PSICOLÓGICA

# Recent advances in the study of the illusion of causality: theory, methods, and practical implications


## María Manuela Moreno-Fernández [1], Fernando Blanco [2], Helena Matute [3]

[1]Department of Developmental and Educational Psychology, University of Granada (Spain)
[2]Department of Social Psychology, University of Granada (Spain)
[3]Department of Psychology, Deusto University, Bilbao (Spain)



## Abstract

Learning causal relations provides the knowledge that allows us to make accurate predictions and to make more effective interventions. Some of these predictions may have a high value for survival, providing us with a body of knowledge that maximizes context adaptation. This is why researchers have tried to understand how people learn about the causal structures in their environment. In this article, we outline some of the most recent advances in the understanding of causal learning, and specifically of the causal learning biases that often affect our decision-making.


*Keywords:* Causal learning, causal illusion, cognitive bias.

## Illusions of causality and contingency learning

Inferring causality is not an easy task as it is not directly observable. Instead, we need to infer it indirectly by using some cues or principles, the most relevant ones being priority, contiguity, and contingency (Hume, 1748). Priority refers to the fact that causes proceed their consequences, while contiguity implies that causes and consequences are often close in time and space. Lastly, contingency refers to cause-effect covariation, that is, causes typically covary with their effects in such a way that the presence of a cause changes the probability of its alleged effect. Thus, in some cases, the presence of the cause increases the probability of the effect (generative causes) and, in other cases, the cause reduces the probability of finding an effect (preventive cause).

Let us illustrate the contingency principle with an example. Imagine that you are trying to find out whether a new drug is able to reduce the symptoms of a disease (e.g., headaches). Therefore, you arrange a set of testing trials to study this relationship. For the sake of simplicity, imagine that any other relevant variables such as presence of other diseases, additional treatments or lifestyle changes are also controlled, and that the only factor varying from one test to another is using or not using the target drug. On these circumstances, there are four possible situations (see Figure 1, panel A): You use the

*Psicológica (2023)*

**DOI**
10.20350/digitalCSIC/15705

**Corresponding author**
Helena Matute
matute@deusto.es

**Edited by**
Miguel Vadillo

**Reviewed by**
José Andrés Alcalá

drug and feel better (type *a* event), you use the drug and do not feel better (type *b* event), you do not use the drug and feel better (type *c* event), or you neither use the drug nor feel better (type *d* event).

A - Contingency table

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | a | b |
| No drug (No cause) | c | d |

B - Null contingency

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | 3 | 3 |
| No drug (No cause) | 3 | 3 |

C - Negative contingency

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | 0 | 6 |
| No drug (No cause) | 6 | 0 |

D - Positive contingency

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | 6 | 0 |
| No drug (No cause) | 0 | 6 |

*Figure 1. Some examples of contingency Tables. A. Contingency table containing the four types of events that can be found when combining one potential cause (drug) and one outcome (healing): a, b, c, and d. B. Example of cell frequencies showing a null contingency. C. Example of cell frequencies showing a negative contingency. D. Example of cell frequencies showing a positive contingency.*

Note that all four types of events contribute to an accurate estimation of the covariation between drug (the potential cause, C) and healing (the alleged outcome, O): if the drug is actually effective against the disease, we will expect that the probability of feeling better after using the drug will be higher than the probability of feeling better after not using the drug. This difference may be computed following Equation 1 (Δp index; Allan, 1980):

$$\Delta p = P(O|C) - P(O|{\sim}C) = [a/(a + b)] - [c/(c + d)] \quad (1)$$

When the contingency index Δp equals zero (null contingency), the probability of feeling better after using the drug equals the probability of feeling better after not using the drug, that is, using the drug does not make any difference. When the difference between the two probabilities is lower than zero (negative contingency), the probability of feeling better after using the drug is lower than the probability of feeling better after not using the drug, that is, using the drug prevents the healing, so it is actually counterproductive (the lower the value of Δp index, the stronger the effect). Finally, when the difference is higher than zero (positive contingency), the probability of feeling better if using the drug is higher than the probability of feeling better if not using the drug, that is, using the drug contributes to health improvement (the higher the value of Δp index, the stronger the contribution). Figure 1 (panels B, C and D) shows an example of null, negative, and positive contingencies respectively.

People are quite good at estimating these covariations that can be used to inform causal learning (Shanks & Dickinson, 1988; Wasserman, 1990), but as we will detail next, when certain conditions are met, biases in covariation estimation may arise (Blanco & Matute, 2015; Chow et al., 2019; Double et al., 2020; Hannah & Beneteau, 2009; Moreno-Fernández et al., 2017).

Following our previous example, let us illustrate what a bias is in this context and which conditions may contribute to its appearance. Imagine that you conduct a test of the drug effectiveness on a sample of 100 patients who suffer from the disease, and find the results depicted in Figure 2, panel A.

A

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | 40 | 10 |
| No drug (No cause) | 40 | 10 |

B

| | Remission (Outcome) | No remission (No outcome) |
|---|---|---|
| Drug (Cause) | 60 | 15 |
| No drug (No cause) | 20 | 5 |

*Figure 1. Two examples of cell frequencies in null contingencies. A. Outcome-density effect (most trials show an outcome occurrence). B. Cause- and Outcome-density effects combined (most trials show a co-occurrence of both cause and outcome).*

Following Equation 1, in this example Δp equals zero, which means that we have no reason to believe that the medicine affects the headache remission rate: it holds at 0.8 regardless of whether the medicine has been used or not. However, this is an example of a situation in which people usually overestimate the relation between the medicine and the recovery. That is, when observing a pattern like this, people usually believe that the medicine is working although the data do not support this conclusion.

The critical feature to bear in mind is that, in this particular example, the expected outcome (being healed) occurs at a high rate. From the table in Figure 2 we can compute the probability of the outcome occurrence, p(O), which in this case equals 0.8, a high value. We know from previous research that a high outcome occurrence probability biases the judgments. Moreover, this type of situation is not atypical: acute headaches, for example, are frequently episodic, lasting for some hours before remitting. Therefore, remissions are frequent. Likewise, the data in this example show a high rate of remissions that could be spontaneous, that is, remissions not attributable to the medicine. Of course, it is possible that something else is causing those remissions, but the relevant part here is that the data clearly show that remissions do not depend on the use of the medicine (Blanco & Matute, 2020).

It has been repeatedly reported that in situations where the outcome (e.g., being healed) appears at a high rate, i.e., p(O) is high, as in the previous example, people tend to overestimate the strength of the causal relation. This effect is known as outcome-density effect (Alloy & Abramson, 1979; Buehner et al., 2003; White, 2004). Additionally, increasing the probability of the cause, p(C), makes a similar effect (cue-density effect; Barberia et al., 2020; Matute et al., 2011; Vadillo et al., 2011). Overestimations are typically stronger when both p(C) and p(O) are high (as shown in Figure 2, panel B; Blanco et al., 2013).

Cue and outcome-density effects are two examples of how people often overestimate causality despite data showing a null contingency that points to an absence of causality. This overestimation is often called illusion of causality or causal illusion, and it refers, in general, to any situation in which people believe that a causal relationship exists in the absence of contingency data that supports such conclusion (i.e., typically, a null contingency).

As we will explain below, these illusions may have a wide range of consequences, some of them could be considered as an advantage, while others may pose a risk (Blanco, 2017). For example, it has been argued that this type of illusions can be positive for self-esteem, as they make us believe, even if mistakenly, that important outcomes are not completely subject to chance. On the other hand, there are other occasions where causal illusions may underlie wrong beliefs that support risky behaviors. A common example is using a bogus treatment with dangerous side effects instead of scientifically proven medical treatments.

Given the practical implications of causal illusions on real life and the consequences that may derive from them, some researchers have tried to describe the mechanisms underlying these illusions: Why do people overestimate causal relations? On a later section we will revise some of the most relevant perspectives in the field and the advances made on this topic in recent years. Nevertheless, the first step to build a comprehensive theory is to have a comprehensive set of data, and to get those data we need to develop methods that allow us to get thorough and high-quality information about the conditions and factors that contribute to the development of causal illusions.

## Measuring causal illusions. Methodological Advances

There are two main methodological procedures traditionally used to assess causal learning from contingency information. These are the Passive and the Active procedures. In both cases, participants are exposed to a series of trials in which a cue may or may not occur, and a given outcome may or may not follow, thus showing a sequence of the four trial types. Then, it is their task to infer whether there is a causal relationship between the two events. The main difference between the two procedures is whether the participant's behavior or an external event plays the role of the cue (potential cause). Below, we describe these two main procedures, along with some of the methodological variations that have been proposed in the literature, and some of the effects that choosing one or the other method may have on the results of the experiment.

### Passive procedure

In the passive procedure, the potential cause is an external event. For instance, in a popular methodological procedure, participants are shown a series of medical records of fictitious patients, one per trial. The fictitious patients may or may not have taken a given medical treatment (potential cause), and then they may report feeling better (outcome) or not. That is, a pair of external events that do not depend on the participant's behavior, such as watching records of patients taking a pill and reporting feeling better, can co-occur or not during a series of trials. The task of the participants is to infer to what degree the relationship between those two external events is causal. The cover stories may also differ from the medical story we just described. They may use, for example, meals that cause allergic reactions (Aitken et al., 2000; Hannah et al., 2009), fertilizers that cause plants' growing (Katagiri et al., 2007; Moreno-Fernández et al., 2017), political parties and economical outcomes (Blanco et al., 2018), chemical substances and alien mutations (Blanco & Matute, 2019), teaching strategies and academic achievement (Double et al., 2020), and many other instances of potential causes and outcomes that may be present or absent in each trial (see Matute et al., 2015, 2019, for reviews).

As we described earlier, there are four possible combinations of the presence and absence of the potential cause and outcome in each trial, which result in four different types of trials. Thus, researchers can control the objective contingency between the two events by manipulating the number of trials of each type.

The null contingency paradigm is the standard experimental procedure for investigating the causal illusion because the goal is to study the factors that affect the development of a causal illusion when the actual causal relationship is zero. Therefore, when asked at the end of the sequence of trials about the degree of causality between the two events in these experiments, participants should provide a causal judgment of zero or close to zero. However, they often give a higher value, overestimating the causal relationship between the two events. Thus, they show an illusion, or bias, of causality (see Matute et al., 2015, 2019, for reviews).

## Active procedure

The other standard procedure is the active procedure. Like in the passive procedure, when the research purpose is to study the illusion of causality, the contingency between the potential cause and the outcome will be set to be zero, that is, the outcome is programmed to occur with the same probability in the presence and in the absence of the potential cause. In this procedure, however, the potential cause is the behavior of the participant. Thus, even though the task of the participant is still to infer whether there is a causal relationship between the potential cause and the outcome, in the active procedure the participant is the one who actively decides in which trials will the potential cause be present or absent. As we will see below, this is a very important difference with respect to the passive procedure.

This procedure has been used in studies of causal judgments since the late seventies by, for example, Allan & Jenkins (1983), Alloy & Abramson (1979), Shanks & Dickinson (1988), and Wasserman (1990). These researchers were interested in studying how people associate their own behavior to its consequences, as well as in how people may overestimate the relationship between their behavior and its consequences, showing an illusion of control. It has also been used in research on information seeking strategies, particularly sampling biases, which can also lead to an illusion of control. The illusion of control is a particular type of the illusion of causality, in which the potential cause is the behavior of the participant, so that people believe that their own behavior is controlling uncontrollable events in their environment.

Thus, the active procedure is generally preferred over the passive procedure when studying causal illusions in which the potential cause is the behavior of the participant. However, the passive procedure has advantages, as it allows for more rigorous experimental control over the number of trials of each type, and therefore, over the objective contingency experienced by the participants. Indeed, when using the active procedure, we should be aware that, because the participant is free to decide when to introduce the potential cause (action), we are losing experimental control and surrendering one degree of freedom to the participant (Gibbon et al., 1974). For instance, if the outcome is programmed to occur with identical probability in trials where the participant introduces the action and in trials in which the action is absent, the relationship between the two events should be zero. However, participants wishing to obtain a desired event tend to behave with a high probability of responding, p(R), that is, they tend to introduce the cause in a high number of trials (Matute, 1996). This behavior exposes them mainly to trials where the outcome occurs in the presence of their behavior (i.e., type *a* and type *b* trials). Moreover, as we previously mentioned, conditions that favor the illusion of causality involve a high outcome probability. Thus, if experimental participants introduce the cause with high frequency, there will be many coincidences, that is, this will cause participants to become exposed mainly to type *a* trials (i.e., both cue and outcome are present). This can easily lead to an illusion of control (or, more generally, to an illusion of causality) that would be reflected on their causal judgment (Blanco et al., 2011; Hannah & Beneteau, 2009). This so-called p(R) effect is a type of cue-density effect that occurs when the potential cause is the participants' behavior, and it has been reported in many experiments that use the active procedure as a way to mimic those real-life cases where participants try to obtain desired outcomes.

## Variations of the active procedure

Many experiments aim to investigate whether people are in principle capable to detect a null contingency between their action and the outcome at least under neutral conditions. In these cases, researchers are exploring ways to reduce the probability of responding in the active procedure to make it more neutral and more similar to the passive procedure, in which the experimenter can program the cause to occur with a 50% probability. Thus, in these experiments, participants are not instructed to obtain the outcome, but to learn the degree of causal relationship between their behavior and the outcome. The goal is to equalize the exposure to what happens when participants respond versus when they do not respond, so that they receive less biased information.

A possible strategy is simply instructing participants to introduce the potential cause in approximately 50% of the trials so that they can learn both what happens when they act and when they do not act (Shanks & Dickinson, 1987; Wasserman, 1990). This instruction helps them to behave in a less biased way, and to become exposed to a more complete contingency information. When these instructions are used, judgments tend to be more analytic and less biased than when participants are simply trying to obtain a desired outcome under more naturalistic instructions (Matute, 1996). Nonetheless, the participants still have some degree of control over the actual probability of the cause, so researchers have looked for additional ways to minimize this potential problem.

One approach is a more actively guided procedure in which researchers not only instruct participants on how much to respond, but also when exactly to respond. In this way, researchers can ensure not only that the actual contingency corresponds exactly with the one that has been programmed, but also that the same contingency is maintained through the experimental session (Hannah & Beneteau, 2009). This variation of the active procedure induces judgments that are more similar to those of the passive procedure, and less biased than when participants respond in the active procedure with a high probability of responding.

Additionally, there are many other variables that have also been shown to affect the probability of responding, and therefore, the illusion of causality as well. Introducing any of these variables in experiments using the active procedure will almost certainly modify the participants' behavior, and therefore the information that they are exposed to, and hence their illusion of causality. For instance, telling participants that the drug has severe (vs. no) side effects, makes them reduce their probability of responding (Blanco et al., 2014).

## Variations in the presentation of the Outcome

Some researchers have sensibly argued that outcomes in real life are not an all-or-none event. Indeed, outcomes are often continuous and graded, as when, for example, people feel slightly better or perfectly recovered after having followed a medical treatment. Thus, experiments have also been conducted using continuous outcomes, which gives a greater ecological validity to the experimental designs (e.g., Chow et al., 2019; Double et al., 2020). Moreover, it has also been noted that the successive learning trials in causal learning experiments are usually compressed over a very short period of time (e.g., 30 minutes), whereas in real life learning trials may occur over much extended periods. Current researchers are also using experimental manipulations that are more realistic and involve learning trials over long periods of time (e.g., one trial per day over a 24-days period; Willett & Rottman, 2021).

Another important aspect of outcomes in real life is that they are not always as straightforward as they are in the laboratory experiments, and their occurrence may often be ambiguous. For instance, in a study by Blanco et al. (2020), the outcome (pictures of –fictitious- samples of human tissue) needed to be classified as positive or negative for a certain (fictitious) disease. The stimuli were programmed to be darker or lighter by varying the number of dark pixels in the picture. Positive cases were defined as those showing a greater proportion of "darker" cells (or lighter, depending on the counterbalancing condition), which were determined by varying the number of pixels of each color. Some stimuli were easy to classify (e.g., those having a 20/80 proportion of dark to light pixels were very easy), whereas others were difficult (e.g., those showing a 40/60 and 60/40 proportion). That is, many instances of tissues where relatively ambiguous, in order to mimic the more complex and noisy conditions that people may face in real life. The results indicate that those people who more often interpreted the ambiguous stimuli as outcome occurrences were the ones who showed stronger illusions, thus suggesting that subjective interpretation of ambiguous information matters. A related line of research (Chow et al., 2019) investigated outcomes that were continuous (e.g., varying in magnitude from 0 to 100) rather than binary (e.g., present / absent), which is often a more realistic setting, and replicated the outcome-density bias in these conditions.

## Other variants

There are several other procedural variations, which also affect the results in these experiments. As an example, the frequency with which the participants' judgments are requested is also a critical methodological aspect that needs to be considered. Causal judgments are usually assessed using a continuous scale (e.g., from 0 to 100, or from -100 to +100), and they are usually assessed once at the end of the experiment, However, they can also be assessed every few trials, or even after every single trial. It has been shown that participants emitting their judgments in a Global (i.e., end of training) mode, tend to integrate the information received throughout the experiment. For instance, participants receiving information that suggests a positive causal relationship during the first (or last) part of an experiment, and the opposite information during the other half of the experiment, will tend to integrate both contradictory pieces of information and their final, global judgment, will tend to be of an intermediate value. However, those receiving that type of contradictory information but requested to emit their judgments at the end of each trial (or at the end of every block of a few trials), will be more influenced by the particular value of the information that they have just received in the most recent trials. If the information in the first half of the experiment suggests that there is a positive causal relationship, and the last part of the experiment suggests just the opposite, then the participants' trial-by-trial judgments will be initially higher and then lower, reflecting the most recent information received before each judgment. Their global (end-of-training) judgments, however, will tend to integrate both pieces of information (Collins & Shanks, 2002; Catena et al., 1998; Matute et al., 2002)[1]. Several explanations have been proposed as to why end-of-training and trial-by-trial judgments should behave differently. Indeed, it has been suggested that people tend to update their beliefs from the last

[1] Studies asking participants to emit their judgments using a continuous scale at the end of training (i.e., global mode), may often add predictive yes/no responses during the individual trails (e.g., "will the outcome occur in this trial? y/n") in order to maintain the participants attention through the task. In this case, it should be noted that those yes/no responses are not judgments and are used for different purposes and assessed differently. That is, they do not inform us about the effect of the frequency of judgments.

trial in which they emitted a judgment (Catena et al., 1998), and that the frequency of judgments could work as a contextual cue that frames the subset of data that should be considered when making a judgment (Vadillo et al., 2004). In any case, it seems that asking participants to judge causation (or contingency or other variable) repeatedly during the experiment creates the expectation that their judgments should change, and/or that they should be making their judgments based on a smaller subset of what they have experienced.

Another procedural variable that strongly influences causal judgments is the wording of the test question that is used to assess the participants' judgments. Differential results can be observed depending on whether the question is worded in terms of causality or prediction, or even whether it is worded in the cause-effect direction (i.e., predictive), or in the effect-cause (diagnostic) direction (Crocker, 1982; Don & Livesey, 2018; Matute et al., 1996; Vadillo & Matute, 2007)  These test-question effects are well known in the causal learning literature and some of them have also been explored – and confirmed- in the area of causal illusions that we are discussing herein.

One more variable that may affect the results is the use of preventive causes instead of generative causes. The examples we discussed above include causes that potentially generate their effects (which implies a positive contingency). But causes can also have preventive relationships with their potential effects (which implies a negative contingency). For instance, taking a certain pill may terminate or even prevent a headache. To our knowledge, although there are few studies directly comparing the illusions in generative vs. preventive scenarios, some suggest that the effects that we have described above are similar regardless of the contingency sign (Blanco & Matute, 2015, Matute & Blanco, 2014), while others indicate that there could be differences (e.g., Aeschleman, 2003; Bloom et al, 2007). This calls for the need of further experiments that test this possibility systematically.

## 3. Explaining causal illusions. Theoretical advances

In general, the advances in the methods and applications of the illusion of causality have not been matched by a parallel improvement in the theories that are used to understand it, at least in recent years. Thus, we continue to use the traditional theoretical approaches that have worked relatively well. However, some recent findings suggest that our theories need to be further improved to account for new data. In this section, we will first review the dominant theories proposed to understand the illusion of causality, and then refer to new findings that challenge these theories.

## Current accounts for the illusion of causality.

Traditionally, the illusion of causality has been understood from the causal learning framework and, consequently, the theories that have been used to explain the bias come from this field. These theories can be divided into two big families that propose different approaches to causal learning: rule-based theories, and associative theories (Perales & Shanks, 2007; Pineño & Miller, 2007).

Rule-based theories describe how people combine pieces of information, with the aim of producing a numerical estimation of the contingency that corresponds to the participant's judgment in a contingency learning task. The primitives that these rules combine to produce an estimation are, typically, the frequencies of the four types of trials that can be observed in this type of task and have been described earlier. Generally, rule-based theories make no specific claims about the psychological mechanisms involved in the computation, but often merely describe what the expected rational output is.

Thus, one of the more straightforward contingency rules is the abovementioned $\Delta P$ (Allan, 1980) (see Equation 1), which just expresses the correlation between a potential cause and an outcome. However, this rule can only account for accurate contingency detection, as it cannot produce the type of biased estimations that we see in the empirical literature. Therefore, some have proposed a weighted version of the rule which incorporates a free parameter for each type of trial (i.e., wa, wb, wc, and wd) (Wasserman et al., 1990). By including these parameters, the rule tries to model the idea of different trial types having different weights or levels of importance when judging causality. This makes the rule flexible enough to account for both accurate contingency detection (i.e. when all weights are similar) and biased contingency detection (i.e., when some of the trial types outweigh the others). Specifically, if type *a* trials (i.e., those which contain both the cause and the outcome) are given relatively higher weight in the computation, then the rule predicts systematic overestimations of contingency and, hence, a illusion of causality. This proposal is somewhat consistent with empirical evidence that suggests that, when judging causality, type *a* trials are paid more attention to, or given more importance, than type *d* trials. This in principle would result in a illusion of causality, particularly when type *a* trials are very frequent (Mandel & Vartanian, 2009; Wasserman et al., 1990). A similar cell-weighting idea is used in some variants of the weighted $\Delta P$ rule, such as the Evidence Integration model (Busemeyer, 1991).

In general, rule-based theories were designed to model accurate contingency/causality detection, and therefore they are not well suited to capture biases. It is true that when these models incorporate weights or other mechanisms, they can indeed predict systematic deviations from true contingency, such as the outcome-density or the cause-density effects. However, they do it at the cost of becoming too flexible (for instance, the weighted version of ΔP adds four free parameters to the model), and thus being potentially able to predict one result as well as its opposite.

Since these models are strongly focused on contingency estimation, other models have been developed to better capture causal inferences (which arguably go beyond mere contingency). A remarkable example is Power PC (Cheng, 1997), which attempts to model some invariant properties of causality that subjects can use to estimate causal power. Causal power is a psychological construct that represents the perceived ability of a cause to produce an effect, and unlike ΔP, it includes certain psychological assumptions about causality. For instance, the model distinguishes between causes and "enabling conditions", so that an event that has high correlation with the outcome can be considered irrelevant from the causal power viewpoint. Thus, the presence of oxygen is necessary for a fire to start (enabling condition), but we usually think that a poorly put out cigarette is a more likely cause of fire than oxygen, even though its correlation with fires is small. This means that, in addition to mere contingency between causes and effects, people need to build a causal model that incorporates prior beliefs and knowledge. This is a feature that makes this model somewhat different from other rule-based theories, as it incorporates psychological elements. Power PC still serves to produce mostly accurate causality estimations, although it can be used to predict certain biases: In particular, given a fixed value of ΔP, Power PC can produce different values of causal power depending on the base-rate of the effect, P(O|~C), thus predicting the outcome-density effect (Buehner et al., 2003). However, the model cannot directly account for outcome-density biases in null contingency settings (unless it is assumed that null contingencies are never perceived as exactly zero).

Similar to Power PC, other theories have proposed to incorporate psychological constrains to model causal learning, in particular in the form of probabilistic dependences and prior beliefs as understood in the Bayesian framework. Thus, the Causal Support theory (Griffiths & Tenembaum 2005) compares two competing causal models: one that includes the target causal link and one that does not include it. The participant then uses the observed data to decide which of the two models is more likely. However, to make this judgment, previous knowledge and beliefs must be incorporated as a prior distribution, in a Bayesian update process. Although Causal Support is also designed to produce rational, accurate judgments of causality, it can be used to understand certain biases. Specifically, according to this theory, prior beliefs can affect the causal inference in different ways, for instance, by making one causal model much more likely than the other, by favoring strong causes (those that are more likely to produce the effect) over weak causes, or by assuming the presence of hidden causes in the background (Lu et al., 2008). This reliance on priors endows the theory with enough flexibility to account for different types of performance (either biased or realistic) depending on details of the scenario. For example, two people could differ on their a priori beliefs about the plausibility of a causal link, one being prone to believe that masks work to prevent virus outbreaks, while the other being skeptical. Even when facing the same covariational information, there could be strong differences in their conclusions due to their different priors concerning the probability of the hypothesis that masks work.

Associative theories, on the other hand focus on modeling a learning mechanism that describes how information is acquired, rather than describing how the pieces of information must be combined to produce rational answers. These theories propose that contingency estimation (and thus perception of causality) is based on a process of strengthening and weakening of the association between the mental representations of the cause and the outcome. This association results from the exposure to occurrences and co-occurrences of the two events. Thus, perhaps one of the most celebrated associative theories, the Rescorla-Wagner model (Rescorla & Wagner, 1972), is able to capture the actual contingency between cause and outcome by updating the strength of the association on a trial-by-trial basis (Equation 2):

$$V_t = V_{t-1} + \alpha\beta(\lambda_t - \Sigma V_{t-1}) \quad (2)$$

That is, the associative strength (V) is updated as a function of the discrepancy between the actual outcome that was observed in the current trial (λ) and the sum of the associative strengths of all potential causes that are present at that moment. The impact of this discrepancy on the update process is weighted by means of two learning rate parameters, α and β, which correspond to the salience (or the "ability to produce learning") of the cause and the outcome, respectively.

One feature of the Rescorla-Wagner model is that it can predict both accurate and biased contingency estimations, at least under some combinations of parameter values (see simulations for different scenarios in Blanco & Matute, 2015; Matute et al., 2019). For example, when many type *a* trials are presented, the model predicts an initial overestimation of the contingency that is interpreted as an outcome-density bias. However, these predictions are pre-asymptotic, which means that the model can only produce biases at the beginning of the training stage, and subsequently converges on the right value of contingency (according to the ΔP rule) as more information is acquired (Danks, 2003; Matute et al., 2019). This prediction has been fulfilled in some studies (Murphy et al., 2011, although this task was not causally framed as it was an illusory correlation experiment), but more recent experiments that aim to provide a systematic test for it have failed to find supporting evidence. For example, Barberia et al., (2019) used a very long training procedure and were unable to document the progressive reduction in the bias that associative models predict (see also Blanco, Vadillo & Matute, 2011). Using a different strategy, Moreno-Fernández et al. (2021) found no effect of training length on the illusion of causality. Thus, a clear prediction derived from associative theory (i.e., that contingency overestimations are fleeting and disappear as more information is provided) seems not well documented in the empirical studies.

Additionally, to fully account for results such as outcome- and cause-density biases that persist after a moderate amount of training, the Rescorla-Wagner model requires certain combinations of parameter values (i.e., typically, higher salience for cause-present trials than for cause-absent trials). In fact, one can mirror the same results that the weighted ΔP rule and related models produce by combining different values for the four learning rate parameter values (i.e., for the presence and the absence of both cause and outcome). This means that the two approaches (the associative one and the rule-based one) end up becoming almost identical explanations for the illusion of causality, only differing in the mechanism: whereas weighted rule-based models propose that certain pieces of information are given more importance than others, associative models assume that certain stimuli, or combinations of stimuli, produce more learning. Thus, although we have two promising theoretical accounts for the illusion of causality, none of them seems clearly better endowed than the other to understand the phenomenon. Moreover, as we will review next, the accumulation of evidence that additional factors might play a role in the bias beg for the incorporation of additional mechanisms or more complex models.

## Findings that are hard to fit by current models.

Although contingency detection remains the primary focus in most research investigating causality bias (and, hence, the contingency learning task is still the most widely used procedure), recent experiments have shifted their interest towards examining different processes that can reveal differences between individuals as between scenarios or situations. For example, (Blanco et al., 2018) presented the same null contingency to a group of participants with varying ideological self-positions, ranging from extreme left to extreme right. They also manipulated the causal scenario so that, in one condition, the null contingency corresponded to a left-wing party proposing measures to improve economic indicators, whereas in another condition the corresponding party was of the opposite ideology (right-wing). The results indicated that, despite observing identical contingency information, the description of the party as left- or right-wing affected the illusion of causality selectively: it was more pronounced when the conclusions aligned with the participant's ideology (i.e., right-wing party condition for participants who self-position themselves on the political right, and left-wing party condition for participants who self-identify with the political left). This result suggests that, in addition to the processing of contingency information that both rule-based and associative theories propose, there must be room for additional factors to play a role. Specifically, motivational factors (e.g., trying to fit the observed contingency into a previous cosmovision, set of beliefs, or expectations) could be important to explain differences in causal judgment between individuals and across situations (Vicente et al., 2023). This implies a shift from theories that only consider bottom-up mechanisms to detect covariation or capture statistical regularities (as is the case of associative theories and most rule-based theories, such as ΔP) to more nuanced and complete theories that also incorporate top-down mechanisms as well, which perhaps interact with other processes in complex ways.

In fact, it is possible for current theories to incorporate ways to account for motivational or top-down factors. In particular, ΔP and similar theories can make use of cell weighting, which is a very flexible mechanism, to explain this type of effects. One only needs to assume that cell weights are sensitive to previous beliefs, attitudes or expectations (e.g., type *a* trials are given more importance when they fit in the participant's expectations than when they do not), so that the contingency that is then estimated includes a bias. Similarly, associative theories can use the equivalent mechanism of assuming that the learning rate parameters vary with previous beliefs, with the identical result of weighting the cells in a flexible way. In addition, some associative theories already incorporate response rules (Denniston, Savastano, & Miller, 2001; Pineño, 2007) that allow further flexibility in the observed behavior. These response rules would, on the other hand, imply that at least some factors affecting the illusion of causality do so by means of performance effects, rather than only through encoding or learning effects (Allan et al., 2008). Finally, both Power PC and, more clearly, Causal Support naturally incorporate the influence of prior beliefs, expectations, and knowledge. For example, Power PC assumes certain knowledge of the world to correctly classify the stimuli as potential causes (e.g., cigarette) or enabling conditions (e.g., oxygen); and Causal Support can handle the effect of knowledge and expectations as prior distributions of different forms. In sum, although top-down mechanisms are not explicitly present in most traditional theories used to explain the illusion of causality, there are mechanisms to include this type of influence that should be explored in future research.

Another type of result that may challenge current theories comes from research on certain learning phenomena such as the learned predictiveness effect and the outcome predictability effect (Beesley & Le Pelley, 2010; Liu et al., 2020; Livesey & McLaren, 2007). Participants learn more rapidly about a novel causal relationship when it involves either cues that have been consistently paired with an outcome, or outcomes that were previously learnt as predictable. These effects were initially considered variants of the same phenomenon, although they are probably independent (Thorwart et al., 2017), and together they suggest that, whatever the formalization we choose to indicate different levels of importance for each piece of information (i.e., via learning rate parameters, as in associative theories, or via cell weights, as in the weighted ΔP and related theories), the importance could change during a learning session to respond to previous learning, expectations, or knowledge (Griffiths, Livesey, et al., 2019; Griffiths & Thorwart, 2017). So far, neither type of theory incorporate such sensitivity in the cell weighting/associability mechanism, but it is a logical path of improvement. In this regard, Causal Support may already account for these phenomena via the inclusion of different sets of prior beliefs that are generated by previous training experiences, although the possibility has not been directly addressed.

## 4. Practical implications of causal illusions.

In the preceding sections, we have described what causal illusions are, how researchers measure them, and the theoretical models that have been proposed to explain them. In the following section, we will highlight some of the practical implications that arise from overestimating causality.

First, causal illusions have been linked to irrational beliefs such as paranormal or superstitious beliefs (Blanco et al., 2011; van Elk, 2013; Griffiths, et al., 2019; Matute, 1995). Individuals may develop an illusory link between their own behavior and desired consequences, such us protection from bad luck or increasing chances of getting a desired outcome. This could induce and maintain superstitious behaviors through action biases that promote high p(C), and therefore a cause density effect (Hannah & Beneteau, 2009; Vadillo et al., 2011). When such effects are combined with a naturally high frequency of the outcome, p(O), the overestimation of the effectiveness of actions to prevent or provoke that outcome is facilitated (Allan & Jenkins, 1980; Blanco et al., 2011). For example, a football fan may wear a lucky shirt while watching the match on TV to "ensure" his team wins (wearing the shirt on every match means that the potential cause is presented with high probability). Although no causal relation exists between wearing the lucky shirt and the team winning the match, the fan may develop a causal illusion if the team wins a high proportion of their matches (i.e., a high density of the effect). That is, the illusion is facilitated by a combination between a cause density effect and an outcome-density effect. Although this is a mere example with no serious consequences, similar conditions with more important implications may occur for instance when people believe that using a pseudomedicine can be effective for treating a serious disease.

Recent research has shown that, when estimating causal relations in laboratory settings, people that score high in paranormal beliefs (including superstitious beliefs) display a biased sampling strategy compatible with positive testing. That is, they look for information that eventually allows them to check whether the effect is present when the potential cause is also present, but disregarding or diminishing the interest on the probability of the effect in the absence of the cause (Blanco et al., 2015). As previously mentioned, this biased strategy may promote density effects, thereby serving to either generate such beliefs or perpetuate pre-existing ones (see Yarritu & Matute, 2015 for an example of how previous beliefs may affect sampling strategy). Furthermore, such sampling biases may also contribute to self-esteem by inducing an illusion of control that may maintain the bias. Thus, the illusion of causality may enhance one's perceived ability to positively influence events, contributing to self-esteem. In line with this, research has demonstrated that individuals exhibiting high levels of depressive symptomatology demonstrate a less biased sampling strategy and, therefore, greater accuracy in establishing causal relationships compared to non-depressed counterparts (Alloy & Abramson, 1979; Blanco et al., 2012).

Research on irrational beliefs, such as paranormal and superstitious ideas, has often used the active procedures that we have described in previous sections, because they allow to investigate how people behave. However, some studies have also revealed the potential effect of individual differences in evidence weighting, and their role on these beliefs. For instance, Griffiths et al. (2019) showed that in addition to the biased sampling strategy, people who are prone to superstitious beliefs also tend to overweight those events that seem to confirm the relation between the potential cause and the alleged consequence (confirmatory events) over events that disconfirm the relation. Similar results have been found when exploring pseudoscientific beliefs using a passive procedure, for example Vicente, Blanco & Matute (2023) have recently found a relation between pseudoscientific beliefs and information weighting. In this research, pseudoscientific beliefs were related to the magnitude of causal illusions on a passive task in which the outcome appeared with high frequency (outcome-density effect), a result that is consistent with previous results reported by Torres et al. (2020). However, these results also showed that the magnitude of that correlation depends on prior beliefs in such a way that the illusion was stronger when information was compatible with prior beliefs than when it was not. Thus, the illusory effect of an alternative medicine varied as a function of general confidence in alternative therapies: those participants with high confidence in alternative medicine developed stronger illusions toward a fictitious alternative product than those participants with a lower confidence in alternative medicine (see also Blanco et al., 2018 for similar results on political appraisal).

As illustrated in the preceding example, one context where pseudoscientific beliefs can have repercussions is in the domain of health. Holding the belief that an ineffective treatment is efficacious may lead to a wide range of harms, not only in terms of health outcomes (e.g., delaying the acceptance of evidence-based treatments), but also in financial terms by increasing healthcare costs. Previous research has focused on investigating factors that may contribute to the utilization of pseudomedicine.

The absence of side effects has also been identified as a potential factor that can influence individuals to use alternative or unproven treatments because these treatments are perceived as safer and less risky. However, a high frequency of use may result in the development of a cause density effect, leading to causal illusions (Blanco et al., 2014). On the other hand, illnesses that tend to resolve naturally without intervention may create an illusion of causality due to the high number of positive outcomes, even if the treatment had no effect (Blanco & Matute, 2020).

The implications that biased information seeking strategies, and together with the effect of prior beliefs, entail for health outcomes are particularly relevant in today's world. The Internet has emerged as a significant source of health-related information, with web search engines being frequently used for accessing information concerning diagnoses and treatments. In this context, not only previous beliefs may guide and bias these searches, but also sampling strategies may be biased (Marcu et al., 2019) facilitating the emergence of cue and outcome-density biases that eventually will contribute to misinterpretation of the causal relation between events (see Moreno-Fernández & Matute, 2020). For example, a person who is concerned about COVID-19 may use the Internet to check information about symptoms. This search is already biased as it is focused on a specific cause (COVID-19), and it will provide mostly type *a* and *b* information. The presence of a common symptom for COVID-19 may induce a causal bias if no information regarding the specificity of that symptom is collected. Thus, for example, a common symptom for COVID 19 is fever, but fever is a very unspecific symptom associated with other infectious diseases, that is, it is a non-specific symptom.

Although there has been a lot of focus on the impact of causal illusions on health and their link to irrational beliefs, such as pseudoscientific beliefs, there has been a recent interest in exploring the contribution of causal illusions on different domains, such as education. For example, the adoption of instructional practices that lack a scientific basis could be problematic, as such strategies may replace evidence-based interventions, leading to negative consequences for academic performance. Double et al. (2020) conducted a series of six experiments to investigate the potential impact of causal illusions on judgments regarding the effectiveness of novel teaching methods. Therefore, the authors adapted the contingency learning paradigm to suit the educational context and requested participants to evaluate the efficacy of a novel instructional technique in a null contingency setting with a high effect density, where there was a high probability of achieving excellent academic performance irrespective of the teaching method employed. Their results consistently showed that causal illusions may emerge in educational scenarios, thereby highlighting a plausible mechanism for explaining why educators may overestimate the effectiveness of teaching methods lacking scientific support.

Finally, another goal of applied research on this topic is to design and test interventions aimed at reducing or preventing the bias. We will mention three promising approaches: First, educational interventions that aim to reduce information sampling biases seem to work fine to expose participants to a more complete set of contingency information and, hence, to improve accuracy in their judgments. These interventions have been tested on high school and college students (Barberia et al., 2013, 2019; Martinez et al., 2023). A second approach is to provide the complete contingency table as a means to refuting erroneous causal conclusions (MacFarlane et al., 2018). Finally, a third approach consists of suggesting alternative causes for the outcome occurrences (Vadillo et al., 2013). These three approaches show promise to reduce or prevent the type of causal biases we have described in this article.

## 5. Challenges and future directions.

In this section, we will outline some future challenges that research on causal illusions will need to address in the following years.

Probably one of the most challenging phenomena that humankind is facing nowadays is related to misinformation dissemination. In the era of the Internet, exposure to fake or misleading information has become a worldwide concern. The potential effects of misinformation on critical areas such as healthcare, education or policymaking has raised the interest of the scientific community, and psychological research has started to describe cognitive, social and affective processes involved in false beliefs vulnerability (Ecker et al., 2022). However, few studies have focused on the role that causal illusions may play in the development of such beliefs. Some exceptions include, for example, the study conducted by Saltor et al. (2023) in which fake news discriminability was predicted by causal illusions vulnerability suggesting that cognitive processes involved in causal illusion may also be playing a role in fake news discriminability. Similarly, more research is needed to extent causal illusion research to other relevant and related field such as science denial (see e.g., Sulik et al., 2020).

Translational research on causal illusions has also experienced a considerable growth during the last years. Understanding biased behaviors may be considered the very first step for developing interventions that may positively contribute to reduce the negative effects associated to these biases or to enhance their positive outcomes. We have already mentioned that basic research has been extended to applied setting (Barberia et al., 2013; Martinez et al., 2023; Rodriguez-Ferreiro et al., 2021) and educational programs aimed at promoting critical thinking has been developed. However, one of the missing aspects of this research is related to the long-term effects of these interventions and their potential generalization to related contexts.

In our opinion, one of the most promising avenues for future theoretical development is the incorporation of explicit top-down processes into the causal learning, as a means to capture the variability that human exhibit in this type of tasks. For example, we have noted above that prior beliefs and expectancies seem to modulate the illusion of causality (e.g., Blanco et al. 2018: Vicente et al., 2023) in a way that reminds of phenomena like motivated reasoning (Kunda, 1990), although these beliefs are not explicitly incorporated into theories like Rescorla-Wagner. Furthermore, the active version of the task that we have described above allows for investigating the potential effect of other top-down influences such as the confirmation bias (i.e., selecting information that confirms, rather than disconfirms, the current hypothesis) (see e.g., Moreno-Fernández & Matute, 2020).

Finally, there are a number of potential methodological advances that could be considered in future research. For example, although many studies use a unidirectional scale (i.e., from 0 to 100) to assess the causal judgment, some previous (Neunaber & Wasserman, 1986) and recent (Ng, Lee, & Lovibond, 2023) experiments suggest that bidirectional scales should be generally preferred as they reduce the bias under some conditions. Also, given the apparent influence of top-down factors and previous beliefs that we commented above, researchers must reflect on the cover-story that is used to represent the task, since it could play a role in the causal illusion (Fugelsang & Thompson, 2023). Another potential methodological improvement relates to the current reliance on online samples for conducting psychological research. Although the Internet has been proven a useful arena for recruiting big and diverse samples, and some studies have been conducted to directly compare online and offline samples in phenomena related to causal illusions (see e.g., Matute et al., 2007), some recent cases have pointed out the dangers in "professional" samples such as M-Turk, as they can compromise data quality (Sulik et al, 2023). Hence, researchers should increase their quality control by including attention checks and other measures that are not yet widely used in this literature.

## Acknowledgments

Support for this research was provided by Grant PID2021-126320NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe, as well as Grant IT1696-22 from the Basque Government.

## Conflict of interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Disclaimer

Formatting, following the templates provided by Psicológica, spelling, grammar-checking and correct referencing are the sole responsibility of the authors.

## References

Aeschleman, S. R., Rosen, C. C., & Williams, M. R. (2003). The effect of non-contingent negative and positive reinforcement operations on the acquisition of superstitious behaviors. *Behavioural Processes, 61,* 37–45. https://doi.org/10.1016/s0376-6357(02)00158-4

Aitken, M. R. F. R., Larkin, M. J., & Dickinson, A. (2000). Super-learning of causal judgements. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology, 53*(1), 59-81. https://doi.org/10.1080/713932716

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15*(3), 147-149. https://doi.org/10.3758/BF03334492

Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Experimental Psychology, 34*(1), 1-11. https://doi.org/10.1037/h0081013

Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14*(4), 381-405. https://doi.org/10.1016/0023-9690(83)90024-3

Allan, L. G., Hannah, S. D., Crump, M. J. C., & Siegel, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General, 137*(2), 226-243. https://doi.org/10.1037/0096-3445.137.2.226

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of experimental psychology. General, 108*(4), 441-485. https://doi.org/10.1037/0096-3445.108.4.441

Barberia, I., Blanco, F., & Rodríguez-Ferreiro, J. (2020). The more, the merrier: Treatment frequency influences effectiveness perception and further treatment choice. *Psychonomic Bulletin & Review, 28,* 665–675. https://doi.org/10.3758/s13423-020-01832-6

Barberia, I., Blanco, F., Cubillas, C. P., & Matute, H. (2013). Implementation and Assessment of an Intervention to Debias Adolescents against Causal Illusions. *PloS ONE, 8*(8), e71303-e71303. https://doi.org/10.1371/journal.pone.0071303

Barberia, I., Vadillo, M. A., & Rodríguez-Ferreiro, J. (2019). Persistence of causal illusions after extensive training. *Frontiers in Psychology, 10*(24), 1-9. https://doi.org/10.3389/fpsyg.2019.00024

Beesley, T., & Le Pelley, M. E. (2010). The effect of predictive history on the learning of sub-sequence contingencies. *Quarterly journal of experimental psychology, 63*(1), 108-135. https://doi.org/10.1080/17470210902831767

Blanco, F. (2017). Positive and negative implications of the causal illusion. Consciousness and *Cognition, 50,* 56-68. https://doi.org/10.1016/j.concog.2016.08.012

Blanco, F., & Matute, H. (2015). Exploring the Factors That Encourage the Illusions of Control The Case of Preventive Illusions. *Experimental Psychology, 62*(2), 131-142. https://doi.org/10.1027/1618-3169/a000280

Blanco, F., & Matute, H. (2019). Base-rate expectations modulate the causal illusion. *PloS ONE, 14*(3), e0212615-e0212615. https://doi.org/10.1371/journal.pone.0212615

Blanco, F., & Matute, H. (2020). Diseases that resolve spontaneously can increase the belief that ineffective treatments work. *Social Science & Medicine, 255,* 113012-113012. https://doi.org/10.1016/J.SOCSCIMED.2020.113012

Blanco, F., Barberia, I., & Matute, H. (2014). The lack of side effects of an ineffective treatment facilitates the development of a belief in its effectiveness. *PLOS ONE, 9*(1) : e84084. https://doi.org/10.1371/journal.pone.0084084

Blanco, F., Barberia, I., & Matute, H. (2015). Individuals who believe in the paranormal expose themselves to biased information and develop more causal illusions than nonbelievers in the laboratory. *PLOS ONE, 10*(7): e0131378. https://doi.org/10.1371/journal.pone.0131378

Blanco, F., Gómez-Fortes, B., & Matute, H. (2018). Causal illusions in the service of political attitudes in Spain and the United Kingdom. *Frontiers in Psychology, 9,* 1033-1033. https://doi.org/10.3389/fpsyg.2018.01033

Blanco, F., Matute, H., & Vadillo, M. A. (2011). Making the uncontrollable seem controllable: The role of action in the illusion of control. *Quarterly journal of experimental psychology, 64*(7), 1290-1304. https://doi.org/10.1080/17470218.2011.552727

Blanco, F., Matute, H., & Vadillo, M. A. (2012). Mediating Role of Activity Level in the Depressive Realism Effect. *PloS ONE, 7*(9). https://doi.org/10.1371/journal.pone.0046203

Blanco, F., Matute, H., & Vadillo, M. A. (2013). Interactive effects of the probability of the cue and the probability of the outcome on the overestimation of null contingency. *Learning & behavior, 41*(4), 333-340. https://doi.org/10.3758/s13420-013-0108-8

Blanco, F., Moreno-Fernández, M. M., & Matute, H. (2020). Are the symptoms really remitting? How the subjective interpretation of outcomes can produce an illusion of causality. *Judgment & Decision Making, 15*(4).

Bloom, C. M., Venard, J., Harden, M., & Seetharaman, S. (2007). Non-contingent positive and negative reinforcement schedules of superstitious behaviors. *Behavioural Processes, 75*, 8–13. https://doi.org/10.1016/j.beproc.2007.02.010

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1119-1140. https://doi.org/10.1037/0278-7393.29.6.1119

Busemeyer, J. R. (1991). Intuitive statistical estimation. In N. H. Anderson (Ed.), *Contributions to information integration theory* (pp. 187-205). Erlbaum.

Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of frequency of judgement and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 481. https://doi.org/10.1037/0096-1523.24.2.481

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*(2), 367-405. https://doi.org/10.1037/0033-295X.104.2.367

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of personality and social psychology, 58*(4), 545-567. https://doi.org/10.1037/0022-3514.58.4.545

Chow, J. Y. L., Colagiuri, B., & Livesey, E. J. (2019). Bridging the divide between causal illusions in the laboratory and the real world: The effects of outcome density with a variable continuous outcome. *Cognitive research: principles and implications, 4*(1), 1-1. https://doi.org/10.1186/s41235-018-0149-9

Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition, 30*, 1138-1147. https://doi.org/10.3758/bf03194331

Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin, 8*, 214-220. https://doi.org/10.1177/0146167282082005

Danks, D. (2003). Equilibria of the Rescorla–Wagner model. *Journal of Mathematical Psychology, 47*(2), 109-121. https://doi.org/10.1016/S0022-2496(02)00016-0

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Hillsdale, NJ: Erl-baum.

Don, H., & Livesey, E. Is the blocking effect sensitive to causal model? It depends how you ask. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 1630–1635). Austin, TX: Cognitive Science Society.

Double, K. S., Chow, J. Y. L., Livesey, E. J., & Hopfenbeck, T. N. (2020). Causal illusions in the classroom: How the distribution of student outcomes can promote false instructional beliefs. *Cognitive Research: Principles and Implications, 5,* Article 34. https://doi.org/10.1186/s41235-020-00237-2

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*(1), Article 1. https://doi.org/10.1038/s44159-021-00006-y

Elk, M. van. (2013). Paranormal believers are more prone to illusory agency detection than skeptics. *Consciousness and Cognition, 22*(3), 1041-1046. https://doi.org/10.1016/j.concog.2013.07.004

Fugelsang, J. A., & Thompson, V. A. (2003). A dual-process model of belief and evidence interactions in causal reasoning. Memory & Cognition *31*, 800–815. https://doi.org/10.3758/BF03196118

Gibbon, J., Berryman, R., & Thompson, R. L. (1974). Contingency spaces and measures in classical and instrumental conditioning. *Journal of the Experimental Analysis of Behavior, 21*(3), 585-605. https://doi.org/10.1901/jeab.1974.21-585

Griffiths, O., & Thorwart, A. (2017). Effects of Outcome Predictability on Human Learning. *Frontiers in Psychology, 8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00511

Griffiths, O., Livesey, E., & Thorwart, A. (2019). Learned biases in the processing of outcomes: A brief review of the outcome predictability effect. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*(1), 1-16. https://doi.org/10.1037/xan0000195

Griffiths, O., Shehabi, N., Murphy, R. A., & Le Pelley, M. E. (2019). Superstition predicts perception of illusory control. *British Journal of Psychology, 110*(3), 499–518. https://doi.org/10.1111/bjop.12344

Hannah, S. D., & Beneteau, J. L. (2009). Just tell me what to do: Bringing back experimenter control in active contingency tasks with the command-performance procedure and finding cue density effects along the way. *Canadian Journal of Experimental Psychology 63*(1), 59–73. https://doi.org/10.1037/a0013403

Hannah, S. D., Crump, M. J. C., Allan, L. G., & Siegel, S. (2009). Cue-interaction effects in contingency judgments using the streamed-trial procedure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 63*(2), 103-112. https://doi.org/10.1037/a0013521

Hume, D. (1748). *An enquiry concerning human understanding*. Clarendon.

Katagiri, M., Kao, S.-F., Simon, A. M., Castro, L., & Wasserman, E. A. (2007). Judgments of causal efficacy under constant and changing interevent contingencies. *Behavioural Processes, 74*(2), 251-264. https://doi.org/10.1016/j.beproc.2006.09.001

Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108(3), 480-498. https://doi.org/10.1037/0033-2909.108.3.480

Liu, W., Livesey, E. J., Lachnit, H., Don, H. J., & Thorwart, A. (2020). Does learning history shape the associability of outcomes? Further tests of the outcome predictability effect. *PloS ONE, 15*(12), e0243434. https://doi.org/10.1371/journal.pone.0243434

Livesey, E. J., & McLaren, I. P. L. (2007). Elemental associability changes in human discrimination learning. Journal of Experimental Psychology. *Animal Behavior Processes, 33*(2), 148-159. https://doi.org/10.1037/0097-7403.33.2.14

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115*(4), 955-984. https://doi.org/10.1037/a0013256

MacFarlane, D., Hurlstone, M. J., & Ecker, U. K. H. (2018). Reducing demand for ineffective health remedies: Overcoming the illusion of causality. *Psychology and Health, 33*(12), 1472-1489. https://doi.org/10.1080/08870446.2018.1508685

Mandel, D. R., & Vartanian, O. (2009). Weighting of contingency information in causal judgement: Evidence of hypothesis dependence and use of a positive-test strategy. *The Quarterly Journal Of Experimental Psychology, 62*(12), 2388-2408. https://doi.org/10.1080/17470210902794148

Marcu, A., Muller, C., Ream, E., & Whitaker, K. L. (2019). Online Information-Seeking About Potential Breast Cancer Symptoms: Capturing Online Behavior With an Internet Browsing Tracking Tool. *Journal of medical Internet research, 21*(2), e12400-e12400. https://doi.org/10.2196/12400

Martínez, N., Rodríguez-Ferreiro, J., Barberia, I., & Matute, H. (2023). A debiasing intervention to reduce the causality bias in undergraduates: The role of a bias induction phase. *Current psychology, 42*, 32456-32468. https://doi.org/10.1007/s12144-022-04197-2

Matute, H. (1995). Human reactions to uncontrollable outcomes: Further evidence for superstitions rather than helplessness. *The Quarterly Journal Of Experimental Psychology, 48B*(2), 142-157. https://doi.org/10.1080/14640749508401444

Matute, H. (1996). Illusion of control: Detecting response-outcome independence in analytic but not in naturalistic conditions. *Psychological Science, 7*(5), 289-293. https://doi.org/10.1111/j.1467-9280.1996.tb00376.x

Matute, H., & Blanco, F. (2014). Reducing the illusion of control when an action is followed by an undesired outcome. *Psychonomic Bulletin & Review, 21*(4), 1087-1093. https://doi.org/10.3758/s13423-014-0584-7

Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(1), 182–196. https://doi.org/10.1037/0278-7393.22.1.182

Matute, H., Blanco, F., & Díaz-Lago, M. (2019). Learning Mechanisms Underlying Accurate and Biased Contingency Judgments. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*(4), 373-389. https://doi.org/10.1037/xan0000222

Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: How they bias our everyday thinking and how they could be reduced. *Frontiers in psychology, 6,* 888. https://doi.org/10.3389/fpsyg.2015.00888

Matute, H., Vadillo, M. a, Vegas, S., & Blanco, F. (2007). Illusion of control in Internet users and college students. *CyberPsychology & Behavior, 10*(2), 176-181. https://doi.org/10.1089/cpb.2006.9971

Matute, H., Vegas, S., & De Marez, P.-J. J. (2002). Flexible Use of Recent Information in Causal and Predictive Judgments. *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(4), 714-725. https://doi.org/10.1037/0278-7393.28.4.714

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. *British journal of psychology*, *102*(3), 392-405. https://doi.org/10.1348/000712610X532210

Moreno-Fernández, M. M., & Matute, H. (2020). Biased Sampling and Causal Estimation of Health-Related Information: Laboratory-Based Experimental Research. *Journal of medical Internet research, 22*(7), e17502-e17502. https://doi.org/10.2196/17502

Moreno-Fernández, M. M., Blanco, F., & Matute, H. (2017). Causal illusions in children when the outcome is frequent. *PloS ONE, 12*(9), e0184707. https://doi.org/10.1371/journal.pone.0184707

Moreno-Fernández, M. M., Blanco, F., & Matute, H. (2021). The tendency to stop collecting information is linked to illusions of causality. *Scientific Reports, 11*(1), 3942. https://doi.org/10.1038/s41598-021-82075-w

Murphy, R. a, Schmeer, S., Vallée-Tourangeau, F., Mondragón, E., & Hilton, D. (2011). Making the illusory correlation effect appear and then disappear: The effects of increased learning. *Quarterly journal of experimental psychology, 64*(1), 24-40. https://doi.org/10.1080/17470218.2010.493615

Neunaber, D. J., & Wasserman, E. A. (1986). The effects of unidirectional versus bidirectional rating procedures on college students' judgments of response-outcome contingency. *Learning and Motivation, 17*(2), 162-179. https://doi.org/10.1016/0023-9690(86)90008-1

Ng, D. W., Lee, J. C., & Lovibond, P. F. (2023). Unidirectional rating scales overestimate the illusory causation phenomenon. *Quarterly journal of experimental psychology* (2006). Advance online publication. https://doi.org/10.1177/17470218231175003

Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review, 14*(4), 577-596. https://doi.org/10.3758/bf03196807

Pineño, O. (2007). A response rule for positive and negative stimulus interaction in associative learning and performance. *Psychonomic Bulletin & Review, 14*(6), 1115-1124. https://doi.org/10.3758/bf03193100

Pineño, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning. *The Quarterly Journal of Experimental Psychology, 60*(3), 310-329. https://doi.org/10.1080/17470210601000680

Rescorla, R. A., & Wagner, A. R. (1972). *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement*. En A. H. Black & W. F. Prokasy (Eds.), Classical Conditioning II: current research and theory (pp. 64-99). Appleton-Century-Crofts.

Rodríguez-Ferreiro, J., Vadillo, M. A., & Barberia, I. (2023). Debiasing Causal Inferences: Over and Beyond Suboptimal Sampling. *Teaching of Psychology, 50*(3), 230-236. https://doi.org/10.1177/00986283211048394

Saltor, J., Barberia, I., & Rodríguez-Ferreiro, J. (2023). Thinking disposition, thinking style, and susceptibility to causal illusion predict fake news discriminability. *Applied Cognitive Psychology, 37*(2), 360-368. https://doi.org/10.1002/acp.4008

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In *The psychology of learning and motivation: Advances in research and theory*, Vol. 21 (pp. 229-261). Academic Press.

Shanks, D. R., & Dickinson, A. (1988). Associative Accounts of Causality Judgment. *The psychology of learning and motivation: Advances in research and theory, 21,* 229-261. https://doi.org/10.1016/S0079-7421(08)60030-4

Sulik, J., Ross, R. M., Balzan, R., & McKay, R. (2023). Delusion-Like Beliefs and Data Quality: Are Classic Cognitive Biases Artifacts of Carelessness? *Journal of Psychopathology and Clinical Science, 6*(132), 794-760. https://doi.org/10.1037/abn0000844

Sulik, J., Ross, R., & Mckay, R. (2020). The contingency illusion bias as a potential driver of science denial. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 829-835). Austin, TX: Cognitive Science Society.

Thorwart, A., Livesey, E. J., Wilhelm, F., Liu, W., & Lachnit, H. (2017). Learned predictiveness and outcome predictability effects are not simply two sides of the same coin. *Journal of Experimental Psychology. Animal Learning and Cognition, 43*(4), 341-365. https://doi.org/10.1037/xan0000150

Torres, M. N., Barberia, I., & Rodriguez-Ferreiro, J. (2020). Causal illusion as a cognitive basis of pseudoscientific beliefs. *British journal of psychology, 111*(4), 840-852. https://doi.org/10.1111/bjop.12441

Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *Quarterly journal of experimental psychology, 60*(3), 433-447. https://doi.org/10.1080/17470210601002520

Vadillo, M. A., Matute, H., & Blanco, F. (2013). Fighting the Illusion of Control: How to Make Use of Cue Competition and Alternative Explanations. *Universitas Psychologica, 12(*1), 261-270.

Vadillo, M. A., Musca, S. C., Blanco, F., & Matute, H. (2011). Contrasting cue-density effects in causal and prediction judgments. *Psychonomic bulletin & review, 18*, 110-115. https://doi.org/10.3758/s13423-010-0032-2

Vadillo, M. A., Vegas, S., & Matute, H. (2004). Frequency of judgment as a context-like determinant of predictive judgments. *Memory & cognition, 32*(7), 1065-1075. https://doi.org/10.3758/BF03196882

Vicente, L., Blanco, F., & Matute, H. (2023). I want to believe: Prior beliefs influence judgments about the effectiveness of both alternative and scientific medicine. *Judgment and Decision Making, 18*, 1-1.

Wasserman, E. A. (1990). Detecting Response-Outcome Relations: Toward an Understanding of the Causal Texture of the Environment. *Psychology of Learning and Motivation - Advances in Research and Theory, 26*(C), 27-82. https://doi.org/10.1016/S0079-7421(08)60051-1

Wasserman, E. A., Dorner, W. W., & Kao, S. (1990). Contributions of Specific Cell Information to Judgments of Interevent Contingency. *Cognition, 16*(3), 509-521.

White, P. A. (2004). Causal judgment from contingency information: A systematic test of the pCI rule. *Memory & cognition, 32*, 353-368. https://doi.org/10.3758/bf03195830

Willett, C. L., & Rottman, B. M. (2021). The Accuracy of Causal Learning Over Long Timeframes: An Ecological Momentary Experiment Approach. *Cognitive Science, 45*(7), e12985. https://doi.org/10.1111/cogs.12985

Yarritu, I., & Matute, H. (2015). Previous knowledge can induce an illusion of causality through actively biasing behavior. *Frontiers in Psychology, 6*, 389-389. https://doi.org/10.3389/fpsyg.2015.00389