of a graph.

Subjects then provided demographic data, could report any technical errors they might have encountered, and finished the study reading a short debriefing screen.

7.3. Results

7.3.1. Subjects' causal strength ratings

Subjects' causal strength ratings are shown in Fig. 8. As can be seen there, subjects again tended to report estimates consistent with a causal strength dilution effect even without having seen a causal graph. The effect was slightly smaller than in Experiment 1b. The mean change in subjects strength ratings was M_{diff} = 0.70 − 0.54 = 0.16, 95% CI [0.11, 0.21], $t_{contrast}$ = 5.83, p < .001 (two-sided), d = 0.52, 95% CI [0.26, 0.78].

7.3.2. Subjects' explanations

The explanations that subjects gave were similar to those we observed in Experiment 1b, which suggests that the way they think about the interaction of causal structure and strength did not dependend on the presentation of an abstract causal graph. Even in the absence of a graph, subjects still tended to report causal strength dilution. Two examples for the kind of explanations that many subjects wrote in the present study are: "It seems logical to me that C1, which has an influence on multiple things, would therefore have a diluted influence on each in order to cover all three effects. Since C2 only influences one thing, it feels logical that it would therefore have a higher concentration of effect since it doesn't have to share its influence among other things", and "I assumed

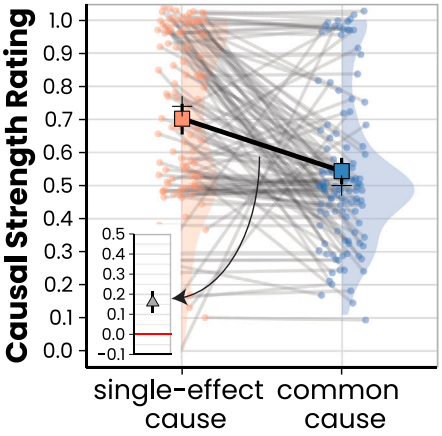

Fig. 8. Subjects' causal strength ratings in Experiment 1c. *Note.* Squares denote means and "+" denote medians. All error bars represent 95% CIs. Jittered dots and lines show subjects' individual ratings, and density plots show their distribution. The difference plot shows the estimated mean change.

that if C2 could only affect E1, the full force of C2 would be on E1. However, C1 can influence multiple things, so I reduced the strength accordingly".

7.4. Discussion

In sum, this experiment demonstrated that perceived causal strength dilution does not require a graphical presentation of the different causal structures. The experiment thus provides further evidence that people's default concept of causality seems to involve the idea that causes possess a certain amount of a continuous causal capacity that they distribute between their effects.[3]

8. Experiment 2

If causal strength dilution is a default assumption that reasoners make in the absence of additional information, a relevant question is whether there are boundary condition for this effect. The goal of Experiment 2 was to address the boundary conditions of the dilution effect by directly testing the factors that we predicted to have an influence on the effect: We manipulated the variable type (binary vs. continuous), the type of causal process (radial vs. channel transmission), and the amount of energy the instructed causes possess (limited vs. unlimited). We expected perceived strength dilution to be particularly strong in the case of continuous variables, and with causes that have a limited amount of energy that they spread via channels. By contrast, we expected no (or at best a weak) strength dilution effect in the case of genuinely binary variables that vary between present and absent. With respect to the type of causal process, we assumed that strength dilution should be mitigated if causes exert their influence radially rather than via channels. Also, causes with unlimited energy should generally elicit weaker perceived strength dilution than causes with a limited amount of energy.

8.1. Methods

8.1.1. Participants

Seven hundred and eighty subjects ($M_{age} = 37.94$, $SD_{age} = 12.73$, age range 18 to 75 years, 378 female, 388 male, 11 non-binary, 3 who preferred not to report their gender) participated in this online study and provided valid data. The sample size decision was guided by the results we obtained in a pilot study with $N = 122$ in which we pretested our materials and obtained first estimates of the effect sizes in the different experimental conditions (a summary of the results of the pilot study can be found at the repository site). Based on the pilot study data, we decided to apply an estimation precision criterion as the stopping rule for data collection: we recruited subjects until the width (upper limit minus lower limit) of the 95% CI of the dilution effect (i.e., the difference between strength ratings for the single-effect cause and the common cause) was < 0.20 in all of the six experimental conditions, and the number of subjects was equal in all conditions. This criterion was reached with $n = 130$ subjects per condition.

[3] We also conducted a supplementary study, Supplementary Study 2 on the repository site, in which we probed and found the strength dilution effect with a probabilistic causal strength query. Also, in a further supplementary study, Supplementary Study 3 on the repository site, we introduced a third cause influencing seven effects. As predicted, we found that causal strength dilution increases with the number of effects served by a cause.

8.1.2. Design

The study had a mixed design. All subjects rated the causal strength of a common cause of four effects and a single-effect cause (within-subject manipulation). The factors that we predicted to influence causal strength dilution were manipulated between subjects, resulting in six different between-subjects conditions. These factors were: (1) the variable types of the variables in the presented causal structures (binary vs. continuous), (2) the type of process via which the causes influence their effect(s) (radial vs. channel transmission), and (3) the energy level of the causes (limited vs. unlimited energy). The energy level factor was manipulated only in the conditions with continuous variables but not in the binary variables conditions. This led to the following six between-subjects conditions:

1. binary variables; radial transmission
2. binary variables; channel transmission
3. continuous variables; radial transmission; unlimited energy
4. continuous variables; radial transmission; limited energy
5. continuous variables; channel transmission; unlimited energy
6. continuous variables; radial transmission; limited energy

8.1.3. Materials and procedure

A demo version of the experiment can be found on the repository site. Subjects first received the description of an abstract causal system in which a cause variable C was described to influence four effect entities E. Along with the description of the variables subjects were shown a picture of variables with C in the middle and the four effects E around it. The figure looked like the left illustrations in Fig. 3, but without the radial representations or the channels. The description subjects were shown was:

> Below you see a schematic, abstract representation of a causal system. This causal system may represent many different kinds of causal relations in the real world, but we do not want to focus on details now. For our purpose, it is important that you understand specific properties of the causal system, which we will describe to you on this and the following screens.
>
> The system consists of a single cause entity "C" that causally influences four effect entities "E".

Subjects then proceeded to a new screen on which we introduced the variable type (continuous vs. binary). In the continuous variables conditions, subjects also received the information about the cause's energy level (limited vs. unlimited). Subjects were shown the same illustration as before. For example, subjects in the conditions with continuous variables and causes with unlimited energy read the following description:

> The cause C and the four effects E represent entities that can have continuously varying levels of activity. Each entity of the system can be more or less intensely active. As C is the cause entity, it influences the activity levels of the four effect entities. Moreover, the cause entity C has an unlimited amount of energy it can emit.

In the continuous variables conditions in which the causes had a limited energy, the last sentence of the description was "Moreover, the cause entity C has a specific amount of energy it can emit"..

In the binary variable conditions, the description that subjects received on the screen was:

> The cause C and the four effects E represent entities that can be in one of two possible states. Each entity of the system can be on or off, but there's nothing in between.

On the next screen, subjects were given a description of the type of causal process. Subjects in the "radial transmission" conditions read:

> The cause entity exerts its influence in a specific way that we call "radial influence". It emits its energy in all directions at once.

Subjects in the "channel transmission" conditions read:

> The cause entity exerts its influence in a specific way that we call "channel influence". It sends its unlimited [specific] energy to the effects through channels.

Along with these descriptions subjects were shown an illustration that looked like the left illustrations shown in Fig. 3.

Subjects then proceeded to a new screen on which we introduced causal structure differences (common cause vs. single-effect cause) by telling subjects that not all four effect entities are always part of the causal system. Subjects read:

> Not all four effect entities are present all the time. For example, it may happen that three effect entities are absent and only one is present. In the left illustration (A) below, you see the same situation as before in which all four effect entities are present. In the right illustration (B), by contrast, you see a situation in which only one effect entity is present and three are absent.

The illustrations subjects were presented on this screen were those shown in Fig. 3 a and b (depending on condition).

Subjects then proceeded to a comprehension test screen on which we probed their understanding of the causal system and the relevant factors that we introduced. Subjects could only proceed to the test phase if they answered all comprehension questions correctly. Subjects who failed the comprehension test returned to the instruction screens and then had a new chance to pass the test. The data of subjects who failed the comprehension test more than four times were coded as invalid; these subjects were excluded from the sample. This was the case for two participants who started the experiment.

On the test screen, subjects were shown the same illustration again as on the previous screen (cf. Fig. 3 a and b). They were asked to imagine that the illustration represents two different situations, one in which all four effect entities are part of the causal system and one in which only one effect entity is part of the system. Subjects then read that in both situations the cause entity would currently be spreading its unlimited [limited] energy radially [via channels] (in the binary variable conditions the reference to the cause's energy level was omitted). For each case, subjects were then asked to rate the causal strength with which the causes influence one of the effects. For example, in the continuous variables conditions, subjects were asked: "According to your intuition, how high is the activity level of the bottom right effect entity?" Ratings were provided on an eleven-point rating scale with the endpoints being labeled "Very low" to "Very high". The order of the two test questions (strength for common cause vs. strength for single-effect cause) was randomized. In the binary variables conditions, we used a probability test query, as a query asking for a continuous state makes no sense in this case. Subjects here were asked: "According to your intuition, what is the probability that the bottom right effect entity is 'on'?" Ratings were provided on an eleven-point rating scale with the endpoints being labeled "0%" and "100%".

After subjects had provided their ratings, they proceeded to a new screen on which we asked them to write short explanations of their previous ratings. Subjects then provided demographic information, were given the chance to report any technical difficulties they may have encountered during the experiment, and then finished the study with a short debriefing screen.

8.2. Results

8.2.1. Subjects' causal strength ratings

An R-Markdown script of the analyses can be accessed at the repository site. Subjects' ratings for the common cause test case and the single-effect cause test case are shown in Fig. 9. Fig. 9a shows their ratings in the binary variables conditions and Fig. 9b shows their ratings in the continuous variables conditions. As can be seen there, overall subjects tended to rate the strength of common causes and single-effect causes similarly in the case of binary variables, whereas their ratings tended to differ in the case of continuous variables. In the case of continuous variables, most effects indicated perceived causal strength dilution: subjects tended to give lower ratings for the common cause than for the single-effect cause. As can also be seen, the magnitude of perceived strength dilution was moderated by our experimental factors in the predicted directions.

To better assess the influence of the manipulated factors, we calculated the mean differences between subjects' strength ratings. These differences along with their 95% CIs are shown in Fig. 10a. As can be seen there, we did not find a strong strength dilution effect when subjects reasoned about binary variables. Also, as predicted, causal strength dilution was most pronounced when subjects were presented with continuous variables and when the cause had a limited amount of energy that it spread via channels.

Fig. 10b summarizes an analysis of the influence of the individual factors. As can be seen there, the different experimental factors had the predicted influence. For binary variables, perceived strength dilution is only very weak, which corroborates our hypothesis that a pronounced strength dilution effect requires continuous variables. For continuous variables, perceived strength dilution is less pronounced if reasoners assume that a cause has an unlimited amount of energy than if they assume that a causes energy is limited. We also found that perceived causal strength dilution is moderated by the type of causal process/ mechanism via which a cause exerts an influence on its effect(s). It is stronger if reasoners believe that a cause exerts directed influences through channels, and weaker if they think that a cause exerts a non-directed, "radial" influence on its effect. Another noteworthy finding is that we never observed a reversal of the effect: In none of our conditions subjects tended to perceive a "strengthening effect", that is, they never had the impression that a common cause has stronger links than a single-effect cause.

Subjects' strength ratings summarized in Fig. 9 and Fig. 10 were further analyzed by fitting two mixed ANOVAs, one for the binary and one for the continuous variables conditions. The models were fitted using the R-package "afex" (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2022). The results of the analysis for the binary variables conditions are summarized in Table 1. The analysis yielded a significant main effect of causal process type (channel vs. radial transmission). Fig. 9a shows that this main effect was obtained because subjects' ratings tended to be higher when the causes generated their effect(s) via channels. The mean differences was $M_{ch} - M_{rad} = 0.07$, 95% CI [0.0132, 0.123]. There was also a significant main effect of cause type (common cause vs. single-effect cause), confirming the existence of a weak strength dilution effect as shown in the top graph of Fig. 10b. There was no significant interaction, however, between causal process and cause type, indicating that the small strength dilution effect does not differ between the two causal process type conditions.

The results of the analysis for the continuous variables conditions are summarized in Table 2. We obtained a significant main effect of causal process type, which resulted from overall higher strength ratings when the causes produced their effect(s) via channels ($M_{ch} - M_{rad} = 0.06$, 95% CI [0.0267, 0.091]). There was also a significant main effect of energy level (unlimited vs. limited): Overall, subjects gave higher strength ratings when the causes were described as having unlimited energy ($M_{unl} - M_{lim} = 0.09$, 95% CI [0.05, 0.12]). A third significant main effect was the effect of cause type (common-cause vs. single-effect cause), indicating that subjects tended to perceive causal strength dilution. The theoretically relevant effects that test our predictions about the influence of causal process type and energy level are the two two-way interaction effects between these factors and the cause-type factor

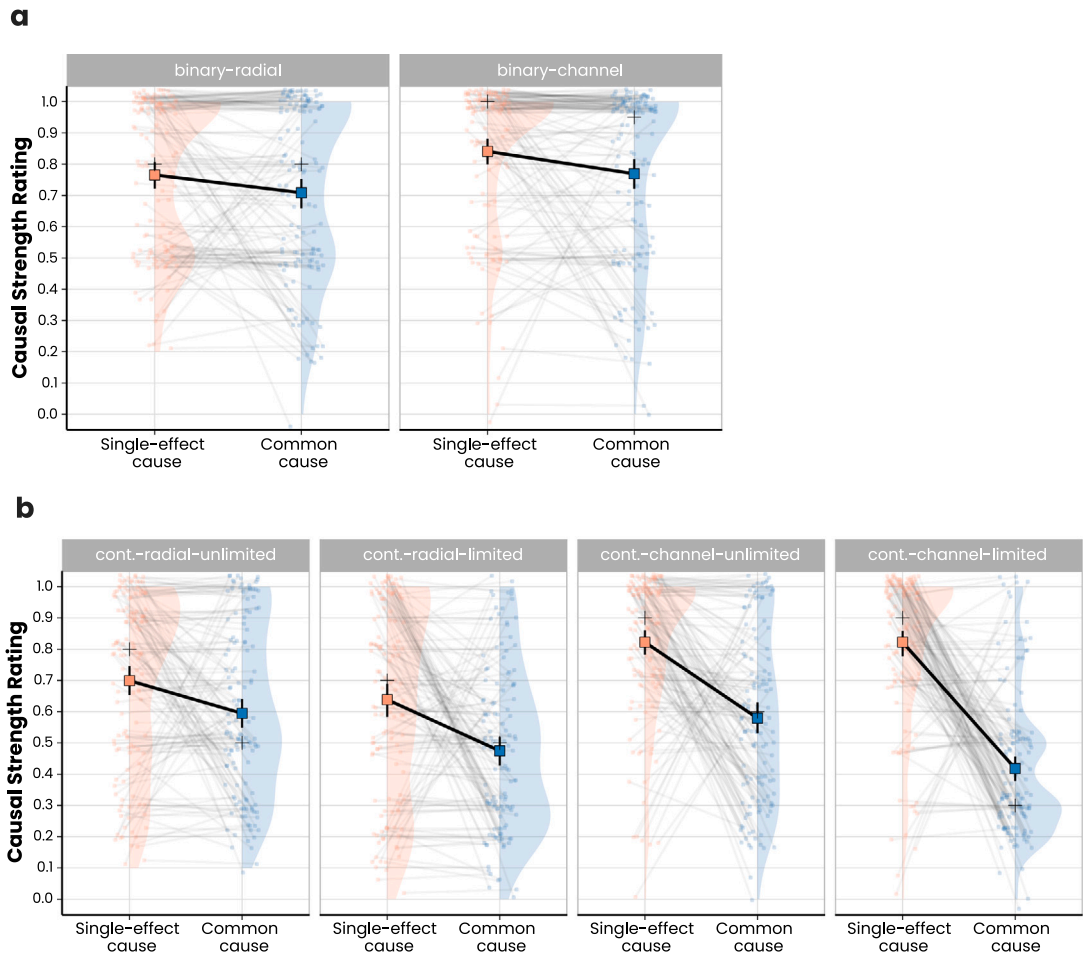

Fig. 9. Subjects' causal strength ratings in Experiment 2. *Note.* Squares denote means and "+" denote medians. All error bars represent 95% CIs. Jittered dots and lines show subjects' individual ratings, and density plots show their distribution.

Table 1
Results of mixed ANOVA for the binary variables conditions of Experiment 2.

Effect	df	MSE	F	ges	p.value
Causal process (between)	1, 258	0.10	5.97*	.016	.015
Cause type (within)	1, 258	0.04	13.51***	.015	<.001
Causal process × Cause type	1, 258	0.04	0.17	<.001	.676

(common cause vs. single-effect cause strength ratings). Table 2 shows that both these two-way interaction effects were significant. These effects reflect the pattern shown in Fig. 10b: the magnitude of perceived strength dilution is moderated by the type of causal process that reasoners assume and by the amount of energy they think a cause possesses. The three-way interaction causal process × energy level × cause type was not significant, indicating that causal process type and energy level moderate the strength dilution effect to a similar degree.

8.2.2. Subjects' explanations

We also analyzed the verbal explanations, subjects provided. The full list of explanations can be found in the "explanation" column of the data set, which can be accessed at the repository site. Our screening of subjects' explanations showed that many explanations tended to reflect the predicted intuitions. As an example for a typical explanation when the causes were described to have unlimited energy, one subject wrote: "As the cause has an infinite amount of energy then surely it makes no difference to the situation regarding how many entities it feeds". Another example is: "Because the amount of energy cause C has, which is infinite, there is no reason to believe that each effect entity gets a finite amount of 1/4 of the energy. That amount could easily be the same as is channeled into entity E. They both could be very high, or they could be different".

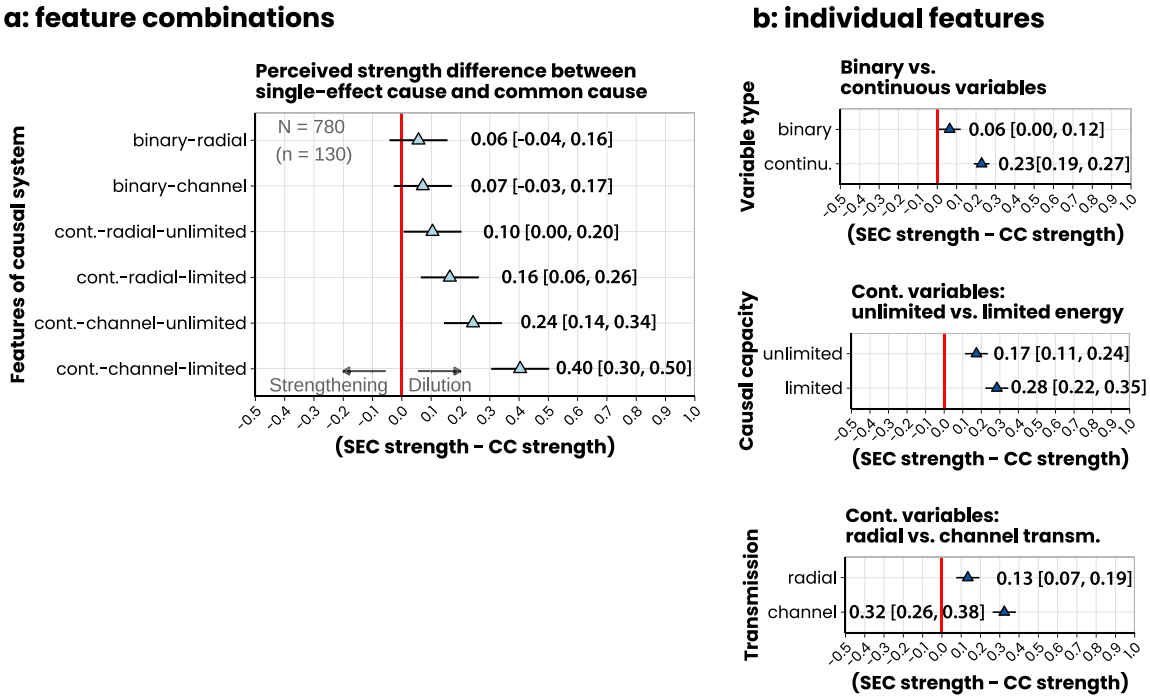

Fig. 10. Causal strength dilution effects in Experiment 2. *Note.* Triangles represent mean differences. All error bars represent 95% CIs.

Table 2
Results of mixed ANOVA for the continuous variables conditions of Experiment 2.

Effect	df	MSE	F	ges	p.value
Causal process (between)	1, 516	7	12.93***	.012	<.001
Energy level (between)	1, 516	7	27.22***	.026	<.001
Causal process × Energy level	1, 516	7	0.09	<.001	.760
Cause type (within)	1, 516	7	196.47***	.160	<.001
Causal process × Cause type	1, 516	7	33.58***	.031	<.001
Energy level × Cause type	1, 516	7	11.35***	.011	<.001
Causal process × Energy level × Cause type	1, 516	7	2.42	.002	.121

Subjects' explanations tended to differ when the causes were described as having a limited amount of energy and generated the effect(s) via channels. The explanations here tended to describe perceived strength dilution. As an example of a typical explanation, one subject wrote: "In the first case, the Cause, having finite amount of energy to give, divides its energy between four Effects, but in the second case the only one Effect gets the full amount of energy that can be given by the Cause". Another wrote: "A specific amount of energy is available so in situation A (All 4 present) I expected there to be lower strength of cause C influencing bottom right effect entity. When only 1 effect entity is present, more energy can be directed and thus higher strength of influence".

Subjects' explanations also reflected sensitivity to our causal process manipulation (directed channel vs. non-directed radial transmission) in the predicted way, although some explanations still expressed strength dilution even in the case of radial transmission. As an example explanation for when the causes were described to exert radial influence, one subject wrote: "Cause C emits the same level of influence radially, the fact that there is only 1 entity present in the 2nd picture has no impact on the amount of energy it receives from Cause c". Another wrote: "From what I understand, the effect entities have no influence on one another, so it's only cause C that has an effect. As such, the influence from cause C remains the same on the bottom right entity regardless of how many other entities are present". Interestingly, some subjects in the radial transmission conditions also mentioned distance to the cause in their explanations. For example, one subject wrote: "They are the same distance away from the entity and that means they'll receive the same energy regardless of how many other entities are there".

Examples for subjects' explanations in the binary variables conditions are: "The effect entities are either on or off and have no relative strengths, therefore regardless of whether the cause influences just one or all four entities, they are all at 100%", "I felt that Cause C influenced the bottom right effect entity in both these situations", and "The study stated that the entity exists in two states either on or off, so if it's on it should cause an effect to the E dots. Regardless of whether it's one of the E dots that are on or all four, if C is on, then E should also turn on. The absence of the other 3 dots of E could be that they simply just aren't around at the present".

Some subjects also provided explanations that signaled a poor understanding of the task. An example is: "I think that if C is off there is a chance that E must be on". Finally, some subjects also provided only pseudo-explanations. An example is: "Intuition, but I think I may answer differently if given the option again".

8.3. Discussion

Experiment 2 successfully demonstrated that the degree of causal strength dilution depends on the factors identified by our theory. As predicted, we found that causal strength dilution is moderated by variable type, causal process intuitions, and assumptions about the amount of "energy" or "capacity" that a cause possesses.[4]

9. General discussion

The results presented in this paper suggest that people's representation of causality is richer than is stipulated by dependency accounts of causal reasoning, such as causal Bayes net theory. We here argued for the view of hybrid causal representations (see also Waldmann & Mayrhofer, 2016). According to this view, reasoners' concept of causality is an amalgam of bits and pieces from different theoretical frameworks rather than a unitary coherent philosophical theory of causality. Since psychological studies tend to test the predictions of theories of causal reasoning that are inspired by unitary philosophical accounts of causality, relatively little is known about how different concepts of causality interact in reasoners' minds. We here proposed and tested the theory that reasoners' causal structure representations are augmented with dispositional intuitions or "mechanism schemas" according to which causes possess a certain amount of causal capacity that they distribute amomg their effects. To test this hypothesis, we investigated whether inferences about causal strength are influenced by causal knowledge about the number of effects a cause influences. Our theory predicts a perceived dilution effect of causal strength with multiple effects when certain boundary conditions hold. We assume that causal strength dilution belongs to the set of default assumptions that reasoners make in the absence of more specific domain knowledge.

We also identified factors that moderate perceived causal strength dilution. First, the occurrence of the effect seems to rely on the variable type; the effect is weak at best in the case of genuinely binary variables that are either present or absent. Moreover, strength dilution is only possible if a continuously varying amount of capacity can be distributed across different effects, which in the case of a cause that can only be on or off is impossible. When the positive value of a binary cause suggests a sufficiently large amount of capacity, though, dilution may be observed (e.g., high vs. low value; see, e.g., Rehder, 2014; Rehder & Burnett, 2005; Rehder & Waldmann, 2017; Rottman & Hastie, 2016, for examples). We also found that the effect is moderated by abstract mechanism schemas: The effect is stronger when subjects assume that the causes initiate causal processes through channels, analogous to a liquid in a water tank distributed through channels. By contrast, in cases in which people have the impression that a cause sends out its "capacity" in all direction at once, perceived causal strength dilution tends to become weaker. Finally, we predicted and found that the magnitude of the effect depends on the amount of capacity a cause is assumed to possess. It is stronger when the amount is finite because then a limited amount of energy needs to be distributed across multiple effects.

Although our experimental results overall corroborate the hypothesis that causal strength dilution results from default assumptions about causality, we also observed interindividual variability. Most subjects reported strength dilution but the magnitude of the effect varied between participants. This is true especially in our initial studies (Experiments 1b and 1c) in which subjects, apart from the causal structure information, received little additional information. However, even there only a minority of subjects did not assume causal strength dilution. Still, among those subjects that assumed causal strength dilution, the magnitude of strength dilution they expected tended to vary. These differences in the assumed magnitude of strength dilution might result from slight differences between subjects in the assumptions about the other factors that according to our theory are relevant (e.g., the assumed type of transmission). This explanation for the observed variability appears plausible also in light of the results of Experiment 2 in which we explicitly manipulated these factors. In the condition in which the continuous causes with a limited capacity exerted their influence via channels, the observed dilution effect was particularly strong and homogeneous.

One interesting finding was that in Experiments 1a - 1c we found the causal strength dilution effect in a context in which causal scope varied "between causes" (i.e., one cause had few effects and another, different, cause had more effects; note, however, that we also manipulated the number of present effect entities within the same cause in Experiment 2). This finding indicates that subjects tended to assume a roughly invariant amount of capacity of different causes which gets transmitted to either a single or to multiple effects. Since we did not provide capacity instructions, it would also have been consistent with the cover stories if subjects had used the number of effects as a diagnostic indicator of the strength of the cause (which a small number of participants might have done). For example, if subjects thought that a common cause of three effects must have three times the capacity of a single-effect cause, then identical strength ratings would result. The stable finding that the cause capacities tended to be assumed to be invariant, whereas the effect strengths were assumed to be variable is an interesting empirical discovery of our project that seems consistent with intuitions about causal asymmetry and the default role of causes as the agents of causal relations. According to dispositional theories of causation, in most cases causes are viewed as active emitters of energy, and effects as passive receivers. However, it should be possible to create cover stories in which the effects either play a more active role (see, e.g., Mayrhofer & Waldmann, 2015), or require a specific threshold amount of energy to be present (e.g., symptoms of a disease; radios). In these cases, subjects may override their default assumption of an invariant cause capacity and use the number of effects as diagnostic indicators of the capacity of the cause: If a cause generates three effects with a threshold, it needs to be stronger than if it only generates one. Note

[4] In two supplementary studies, Supplementary Studies 4a and 4b on the repository site, we additionally tested binary variables in the context of specific (yet fictitious) scenarios. Again, participants did not report causal strength dilution in these scenarios.

that such a finding, based on diagnostic reasoning, would also be consistent with the general mechanism underlying the strength dilution effect.

9.1. Is perceived causal strength dilution a normative effect?

A question that we left unaddressed in the introduction is whether the dilution effect is normative or must be considered a reasoning bias. Expecting a dilution of causal strength seems to be correct at least in cases in which a cause actually represents a limited quantitative source that gets distributed. Whether a dilution of causal strength should generally be expected also in other domains, such as in the physiological or biological scenarios that we tested in some of our studies, seems less clear to us. Our studies suggest that strength dilution is a default assumption that reasoners have, but this default assumption may result from an overgeneralization. It is possible that the default notion of perceived causal strength dilution is primarily based on analogies to physical or related phenomena, which in some cases may lead to erroneous judgments.

One potential line of future investigations could be to test how reluctant people are to overcome the tendency to expect causal strength dilution in situations in which available evidence suggests that causal strength dilution does not occur.

9.2. Implications for studies on causal learning and perspectives for future research

Our findings may have implications for studies on causal learning, especially those that test Bayesian models. For example, different studies that have investigated people's causal strength priors (e.g., Lu et al., 2008; Yeung & Griffiths, 2015) have not explicitly taken the number of effects of a cause into account. Our results suggest that strength priors might also be influenced by the number of effects. It would be interesting to test how easily causal strength priors will be revised in light of statistical data. While our studies suggest that the strength dilution effect has an influence on people's causal strength priors, it is an open question how stable such priors will be in light of contradicting learning data.

Our results also suggest that it is important to be aware of how the instructed variables are internally represented. We showed that depending on whether reasoners believe that variables are naturally continuous or binary influences their inferences. It is important to note that the difference between continuous and binary is not always clear-cut. Binary variables can either be represented as genuinely binary (e.g., present vs. absent), or they can be represented as dichotomized continuous variables (e.g., retirement savings that can be either "high" or "low"; see, e.g., Rehder, 2014; Rehder & Burnett, 2005; Rehder & Waldmann, 2017; Rottman & Hastie, 2016). Our studies suggest that inferences may differ depending on how the variables are internally represented by a reasoner.

10. Conclusion

People's concept of causality is richer than is postulated by dependency theories of causality. A default assumption about causality that reasoners have is that causes possess a certain amount of a continuous causal capacity that they spread across their causal pathways. In line with this view, reasoners tend to infer a structure-dependent dilution of causal strength: Individual causal links are assumed to become weaker, the more causal links are served by the cause.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Sarah Placı, Patricia Cheng, and an anonymous reviewer for helpful comments. Portions of this work have been presented in the *Proceedings of the Cognitive Science Conference*:

Stephan, S., & Waldmann, M. R. (2020). Causal scope and causal strength: The number of potential effects of a cause influences causal strength estimates. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 3426–3432). Austin, TX: Cognitive Science Society.

This work was supported by a research grant (WA 621/24-1) and a Reinhart Koselleck project (WA 621/25-1), both funded by the Deutsche Forschungsgemeinschaft (DFG).

References

Beebee, H., Hitchcock, C., & Menzies, P. (2009). *The Oxford handbook of causation*. Oxford University Press.

Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge, UK: Cambridge University Press.

Cartwright, N., & Pemberton, J. (2012). Aristotelian powers: without them, what would modern science do? In J. Greco, & R. Groff (Eds.), *Powers and capacities in philosophy: the new aristotelianism* (pp. 93–112). New York: Routledge.

Chater, N. (2018). *The mind is flat: the illusion of mental depth and the improvised mind*. UK: Penguin.

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

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

Cheng, P. W., & Lu, H. (2017b). Causal invariance as an essential constraint for creating representation of the world: Generalizing the invariance of causal power. In M. Waldmann (Ed.), *The Oxford Handbook of Causal Reasoning* (pp. 65–84). New York: Oxford University Press.
Cheng, P. W., & Novick, L. R. (2005). Constraints and nonconstraints in causal learning: Reply to White (2005) and to Luhmann and Ahn (2005). *Psychological Review*, *112*(3), 694–706.
Davis, Z. J., & Rehder, B. (2020). A process model of causal reasoning. *Cognitive Science*, *44*(5), Article e12839.
Eells, E. (1991). *Probabilistic causality*. Cambridge, MA: Cambridge University Press.
Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, *128*, 936–975.
Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.
Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*(1), 3–32.
Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–384.
Hall, N. (2004). Two concepts of causation. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 225–276). Cambridge: The MIT Press.
Heckerman, D. (1998). A tutorial on learning with Bayesian networks. In M. I. Jordan (Ed.), *Learning in graphical models* (pp. 301–354). Springer.
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).
Lewis, D. (1973). Causation. *The Journal of Philosophy*, 556–567.
Liljeholm, M., & Cheng, P. W. (2007). When is a cause the "same"? Coherent generalization across contexts. *Psychological Science*, *18*(11), 1014–1021.
Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*, 955–982.
Mayrhofer, R., & Waldmann, M. R. (2015). Agents and causes: Dispositional intuitions as a guide to causal structure. *Cognitive Science*, *39*(1), 65–95.
Mayrhofer, R., & Waldmann, M. R. (2016). Sufficiency and necessity assumptions in causal structure induction. *Cognitive Science*, *40*(8), 2137–2150.
Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*(3), 277–301.
Mumford, S., & Anjum, R. L. (2011). *Getting causes from powers*. Oxford University Press.
Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review*, *111*(2), 455–485.
Park, J., McGillivray, S., Bye, J. K., & Cheng, P. W. (2022). Causal invariance as a tacit aspiration: Analytic knowledge of invariance functions. *Cognitive Psychology*, *132*, Article 101432.
Paul, L. A. (2009). Counterfactual theories. In H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 158–184). New York: Oxford University Press.
Paul, L. A., & Hall, E. J. (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, England: Cambridge University Press.
Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, *72*, 54–107.
Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, *50*, 264–314.
Rehder, B., & Waldmann, M. R. (2017). Failures of explaining away and screening off in described versus experienced causal learning scenarios. *Memory & Cognition*, *45*(2), 245–260.
Rescorla, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Current Research and Theory*, 64–99.
Rottman, B. M. (2017). The acquisition and use of causal structure knowledge. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 85–114). New York: Oxford University Press.
Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*, 109–139.
Rottman, B. M., & Hastie, R. (2016). Do people reason rationally about causally related events? Markov violations, weak inferences, and failures of explaining away. *Cognitive Psychology*, *87*, 88–134.
Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521–562.
Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2022). afex: Analysis of factorial experiments. Retrieved from https://CRAN.R-project.org/package=afex, R package version 1.1-1.
Sloman, S. (2005). *Causal models: How people think about the world and its alternatives*. New York: Oxford University Press.
Sloman, S., & Fernbach, P. (2017). *The knowledge illusion: The myth of individual thought and the power of collective wisdom*. Pan Macmillan.
Sloman, S., & Lagnado, D. (2015). Causality in thought. *Annual Review of Psychology*, *66*, 223–247.
Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. New York, NY: Springer-Verlag.
Stephan, S., Mayrhofer, R., & Waldmann, M. R. (2020). Time and singular causation – A computational model. *Cognitive Science*, *44*(7), Article e12871.
Stephan, S., & Waldmann, M. R. (2018). Preemption in singular causation judgments: A computational model. *Topics in Cognitive Science*, *10*(1), 242–257.
Sussman, A. B., & Oppenheimer, D. M. (2020). The effect of effects on effectiveness: A boon-bane asymmetry. *Cognition*, *199*, Article 104240.
Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*(1), 49–100.
Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. L. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation* (pp. 47–88). New York: Oxford University Press.
Waldmann, M. (Ed.), (2017). *The Oxford handbook of causal reasoning*. New York: Oxford University Press.
Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 216.
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.
White, P. A. (1997). Naive ecology: Causal judgments about a simple ecosystem. *British Journal of Psychology*, *88*(2), 219–233.
White, P. A. (1998). The dissipation effect: A general tendency in casual judgments about complex physical systems. *The American Journal of Psychology*, *111*(3), 379–410.
White, P. A. (1999). The dissipation effect: A naive model of causal interactions in complex physical systems. *The American Journal of Psychology*, *112*(3), 331–364.
White, P. A. (2000). Naive analysis of food web dynamics: A study of causal judgment about complex physical systems. *Cognitive Science*, *24*(4), 605–650.
White, P. A. (2017). Visual impressions of causality. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 245–264). New York: Oxford University Press.
Williamson, J. (2009). Probabilistic theories of causality. In H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 185–212). New York: Oxford University Press.
Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, *136*(1), 82–111.
Wolff, P., Barbey, A. K., & Hausknecht, M. (2010). For want of a nail: How absences cause events. *Journal of Experimental Psychology: General*, *139*(2), 191–221.
Wolff, P., & Thorstad, R. (2017). Force dynamics. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 147–168). New York: Oxford University Press.
Yeung, S., & Griffiths, T. L. (2015). Identifying expectations about the strength of causal relationships. *Cognitive Psychology*, *76*, 1–29.