

Causal Representation and Behavior: The Integration of Mechanism and Covariation

José C. Perales^{1,*}, David R. Shanks² and David Lagnado²

¹*Department of Experimental Psychology Faculty of Psychology University of Granada, Granada, Spain*

²*Division of Psychology and Language Sciences University College London, London, England*

Abstract: Causal knowledge can be based on acquired information about the statistical relationship (covariation) between a cause and effect or on knowledge of the mechanism by which causal power is transmitted between the cause and effect. A key issue is the functional significance of this distinction. In this article, we review recent research in which the influence of covariational evidence on prior beliefs was analyzed. We argue that the way in which covariation influences prior beliefs is independent of whether those beliefs are based on covariation or mechanism information, and that convincing demonstrations of the dissociability of the two types of causal knowledge have not been obtained. We argue that although there are several ways in which causal knowledge can be acquired, that knowledge shares a common representational basis.

Keywords: Causation, contingency, causal reasoning, human learning.

INTRODUCTION

What do we mean when we say that smoking causes cancer or that gas emissions to the atmosphere cause global warming? And how do we substantiate statements such as these? The answer to these important questions about the representation and origin of causal knowledge can incorporate two possible types of information. On the one hand, causation implies a consistent statistical relation between the cause and the effect, and on the other hand, it implies some (usually hidden) mechanism linking the events. To illustrate, consider two ways to argue with an astrologist. You can use current scientific knowledge to demonstrate the implausibility of a mechanism linking the motions of the planets with the ups and downs of our daily lives, or you can question whether their predictions turn out to be true, that is, you can produce statistics demonstrating the absence of a consistent correlation between astrological predictions and events in real life. This article is concerned with the psychological implications of the mechanism/covariation distinction.

Arguments between proponents of mechanism- and covariation-based approaches to causality have been prominent in psychological [1-5] as well as philosophical [6, 7] discussions about causal inference. Fugelsang and Thompson [4] proposed that the two approaches can be integrated in a dual-representation model, in which mechanism-based and covariation-based causal knowledge are coded separately, and have different roles in determining the strength and nature of our causal beliefs. This means that two essentially different types of causal knowledge are assumed to exist. According

to their model (1) causal beliefs are represented both in terms of covariation (the cause and the effect co-occur), and in terms of mechanism (the cause is connected to the effect by a generative or preventative connection), and (2) the two types of knowledge remain separated in the reasoner's representational system. The main prediction of this model is that if reasoners are exposed to a certain degree of covariation between two events, but have prior knowledge that there is no plausible causal mechanism linking them, the covariational information will have a very limited impact. However, if reasoners have knowledge about a possible mechanism, or previous beliefs are exclusively based on covariational information, then new covariational information will be additively integrated with previous beliefs. In other words, pre-stored causal knowledge about causal mechanisms works as a filter or gate that determines the degree to which covariation is incorporated into causal beliefs.

An important corollary of the dual-representation hypothesis is that knowing about the existence of a mechanism is not strictly necessary for deriving causal knowledge from covariation; put differently, a causal belief can be exclusively based on covariational information. Although the belief that there is no plausible mechanism can preclude the integration of new covariational information, the absence of any knowledge about a plausible causal mechanism should not prevent causal knowledge being represented in exclusively covariational terms. If it did prevent it, the dual hypothesis would not be dual at all, as it would assign a primary epistemological role to mechanisms, and only a secondary, confirmatory or disconfirmatory role to covariation.

Our perspective is in contrast to the dual-representation hypothesis: we argue here that it is not necessary to assume the existence of two types of causal representation, but just one. Directly experienced covariation and information about

*Address correspondence to this author at the Departamento de Psicología Experimental y Fisiología del Comportamiento, Universidad de Granada, Campus de Cartuja, s/n, 18071 Granada, Spain; Tel: +34 958 249609; Fax: +34 958 246239; E-mail: jcesar@ugr.es

mechanisms “feed” a common causal representation. Most importantly, on this viewpoint, *the prerequisites for the updating of causal knowledge in the face of new evidence are the same, regardless of whether such evidence is mechanism- or covariation-based*. Obviously, causal knowledge has several origins, but the conditions under which new information alters prior beliefs are independent of its source. In other words, we propose that *the origin of knowledge is irrelevant in belief updating*. This principle is shared by theories of general purpose causal induction, according to which humans are equipped with causal induction mechanisms that are content- or field-independent [2, 3, 8-10]. We are not arguing here that reasoners discard all information about the sources from which their causal knowledge is extracted. Rather, the common representation hypothesis refers to the functional role of prior causal knowledge on updating. In contrasting these two approaches, any evidence showing that mechanism-based and covariation-based causal knowledge exert differential behavioral impacts (for example, on judgments or predictions), other things being equal, favors the dual-representation hypothesis.

Before getting into the details of the evidence reported to confirm (or disconfirm) the dual hypothesis, we will consider some important background issues more carefully. First, it is common knowledge that not all observed covariations are causal, and that reasoners are sensitive to this fact. For example, the dial of a barometer changes before it rains, but the change in the dial does not *cause* the rain. Therefore, causal knowledge is intuitively different from mere covariational knowledge [2, 3, 11]. However, this difference between causal and covariational knowledge does not rule out the possibility that the latter can provide direct evidence for or against the former. Often, the translation of covariation into causation will depend just on the circumstances in which covariation is computed. People not only think that causes systematically precede their effects, but also that effects follow their causes when other potential causes of the same effect are removed; consequently, covariation computed under certain conditions can be interpreted as causal. Analogously, in science, the power of new evidence to change previous theories is greater if such evidence has been obtained in controlled situations. Furthermore, even when a mechanism is considered implausible, this belief can be altered if the reasoner has access to covariational information that is considered reliable and is obtained in controlled conditions. For instance, most of us do not think that the telepathic transmission of thought is plausible, but, eventually, we could be convinced by appropriate controlled experiments that remove alternative causal pathways. Therefore, when we use the term “covariation” we are broadly referring to any information on co-occurrence between a potential cause and an effect directly extracted from interacting with the world, and the circumstances in which that information has been obtained.

Second, it is difficult to define what a mechanism is. For instance, if we ask someone to explain why drunkenness increases the probability of having an accident, the answer could be that drunkenness increases reaction latency to unexpected events. If we ask further why that happens, and our informant has some specialized knowledge, he or she might produce an answer in neurochemical terms. Eventually, however, there will come a point at which there are no more

mediating variables between two factors in the chain. When that point is reached, the only explanation of why the cause and the effect occur together is just that “it happens”. So, we will use the term “mechanism” to mean any information about a chain of causal factors supporting a causal link, obtained either from verbal communication (e.g., scientific readings, word-of-mouth explanations, cover stories in instructions in lab tasks) or any other means (such as, for example, direct causal impressions; see Fugelsang and Roser’s paper in the present volume).

And third, we are not making any particular claim about the essential nature of the representations supporting causal knowledge. Causal knowledge can adopt different formats (see [6, 12, 13]). Evidence supporting the different theoretical options has become highly controversial, and holding a specific position in that controversy goes beyond the aims of the present article. The important point here is that, in psychological terms, a link of the form ‘A causes B’, however it is instantiated, cannot be reduced to a simpler one, but from it, it follows that A and B covary under certain circumstances. This means that the notion of causality is psychologically irreducible, but *implies* covariation. In addition, if a prior causal belief exists, and new covariational evidence is observed, the circumstances under which that new evidence has been obtained will be assessed. Consequently, the power of new evidence to alter previous beliefs will depend exclusively on (1) the amount and reliability of the new evidence, (2) the degree to which that evidence is interpretable as causal; and (3) how firmly established the previous belief was; but, crucially, not on the sort of information (mechanism or covariation) on which that belief is based. Conditions 1 and 3 are self-evident but, probably, condition 2 requires some extra elaboration. As noted above, not all covariational evidence is equally convincing in causal terms (see [14]). The evidential power of covariation is larger when it is obtained while controlling for potential confounds or as a result of an intentional manipulation on the environment. The *degree to which statistical evidence is interpretable as causal* mostly depends on the circumstances in which that evidence has been obtained.

In any case, these three factors can account for apparently contradictory results in previous studies. For example, Oberauer, Weidenfeld, and Fischer [15] have shown that covariational information can readily overpower prior causal beliefs in judgments of the believability of conditionals; Fugelsang and Thompson [4] showed covariation-based causal beliefs to be more easily overpowered by covariational evidence than mechanism-based beliefs; and Müller *et al.* [16] have shown that the sensitivity of mechanism-based causal beliefs to covariation depends on whether those beliefs are from a medical context or from a stock-market context. Similarly, these studies differ in the amount of new evidence presented to the participants, and the circumstances in which this information is presented. In other words, in all of these works, the three factors discussed above were neglected. For example, it would be reasonable to assume that prior causal beliefs are more robust and subjectively reliable when brought to the experiment by the participant that when generated *in situ* by means of instructions. Similarly, quite likely, average Psychology students are more confident in their medical beliefs than in their stock-market beliefs. So, it is firmness of (or confidence in) the previous causal belief,

and not its kind or origin per se, what counts for it to be more or less substantially altered by new evidence. To our knowledge, the only study that has independently assessed strength and reliability (or confidence) for prior causal beliefs and new evidence separately is Perales, Catena, Maldonado and Candido's [17] and, as will be discussed later, their results show that these measures explains the effect of knowledge origin (information on potential mechanisms, covariation) away.

In summary, we propose that all causal knowledge has the same functional implications. We agree with Cheng [2,3] that people assume that there are causes and effects in the world, and that certain covariations reveal causal relations and others do not, in such a way that some covariations can be used to discover causal links. Hence, we accept the functional dissociation between causality and covariation, but not between different *kinds* of causal knowledge¹.

The aim of the present article is to review the available evidence on whether mechanism- and covariation-based causal knowledge are functionally different or equivalent. In other words, our focus is on the functional properties of causal knowledge – the degree to which and how it is modifiable by new evidence – but less so on its nature. Our causal knowledge can be in the form of a set of propositions, mental models, nets, or other formats, but the important claim here it that its updating obeys the same set of rules, regardless of the source from which it is obtained. Hence, we will focus on two sources of evidence: the series of experiments by Perales and colleagues supporting the common representation hypothesis [17, 18] and the series by Fugelsang and colleagues supporting the dual-coding hypothesis [4, 19, 20]. To our knowledge, these are the only studies in which the functional role of causal representations in the updating of causal beliefs has been directly addressed. Among the experiments in these studies, only some of them directly tackle the key issue, namely, *whether the origin of causal knowledge (mechanism-related information or directly experienced covariation) is relevant for updating causal knowledge*. So, although this review devotes much attention to a limited set of data, its aims go beyond a mere procedural critique: to frame appropriately the arguments presented and results obtained so far, and to establish the basic set of premises for future research.

EVIDENCE CONCERNING THE DUAL-REPRESENTATION HYPOTHESIS

In a recent study, Perales, Catena, Maldonado, and Candido [17] directly tested the influence of mechanism- and covariation-based information on causal judgments. Specifically, their participants were first induced to believe that a given cause (a drug, smoking) generated an effect (stomach-ache, cancer). That belief was generated by means of a cover story in which a mechanism linking the cause and the effect was described (mechanism-based information), the level of co-occurrence between the cause and the effect was quantified (covariation information), or both types of information were provided (mixed information). Before providing any

further information, participants were asked about the degree to which they thought the candidate cause actually caused the effect (prior causal strength), and about how confident they were that the available information allowed them to conclude that such a causal link existed (prior confidence). In a second stage of the study, participants were presented with covariational evidence about the contingency (high or low) between the cause and the effect. After receiving such information, they were asked to make a judgment about how reliable that information was (new evidence reliability), and to judge the degree to which the cause generated the effect, taking all the available evidence into account (integrative causal judgment).

The results of this study showed that integrative judgments directly depended on only three factors: prior strength, new covariational evidence, and new evidence reliability. More specifically, new covariational evidence summed with prior belief, but this additive effect was modulated by how reliable the source of the covariation information was. In other words, once reliability of the sources was controlled, the origin of the prior belief (mechanism-based, covariation-based, mixed) was found to be irrelevant in the generation of new causal beliefs, which is in total accordance with the common representation hypothesis we are advocating here.

The conclusions of that study were reinforced by a second one by Catena, Maldonado, Perales, and Candido [18]. In this case, prior beliefs about the relationship between two potential causes and an effect were generated by means of a cover story in which the plausibility/implausibility of a mechanism linking each cause and the effect was described; that is, prior beliefs were based in all cases on mechanism information. Subsequently, either ambiguous or unambiguous information about the covariation between the causes and the effect was provided. Ambiguous covariational information was ineffective in changing prior beliefs. If a candidate cause was previously believed to be effective or ineffective at generating the effect, new ambiguous covariation evidence did not alter the belief. More interestingly, unambiguous information (normatively portraying causal evidence) about the effectiveness of a causal candidate made people judge that candidate as effective, even if the causal link between that candidate and the effect was previously considered implausible. Conversely, unambiguous information about the ineffectiveness of a causal candidate made people judge that candidate as non-causal, even if the causal link between the candidate and the effect was previously considered plausible. In other words, covariational evidence can be as effective as mechanism-based information at generating or updating causal beliefs, so long as that information concerns causal evidence. Whether covariational evidence is interpreted as a vehicle for causal evidence, in turn, depends on whether that information has been obtained according to certain intuitive normative principles (representativeness, avoiding confounds, etc; see [14]).

These studies provide strong support for the idea of common coding of covariational and mechanism information. The major evidence in favour of the contrasting dual-representation hypothesis comes from a study by Fugelsang and Thompson [4]. We therefore begin by describing their Experiment 1. To our knowledge there is no other work in

¹ A detailed description of the philosophical and psychological roots of this position is beyond the aims of this article, but see [2, 14].

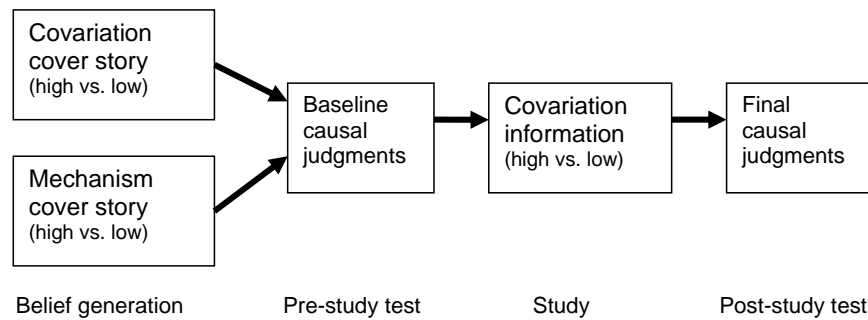


Fig. (1). Summary of the design used in Fugelsang and Thompson's (2003) Experiment 1.

which this or a similar hypothesis has been tested and thus we offer a detailed description and analysis of this research.

FUGELSANG AND THOMPSON (2003, EXPERIMENT 1): SUMMARY OF DESIGN AND PREDICTIONS

The aim of the experiment was to demonstrate that the nature of a prior causal belief (either mechanism-based or covariation-based), and not just its dimensions of strength and reliability, determines how new information about the covariation between a putative cause and an effect is considered and integrated. If a causal belief is based on mere covariation between two events, new evidence regarding that covariation will just be additively integrated with that previous belief. However, if the previous belief is based on the existence of a mechanism binding the two events together, covariational evidence will interact with and add to the previous belief if the mechanism is considered plausible, but not if it is considered implausible. Therefore, the differential effect of covariation on different types of previous beliefs would demonstrate the dissociability of those beliefs.

A summary of the design of Experiment 1 is shown in Fig. (1). In the first stage (belief generation) different cover stories were used to induce causal beliefs based on either covariation or mechanism. These causal beliefs also varied in strength (high vs. low). For instance, in one of the cover stories (corresponding to the covariation-based belief condition), reasoners were told that the presence of a certain chemical had been observed to strongly (or weakly) covary with the slipperiness of a road, but no further information was provided about the possible explanation of that covariation. In the equivalent mechanism-based belief condition, no information was given about covariation, but reasoners were told that the chemical could alter the freezing point of water, and thus make slippery roads more (or less) likely. In all cases, the cover stories involved two single events, namely, a candidate cause and a possible effect.

In a second stage (study), people were exposed to two levels of covariation between the target pair. Covariational information was presented by means of summary contingency tables that displayed the frequencies of the instances in which both the cause and the effect were present (type *a* trials), of those in which the cause was present and the effect was absent (type *b* trials), of those in which the cause was absent and the effect was present (type *c* trials), and, finally, of those in which both the cause and the effect were absent (type *d* trials). Participants were exposed to two different levels of contingency. In the low contingency condition, the

covariation between the cause and the effect, measured as ΔP^2 , was .1, whereas in the high-contingency condition ΔP was .9. Causal strength judgments (about how strongly the participants believed that the effect was attributable to the candidate cause) were collected both before (baseline judgments) and after (final judgments) participants were exposed to the covariational information.

In sum, the three main variables were Belief level (high, low), Belief modality (covariation-based, mechanism-based), and Contingency level (high, low). Belief modality was manipulated between-subjects, whereas the other variables were manipulated within-subjects. The key prediction involved a second-order interaction between these three factors. The effect of covariation on covariation-based beliefs was expected to be additive, that is, it was predicted not to depend on the previous level of belief. As noted above, this is based on the assumption that new information and old beliefs share a common representational code, and therefore the final belief will result from averaging the new evidence with the previous belief level. In contrast the effect of covariation on mechanism-based beliefs was expected to depend on the previous level of belief. In other words, if the mechanism linking the potential cause and the effect was considered plausible, the effect of new covariational evidence was expected to be larger than if such a mechanism was considered implausible.

RESULTS AND REINTERPRETATION

Covariation had a large significant effect in all belief conditions (see Fig. 2). Final judgments were significantly higher in high-covariation conditions than in low-covariation ones independently of the type and level of the previous belief. Most importantly, despite the large number of participants in the experiment ($N=171$) the interaction between the three factors was only close to significance ($p=.055$).

The observed marginally significant interaction was due to the fact that the effect of covariation on high mechanism-based beliefs was slightly larger than in the other three belief conditions; put differently, the effect of prior belief was smaller in the low-contingency mechanism-based condition than in the other three conditions (see Fig. 2). However, a crucial observation is that this difference was already present in the baseline judgments. The difference between the high

² ΔP is a common measure of covariation: $\Delta P = P(E/C) - P(E/-C) = a/(a+b) - c/(c+d)$, where $P(E/C)$ is the probability of the effect given the cause and $P(E/-C)$ is the probability of the effect in the absence of the cause.

and the low mechanism-based beliefs in the low contingency condition was already slightly (though nonsignificantly) smaller than the same difference in the high contingency condition. In other words, the pattern of differences between high and low mechanism-based beliefs in the high and low contingency conditions at baseline is consistent with the pattern of final differences, and therefore it cannot be claimed that the (borderline) interaction between belief level and contingency level arose as a consequence of the studied information. This can be illustrated graphically if we re-plot the critical data shown in Fig. (2) as difference scores between prior beliefs (baseline) and final judgments. This is done in Fig. (3), which shows that study contingency has an almost-perfect additive effect on both mechanism-based and covariation-based beliefs.

Despite Fugelsang and Thompson's interpretation in favour of the dual-representation hypothesis, the foregoing analysis reveals that the pattern of data they report is actually compatible with a different explanation. On the one hand, judgments in this experiment relied more heavily on new covariational evidence than on prior beliefs (as the effect of covariation on final judgments was much larger than the effect of previous beliefs) suggesting that the information presented second was considered more informative than the information received first. In spite of this, there was a significant effect of belief level, which vanished only in the low contingency, mechanism-based belief condition. However, the effect of belief level in that condition was already smaller in baseline judgments (before participants were exposed to new covariational information), an anomaly presumably explained by random sampling. Most importantly, the results failed, despite the large number of participants, to convincingly show the crucial second-order interaction between Belief level, Belief modality, and Contingency level. The main conclusion of the experiment, that covariational evidence interacts differently with mechanism-based and covariation-based causal beliefs, relies on an interaction which is not conventionally significant and which can be explained by baseline differences.

FUGELSANG AND THOMPSON'S (2003) EXPERIMENTS 2 AND 3

Summary of Design and Predictions

In Experiment 1, prior causal beliefs were generated during the initial phase of the experiment. In contrast in Experiment 2 causal beliefs were simply selected, that is, the materials used during the task referred to pairs of events about which participants already had causal knowledge (for example, ice storms and slippery roads, or taking iron supplements and cancer). A group of students was first surveyed about a number of pairs of events. For each pair, they were asked to judge the extent to which they thought the two events were correlated, and the extent to which they were plausibly connected by a causal mechanism in the real world.

Three types of pairs were selected: (1) *Low belief* pairs (e.g., red pots – plants' blooming) were evaluated as weakly correlated, and weakly causally related, (2) *High covariation-low plausibility* pairs (e.g., having chills – fever) were evaluated as strongly correlated, but weakly causally related (although the level of causal plausibility was higher than in the low belief pairs), and (3) *High covariation-high plausibility* pairs (e.g., snow storms – slippery roads) were evaluated as strongly correlated and strongly causally related. A different group of participants was subsequently presented with the three types of pairs, in three different conditions, and exposed to two levels (.1 and .9) of new covariational information involving the three pair types, in order to evaluate the effect of covariation learning on the three types of beliefs. Note that there were not two separate conditions in this experiment for low mechanism-based beliefs and low covariation-based beliefs (instead, there was a single low belief condition).

The strength of the previous belief (measured in the experiment as the *plausibility* of the mechanism) was different for the three conditions. Mean plausibility values were 1.40 for the Low belief condition, 4.56 for the Low plausibility-

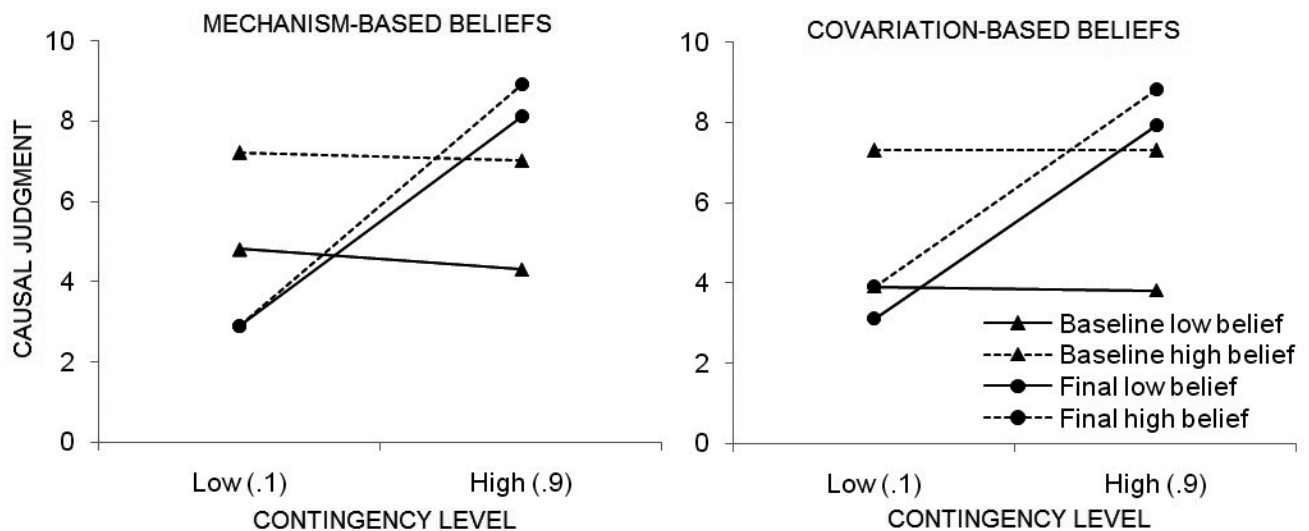


Fig. (2). Mean perceived causal efficacy (causal judgments) for the two belief level conditions and the two covariation (ΔP) levels as a function of belief modality in Experiment 1 (adapted from Fugelsang & Thompson, 2003).

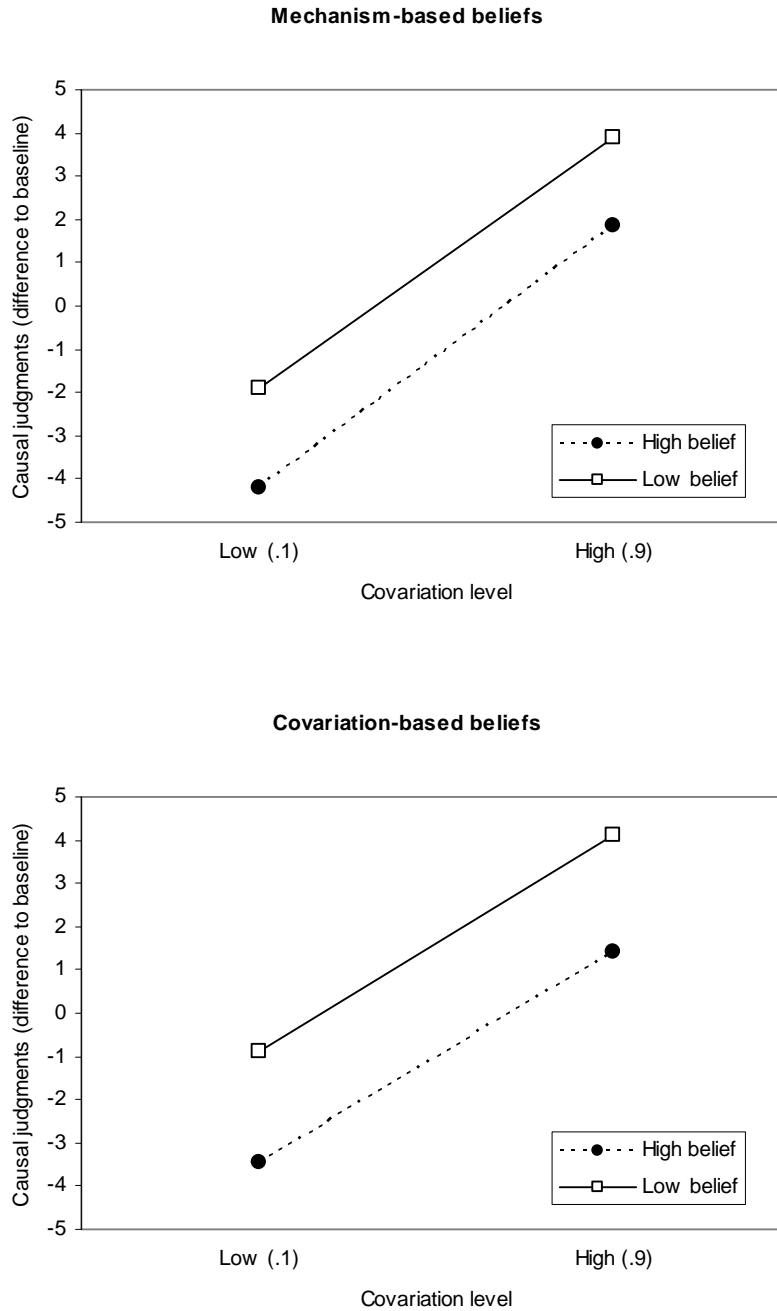


Fig. (3). Results from Fugelsang and Thompson’s (2003) Experiment 1 re-plotted as difference scores between prior judgments (baseline) and final judgments. This way of representing the data shows that learning contingency has an additive effect on both mechanism-based and covariation-based beliefs (that is, the Belief level x Belief Modality x Contingency interaction disappears).

high covariation condition, and 8.62 for the High plausibility-high covariation condition, on a scale ranging from 0 to 10 (as reported by the authors). The results from Experiment 2 (see Fig. 4) showed that the size of the effect of contingency on judgments was directly related to how firmly established the previous belief was (or how plausible the mechanism was, in Fugelsang and Thompson’s terminology). In general terms, the more believable a given causal relation, the larger was the effect of contingency, measured as the difference between causal judgments in the low and high contingency conditions. This is illustrated in Fig. (4) by

the fact that the lines are not parallel but slope slightly (but significantly) more sharply for more believable relations.

REINTERPRETATION

This experiment, however, seems inadequate to test the dual-representation hypothesis, because a crucial condition is missing. In Experiment 1 the interactive effect of contingency and belief level in the mechanism-based belief condition could be compared to the same interactive effect in the covariation-based belief condition, thus making it possible to

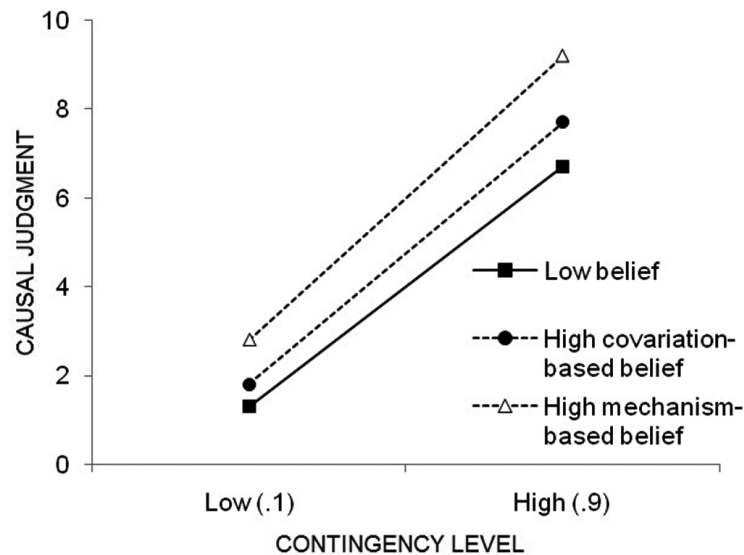


Fig. (4). Mean perceived causal efficacy (causal judgments) for the three belief conditions and the two covariation (ΔP) levels in Experiment 2 (adapted from Fugelsang & Thompson, 2003).

corroborate the existence of the crucial second order interaction. In Experiment 2, however, the low covariation-based belief condition and the low mechanism-based belief conditions were confounded. That is, the low belief in the low belief condition is based both on low covariation and low mechanism plausibility.

The results show that the stronger the belief that a causal link does not exist (or the weaker the belief that there exists a plausible mechanism), the smaller the impact of new covariational information. This effect supports our idea that considering and integrating new covariational evidence partially depends on how firmly established the previous belief is (as shown by [17]), and how reliable the new evidence is (that is, the relative reliability of the sources), but does not demonstrate that mechanism-based and covariation-based causal beliefs are represented separately. The explanation of the fact that this effect appeared in Experiment 2 (and 3), but not in Experiment 1 is unclear. Tentatively, it can be assumed that the difference is due to the fact that prior beliefs were induced in Experiment 1 but selected in Experiment 2; it is not unreasonable to assume that recently induced beliefs, which are based on limited evidence (provided by the experimenter, in this case), are more malleable and easy to modify on the basis of new covariational evidence than real-life beliefs.

In addition, participants in this second experiment were not only asked to make causal judgments, but also covariation judgments. Covariation judgments did not show the Belief type \times Covariation level interaction shown by causal judgments. Importantly, this second result demonstrates the existence of a psychological discrimination between representations of causality and contingency, as we claimed in the introduction. Covariation is not always translated into causality, because causality, in intuitive as well as scientific terms (see [21]) requires covariation to be computed under controlled conditions (as also shown by Catena *et al.* [18]). Again, however, this dissociation does not imply separate representations of covariation-based and mechanism-based causal beliefs.

The results of Experiment 3 were fully consistent with those of Experiment 2. In this case, only high belief and low belief pairs were selected. As expected, the effect of new covariational information was larger in the high belief condition than in the low belief condition. If the two events were thought to be causally unrelated the effect of new evidence was small, whereas if they were thought to be possibly related, the effect of new evidence was much larger. Again, however, this result provides no particular evidence in favor of the separate representation of mechanism- and covariation-based causal beliefs.

In summary, the dual-representation hypothesis does not find support in the analysis of causal judgments in Fugelsang and Thompson's experiments [4]. In Experiment 1, the effect of covariational evidence was comparable in all the belief conditions, and judgments relied more strongly on new covariational evidence than on previously induced beliefs. In Experiments 2 and 3, the influence of covariational information on previous (real) beliefs was evaluated. In both cases, the more plausible a causal link was the larger was the effect of new evidence, which coincides with the results of Perales *et al.* [17]. In these last two experiments, however, the main hypothesis of the dual-representation hypothesis, namely, that the effect of new covariational evidence depends on the nature (not only on the strength) of previous beliefs, was not directly tested.

UNCONSCIOUS BELIEF-BASED AND CONSCIOUS EVIDENCE-BASED PROCESSING

In Fugelsang and Thompson's Experiments 1 and 3 [4], participants were explicitly asked to assess how strongly their previous beliefs or the observed covariational information had determined their causal judgments. In Experiment 1 a significant correlation was found between the objective size of the effect of covariation (measured as the difference between judgments in the high and low contingency conditions for each subject) and the subjective estimates of that effect. However, the subjective estimates of the influence of previous beliefs on causal judgments did not correlate sig-

nificantly with the actual size of that effect (measured as the difference between high and low belief conditions for each subject). This pattern did not depend on belief modality. That is, the effect of mechanism-based beliefs on causal judgment was as difficult to perceive as the effect of covariation-based beliefs. Thus, again, no evidence of dissociation between covariation-based causal beliefs and mechanism-based causal beliefs was found.

In Experiment 3, the actual size of the effect of previous belief was larger than in Experiment 1, which precludes the possible explanation that people failed to detect it just because the effect itself was small. Additionally, information presentation facilitated retrieving the relevant beliefs. However, contrary to what happened in Experiment 1, the correlation between subjective assessments of the impact of previous beliefs and the actual size of that effect was significant.

Fugelsang and Thompson interpreted participants' difficulty in assessing the impact of previous beliefs on judgments as proof that the integration of beliefs with new information is automatic, and largely unconscious, whereas covariation computation is controlled and mainly conscious. However, post-training questions are poor tests of learning awareness (see [22] for a review), and this is especially likely to be true in this situation. By definition, previous beliefs about the connection between two events are stored in long-term memory, and the circumstances in which that information was acquired are probably poorly remembered. In contrast, covariational information is a salient feature of the current experimental task and is superimposed on previous beliefs in the study phase. It is not surprising then that after the study phase recalling recently presented covariational information was easier and more direct than recalling a prior belief. In other words, new information interferes with the recall of previously stored information. That interference effect is strong enough by itself to explain the differences between subjective estimation of the impact of covariation and previous beliefs on causal judgments. The fact that facilitating the retrieval of beliefs also improved participants' ability to measure the impact of such (pre-study) beliefs on final (post-study) beliefs supports the interference hypothesis.

PREVIOUS EVIDENCE

Two of the experiments reviewed here (Experiments 2 & 3 in [4]) show that the effect of contingency manipulations on judgments is weaker when previous causal knowledge dictates that a causal link between the two covarying events is implausible than when that causal link is considered plausible. This effect has also been replicated in [19] and [20], in the same laboratory. It is an intrinsically interesting property of the belief-updating process (see also [23-25]).

The observed pattern of interaction between previous beliefs and contingency information eliminates the possibility that previous causal beliefs work as a yes/no gate that determines whether new covariation information is taken into account in belief updating. Instead, people integrate previous beliefs with new covariational information, and their judgments reflect the effect of both sources of evidence. In other words, old beliefs are revised on the basis of new in-

formation. Which of the two sources of information exerts the stronger influence varies across experiments (and even across participants in a single experiment) and can depend on the order in which the information is presented (See [19]); but, in almost every case, the belief x contingency interaction pattern is strikingly similar.

At some points, Fugelsang and Thompson seem to defend the 'yes/no' gate version of the updating mechanism: "Our first goal was to test the hypothesis that mechanism-based beliefs per se are used to restrict the set of causal candidates about which covariation-based data are considered" ([4], p. 803). However, it is fair to acknowledge that references to this hypothesis are ambiguous, and that their main concern is to defend a weaker version of this claim, according to which beliefs of implausibility downgrade but do not eliminate the effect of contingency [19].

As argued above, the interaction pattern (whereby contingency has a weaker effect on implausible than on plausible beliefs) does not require assuming functionally different representations for different kinds of causal knowledge. In fact, the pattern can be derived from covariation-based theories (see, for example, [25]). The only requirement for accommodating the interaction is the assumption that not only the strength of the previous causal relation, but also its reliability and the degree to which any previous information is thought to provide causal evidence, are taken into account when updating knowledge on the basis of new evidence. Reliable (or plausible, in Fugelsang and Thompson's terminology) beliefs are in general more resistant to modification than unreliable beliefs, and this is entirely rational from a Bayesian perspective. For example, smoking and cancer are only moderately related, but our knowledge of that relation is highly reliable and difficult to revise.

The reason why we have devoted so much attention to Fugelsang and Thompson's study [4] in this article is that it is unique in explicitly formulating and testing the dual-representation hypothesis. In their own words: '*These [their] findings support the conclusion that mechanism-based and covariation-based beliefs are distinct and contribute in different ways to causality judgments... Specifically, when ΔP_c is evaluated in light of covariation-based beliefs, it appears that the two sources of information are simply added together, perhaps because they are derived from the same modality. In contrast, when ΔP_c is evaluated in light of mechanism-based beliefs, it is weighted more heavily for believable than for unbelievable causal candidates. This pattern is consistent with our hypothesis that mechanism-based beliefs are used to restrict the set of candidates for which covariation information is considered*' (p. 806). Yet the only experiment in which that hypothesis is specifically and directly tested is Experiment 1, and, as shown above, the results of that experiment do not support the dual-representation hypothesis.

As discussed above, the results from the studies of Fugelsang, Thompson, and their colleagues are complemented by those of Perales *et al.* [17], in which a similar paradigm was used. In this study, the effect of prior information type (mechanism- or covariation-based) on final integrative causal judgments was completely explained away by the reliabilities of the sources (for prior knowledge and new co-

variational evidence). Still, it is important to acknowledge that the origin of prior information had an effect on the judged reliability of new evidence. Specifically, mechanism based information made people consider new disconfirming covariation evidence as slightly less reliable (e.g., if one knows a mechanism by means of which smoking causes lung cancer, the reliability of any new covariation evidence disconfirming it will be slightly downgraded). In turn, reliability of new evidence had an effect on the power of that evidence to change prior beliefs. However, by means of this indirect connection, the effect of information type on final judgments was negligible, and accounted for less than 7% of the variance. This low percentage could indicate that the power of new evidence to change prior beliefs is not completely captured by reliability measures and that there might be other ways in which information from different sources can differ (e.g., structural certainty, as defined by Griffiths & Tenenbaum [9]).

SUMMARY AND DISCUSSION

According to our reinterpretation, Fugelsang and Thompson's Experiment 1 failed to demonstrate that the impact of new covariational evidence depends on the modality (mechanism- or covariation-based) of such beliefs. Experiments 2 and 3, we have argued, were not completely adequate tests of the dual representation hypothesis, and showed that the influence of new covariational information (when the effect of previous belief is sufficiently strong) is determined by how firmly established was the previous belief that the target events are causally related. Finally, the claim that the integration of previous beliefs and new covariational information is automatic and unconscious, whereas the computation of new evidence is conscious and controlled, was not strongly indicated by the results, because the assessment of awareness was poor [22], and participants were able to track the influence of pre-training beliefs on current post-training beliefs when retrieval conditions were improved (Experiment 3).

On the other hand, the results from Perales and colleagues (and more specifically, [17]) indicate that the origin of causal knowledge is irrelevant for updating, and clearly point to certain parameters affecting the strength of causal beliefs and the evidential power of new evidence as the crucial factors in updating causal representations. The argument is thus clear: because of the commonality of the rules that determine the updating of causal representations, and the fact that different sources of evidence appear to affect causal representations in similar ways, it follows that the mechanism involved in updating causal representations is the same, whatever the source of information.

Considered together, the entire set of results thus supports the notion that all causal knowledge shares a common representational basis and has similar functional implications. Our claim is that such a basis is mainly composed of a network of links of the form 'A causes B'. In accordance with Cheng [3] and with the Bayes' causal nets approach [6, 8, 9, 11], the causal nature of the world needs to be assumed *a priori*, and covariation can under some circumstances signal the presence of causal links. Both covariational information and explanations in the form of mechanisms can be used

to generate new causal statements or to change existing ones. In the latter case, the strength of the prior belief and the evidential power of new information – including in the case of covariational information the conditions under which covariation was computed – will determine how old beliefs and new information interact³.

THE ROLE OF CAUSAL KNOWLEDGE IN HUMAN REASONING AND BEHAVIOR

Causal knowledge penetrates many daily-life reasoning processes and behavior. For example, in the previously cited paper by Oberauer *et al.* ([15]; see also Oaksford and Chater's paper in this volume) the believability of conditionals was indirectly influenced by the presence of a causal link between the two events in the conditional sentence. In a similar fashion, causality can play an important role in heuristic-based decision making (García-Retamero *et al.*, present volume). However, despite its evident importance, the relationship between causal knowledge and everyday behavior is still unclear.

In laboratory experiments, unless specific instructions are provided, reasoners update their causal representations whenever new relevant information is provided; old beliefs are replaced by new ones which in turn serve as the basis for further updating [18]. However, this updating process is very sensitive to (even incidental) task demands and instructions. If asked or primed to do so, people can ignore part of their crystallized causal knowledge and recall directly from their memory the relevant information in order to re-integrate it in a new judgment or choice, in such a way that causality-based behavior can vary as a function of time, context, and even response-type factors [25]. In laboratory experiments all the relevant information is salient and recall from memory direct and easy. In daily life, however, things are rarely linear and sequential; quite often, causal knowledge is not crystallized until it is necessary. This means that the causal representations of the world that underlie our opinions, behaviors, and choices are probably much more flexible than reflected by laboratory experiments.

Much work is still to be done in understanding the integration of causal information in complex real-life environments, and this volume is a perfect example of potentially enriching future research. In this context, the basic assumption of the present review can be considered a starting point: causal links may be the final product of a vast array of information sources, and sometimes those sources are revisited (reasoners, for example, can recall relevant instances from episodic memory), so causal knowledge is continuously subject to updating and change. Nevertheless, the final product – a causal link – can penetrate conditionals, counterfactuals, choices, interventions, and even perception. From our point

³An interesting question (raised by an anonymous reviewer), and a potential line of investigation, is what would happen if a reasoner generates a causal belief from contingency information alone (for example, if a scientist finds an unexpected effect in one of his/her experiments). Most likely, that would trigger attempts to gather more information on such an effect, either evidence from other experiments or explanations in terms of mechanism. Novelty surely motivates curiosity and thus information search. From our point of view, the new information will interact with the prior information following the same rules described above. To our knowledge, however, there are no published works in which prior beliefs are based on covariation only, and new information is presented in terms of mechanisms (that is, when the explanation is given after the evidence).

of view, its centrality crucially depends on its irreducibility. In line with Brunswick [26], the main task of a learner is to capture the causal texture of the environment.

AUTHOR NOTE

J. C. Perales' research is supported by a MICINN grant (reference PSI2009-13133). We would like to thank an anonymous reviewer and the guest editor, Magda Osman, for her role in significantly improving the present paper.

REFERENCES

- [1] Ahn WK, Kalish CW. The role of mechanism beliefs in causal reasoning. In: Keil FC, Wilson RA, Eds. *Explanation and cognition*. Cambridge (MA, US): The MIT Press; 2000; pp. 199-225.
- [2] Cheng PW. Separating causal laws from casual facts: pressing the limits of statistical relevance. In: Medin DL, Eds. *The psychology of learning and motivation. Advances in research and theory* (Vol. 30). San Diego (CA, US): Academic Press; 1993; 215-164.
- [3] Cheng PW. From covariation to causation: a causal power theory. *Psychol Rev* 1997; 104: 367-405.
- [4] Fugelsang J A, Thompson VA. A dual process model of belief and evidence interactions in causal reasoning. *Mem Cogn* 2003; 31: 800-815.
- [5] Shanks DR, Holyoak KJ, Medin DL, Eds. *Causal Learning*. San Diego, CA: Academic Press; 1996. (The psychology of learning and motivation; vol 34).
- [6] Glymour C. Learning causes: psychological explanations of causal explanation. *Mind Mach* 1998; 8: 39-60.
- [7] Harré R, Madden EH. *Causal powers: a theory of natural necessity*. Oxford (UK): Blackwell 1975.
- [8] Gopnik A, Glymour C, Sobel DM, Schulz LE, Kushnir T, Danks, D. A theory of causal learning in children: causal maps and Bayes nets. *Psychol Rev* 2004; 111: 1-31.
- [9] Griffiths TL, Tenenbaum, JB. Structure and strength in causal induction. *Cogn Psychol* 2005; 51: 334-384.
- [10] Pearl J. *Causality: models, reasoning, and inference*. New York (US): Cambridge University Press 2000.
- [11] Lagnado DA, Waldmann MR, Hagmayer Y, Sloman SA. Beyond covariation: cues to causal structure. In: Gopnik A, Schultz L, Eds. *Causal learning: psychology, philosophy, and computation*. New York (US): Oxford University Press 2007; pp. 154-72.
- [12] Goldvarg E, Johnson-Laird PN. Naïve causality: a mental model theory of causal meaning and reasoning. *Cogn Sci* 2001; 25, 565-610.
- [13] Mitchell CJ, De Houwer J, Lovibond PF. The propositional nature of human associative learning. *Behav Brain Sci* 2009; 32: 183-98.
- [14] Perales JC, Catena A. Human causal induction: a glimpse at the whole picture. *Eur J Cogn Psychol* 2006; 18: 277-320.
- [15] Oberauer K, Weidenfeld, A, Fischer, K. What makes us believe a conditional? The roles of covariation and causality. *Think Reas* 2007; 13: 340-69.
- [16] Müller SM, García-Retamero R, Perales JC, Jara E, Maldonado A. Procesos de selección de información en toma de decisiones: la influencia del conocimiento causal / Information selection in decision making: the influence of causal knowledge. Unpublished communication in the XXI Meeting of the Spanish Society for Comparative Psychology 2009; Salamanca, Spain.
- [17] Perales JC, Catena A, Maldonado A, Cándido A. The role of mechanism and covariation information in causal belief updating. *Cognition* 2007; 105: 704-14.
- [18] Catena A, Maldonado A, Perales JC, Cándido A. Interaction between previous beliefs and cue predictive value in covariation-based causal induction. *Acta Psychol* 2008; 128: 339-49.
- [19] Fugelsang JA, Thompson VA. Strategy selection in causal reasoning: When beliefs and covariation collide. *Can J Exp Psychol* 2000; 45: 15-32.
- [20] Fugelsang JA, Stein CB, Green AE, Dunbar KN. Theory and data interaction of the scientific mind: evidence from the molecular and the cognitive laboratory. *Can J Exp Psychol* 2004; 58: 86-95.
- [21] Spellman BA. Conditionalizing causality. In: Shanks DR, Holyoak KJ, Medin DL, Eds. *Causal Learning*. San Diego (CA, US): Academic Press; 1996; 167-206. (The psychology of learning and motivation; vol 34).
- [22] Lovibond PF, Shanks DR. The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *J Exp Psychol Anim Behav Process* 2002; 28: 3-26.
- [23] Catena A, Maldonado A, Megías JL, Frese B. The frequency of judgment effect: belief revision and serial processing of causal information. *Q J Exp Psychol B* 2002; 55B: 267-81.
- [24] Fugelsang JA, Thompson VA. Belief-based and covariation-based cues affect causal discounting. *Can J Exp Psychol* 2001; 55: 70-6.
- [25] Lien Y, Cheng PW. Distinguishing genuine from spurious causes: a coherence hypothesis. *Cogn Psychol* 2000; 40: 87-137.
- [26] Brunswick E. The organism and the causal texture of the environment. *Psychol Rev* 1935; 42: 43-77.

Received: October 02, 2009

Revised: January 07, 2010

Accepted: January 08, 2010

© Perales *et al.*; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.