How do Participants Interpret Trials from Individual Cells in a Causal Illusion Task?

Peter F. Lovibond (p.lovibond@unsw.edu.au)[1]

Julie Y.L. Chow (julie.chow@unsw.edu.au)[1]

Jessica C. Lee (jessica.c.lee@sydney.edu.au)[1 2]

1 School of Psychology, University of New South Wales, NSW 2052, Australia

2 School of Psychology, University of Sydney, NSW 2006, Australia

Abstract

In a causal illusion task, participants rate a cue that has an objectively null contingency with an outcome as causal. Trials are usually organized according to a 2x2 table representing the presence/absence of a binary cue and a binary outcome. Cell A outcomes (cue, outcome) can be attributed to the cue. But how do participants interpret trials from cell C (no cue, outcome), where the cause of the outcome is unspecified? In two experiments we asked participants to provide causal explanations for cell A and C trials in a medicine-recovery causal illusion task. Participants who reported that the cause of cell C outcomes (e.g., strong immunity, spontaneous recovery) did not also apply to cell A outcomes showed the strongest causal illusion. Such a causal reasoning process undermines the logic behind the delta P metric typically used to define a contingency, and thereby provides a potential normative account of causal "illusions".

Keywords: causal illusion; contingency; causal reasoning; delta P; hidden cause

In a typical contingency learning task, participants are presented with a series of trials involving a single binary cue (e.g., administration of a medicine) and a single binary outcome (e.g., recovery from a disease). When present, the cue typically precedes the outcome. Trials are classified in a 2x2 contingency table according to the presence or absence of the cue and outcome (Table 1). The objective contingency is calculated by delta P, the difference in the probability of the outcome in the presence and the absence of the cue (Ward & Jenkins, 1965).

Table 1: Cells in a 2x2 contingency table.

	Outcome	No outcome
Cue	A	B
No cue	C	D

A special case is when these two probabilities are the same, representing a null contingency. Under these conditions, despite the fact that the outcome is equally likely regardless of whether the cue is present or absent, participants tend to judge the relationship as positive. This empirical phenomenon is referred to as an illusory correlation or causal illusion (Matute, Blanco & Díaz-Lago, 2019). It is strongest when the overall cue density (proportion of trials on which the cue is presented) and outcome density (proportion of trials on which the outcome is presented) are both high (e.g., .75; Blanco, Matute & Vadillo, 2013). Causal illusions have attracted considerable interest both as an example of biased cognitive processing and as a potential mechanism for the acquisition of social stereotypes and false causal beliefs (e.g., Goldwater, 2020; Hamilton & Gifford, 1976; Lassiter, 2002; Torres, Barberia & Rodriguez-Ferreiro, 2020).

Theories of causal illusion focus on biased processing of certain trial types, in particular excessive weighting of trials from cell A (cue followed by outcome; e.g., White, 2003) or a reliance on marginal frequencies (Bott, Kellen & Klauer, 2021; Fiedler, Kutzner & Vogel, 2013). However, an alternative approach is to consider contingency learning as a causal reasoning problem. From this perspective, cell C trials (no cue followed by outcome) are ill-defined because it is not clear what caused the outcome. Previous research suggests that under such ambiguity, participants may infer the presence of an additional hidden cause to explain the outcome (e.g., Hagmayer & Waldmann, 2007).

Associative theories (e.g., Rescorla & Wagner, 1972) avoid this problem by assuming that the experimental context acts as a stable background cue which can enter into associations with the outcome. Indeed, it is this feature that allows the theory to predict learning in line with delta P (at asymptote). In causal terms, the context serves as an omnipresent potential cause that can account for any outcomes that occur in the absence of the target cue (Shanks & Dickinson, 1987). However, it is not clear if participants make causal inferences along these lines when an uncued outcome occurs. For example, they might assume that whatever factor causes the outcome on cue absent (cell C) trials is, unlike a background cue, *not* also present on cell A trials. Such inferences could potentially help explain non-normative judgments such as causal illusions.

Accordingly, we set out to record the causal inferences that participants report when faced with outcomes that occur in the presence and absence of the cue in a causal illusion (null contingency) task, using the medicine-recovery scenario. Our approach was informed by the study of Luhmann and Ahn (2011), who used probe questions to assess participants' explanations for trials from cells A and B. We extended this approach to cell C as we were particularly interested in participants' causal explanations for outcomes that occurred in the absence of the target cue.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

Experiment 1

In this experiment we assessed participants' spontaneous causal inferences through open-ended probe questions after individual trials from cells A, B and C, as well as open-ended post-experimental questions about each of these types of trial. To test whether asking the probe questions during trials might influence participants' causal ratings, we included a No Probe control group in which these questions were omitted.

Method

Participants One hundred and seventy participants (54 female, 114 male, 2 other) were recruited on the Prolific online testing platform and paid for their participation (17 min at £6GBP/hr). They were randomly allocated on a 2:1 ratio to a Probe group and a No Probe group. Data from 22 participants were excluded due to failing instruction checks, leaving 97 in the Probe group and 51 in the No Probe group.

Apparatus and Stimuli The experiment used a medical scenario developed by Matute, Yarritu & Vadillo (2011). The experiment was programmed using the jspsych library (de Leeuw, 2015), hosted using JATOS (Lange, Kühn, & Filevich, 2015) and run on participants' web browsers.

Procedure The project was approved by the UNSW research ethics committee, and participants provided online consent. Instruction screens explained that their task was to play the role of a doctor assessing whether a new drug "Cloveritol" was effective as a medicine in treating patients with a disease "Linda syndrome". Participants were told they would see the results from a series of patients and would be asked to predict the outcome for each patient, before being told the actual outcome. After instruction checks, they were then shown a series of 64 trials consisting of four blocks of 16 trials in which the frequencies for cells A/B/C/D were 9/3/3/1 (see Figure 1). The order of trials within each block was randomized. These frequencies represent a cue density of 0.75 (that is, 75% of patients received the drug), an outcome density of 0.75 (75% of patients recovered), and a delta P of zero: p(recovery|drug =p(recovery|no drug) = 0.75.

Each trial started with the text "The patient was given:". Below the text was either an image of a pill bottle with the word "Cloveritol" underneath, or a grayed-out picture of the pill bottle with "No Treatment" underneath. Below that was the instruction "Please rate the likelihood that this patient will recover from their illness" together with a sliding scale labelled "Definitely WILL NOT RECOVER" at the left extreme and "Definitely WILL RECOVER" at the right extreme. Participants clicked on the scale to record their prediction, and then on a "Continue" button to proceed. The prediction instruction and scale were then replaced with the actual outcome: either "Patient recovered" in green text or "Patient did not recover" in red text. After 2.5 sec the screen turned blank and after a 1-sec inter-trial interval, the next trial started. The primary purpose of the prediction rating was to ensure participants maintained their attention to the task, and the data will not be presented here.

Immediately after one trial from each of cells A, B and C, participants in the Probe group were given the question: "You just observed a patient who recovered/did NOT recover after being given Cloveritol/No treatment. Please describe why you think this patient recovered/did not recover" followed by a text box. This strategy was based on the approach developed by Luhmann and Ahn (2011), but used an open-ended response format and included cell C trials as well as A and B. Each type of probe trial was randomly selected to occur on one of the final three blocks of training, with only one probe trial presented per block.

Immediately after the final trial, participants were asked to rate "to what extent Cloveritol prevented vs caused recovery from illness relative to no treatment" using a bidirectional scale labelled "Strongly PREVENTED RECOVERY" at the left extreme, "No effect" in the middle, and "Strongly CAUSED RECOVERY" at the right extreme. Ratings were recorded on a scale from -100 to 100. These causal ratings provided the primary measure of causal judgment and hence, potentially, causal illusion. Participants were then asked how many patients fell into each of the categories defined by the 4 contingency table cells, as well as to predict how many of 100 new patients given the drug, and 100 new patients given no treatment, would recover (Barberia, Vadillo & Rodriguez-Ferreiro, 2019). These frequency measures were used to investigate a separate issue and will not be presented here.

The final questions asked participants to report what they were thinking during the experiment when they saw patients from each of cells A, B and C. For example, the cell A question was "On some occasions, some patients recovered from Linda syndrome after taking Cloveritol. Why did you think this happened?", followed by a text box. All participants were presented with the three post-experimental questionnaires in a randomized order.

Results and Discussion

The mean causal rating for the Probe group was 24.08, and for the No Probe group it was 30.35. Each of these means was significantly greater than the normative value of zero (Probe: $t(96)=5.85$, $p<.001$; No Probe: $t(50)=5.36$, $p<.001$). However the difference between them was not significant ($t<1$). Thus there was no detectable effect of having answered the probe questions on the size of the causal illusion.

Two raters blind to participants' causal ratings carried out a content analysis of the open-ended text answers to each of the post-trial probe questions and each of the post-experimental open-ended questions, one at a time. The raters then conferred and agreed on a small number of categories to use to classify participants' answers. They then re-coded the answers into the agreed categories. Data for the probe questions were similar to those for the post-experimental questions, but somewhat more variable, so we focus here on classification of answers to the post-experimental questions, which were available for participants in both groups (Probe and No Probe). The agreed categories for each question are listed in Table 1. Participants who gave no answer or an uninterpretable answer were excluded from analysis.

For cell A, more than half the participants attributed recovery to the medicine alone and did not mention any other factor. Some mentioned the medicine plus another factor such as spontaneous recovery, and some asserted that only the other factor was causal. For cell B (medicine with no recovery), most participants (83 of 148) cited an extenuating factor such as these patients having weak immune systems or more severe illness. Some stated that the medicine was only partially effective, and a small number used the opportunity to state that the medicine was not effective at all. Cell C had the most diverse explanations, including that these patients had stronger immunity, less severe illness, or they recovered spontaneously (e.g., "by themselves"). A small number cited other causes and some again stated the medicine was ineffective. At this point, before seeing the causal rating data for each category, we planned contrasts to compare mean causal ratings across categories.

For cell A, we expected that participants who mentioned only the medicine as a cause for recovery would give higher causal ratings than those who considered other causes. As shown in the final column in Table 2, this difference was in the expected direction but did not reach significance, $F(1, 144)=3.045$, $p=.083$. For cell B, we expected those who attributed lack of recovery to the medicine (partially effective plus ineffective) to give lower causal ratings than those who attributed it to another factor such as the type of patient. However, this pattern was not observed and the contrast was not significant, $F<1$. For cell C, we expected the minority who stated the medicine was ineffective to give lower causal ratings than the majority who attributed it to some other factor, but again this pattern was not observed and the contrast was not significant, $F<1$.

Table 2: Sample size and mean causal rating for each category of explanation for cells A-C in Experiment 1.

Cell	Category	n	Rating
A	Medicine	85	33.88
	Medicine plus Other	25	20.92
	Other	27	10.67
	Excluded	11	
B	Other	83	24.33
	Medicine Partially Effective	38	39.16
	Medicine Ineffective	12	10.85
	Excluded	15	
C	Immunity	62	24.44
	Less Severe	32	20.88
	Spontaneous Recovery	20	28.25
	Other	10	46.20
	Medicine Ineffective	13	24.92
	Excluded	11	

Thus far, the results for cell A supported our conjecture that many participants (85/148) did not consider factors other than the drug as contributing to recovery in patients who received the drug, even though cell C trials implied the existence of such additional factors and participants readily identified them. However, classification based on cell A alone did not yield a statistically significant difference between these participants and the remainder. We therefore conducted a second classification of cell A attributions, based on participants' responses for *both* cells A and C, again blind to the causal ratings given by each participant. We reasoned that seeing both responses together would allow us to better classify the overlap between cell A and cell C inferred causes. We found that all of the "other" causes nominated for cell A outcomes aligned with the cause/s nominated by that participant for cell C outcomes. Furthermore, we were able to re-code the cell A responses for some participants based on the combined information from the two cells. For example, one participant's cell A response was "The Cloveritol helped their immune system", which was originally coded as "Medicine" since the mention of immune system could have referred to a mechanism by which the medicine worked rather than a separate cause. This participant's cell C response was "The immune system can do all the job", which allowed us to re-classify their cell A response as Medicine plus Other. Finally, we were able to make a finer distinction within the Medicine plus Other subgroup, on the basis of the order in which they described these two causes or the relative importance they ascribed to them in their answers. Thus, we had four categories for cell A attributions based on joint consideration of participants' cell A and cell C responses: Medicine, Medicine then Other, Other then Medicine, and Other.

When we analyzed causal ratings as a function of this new categorization, we saw our predicted ordering, with ratings decreasing as a function of the degree of emphasis participants gave to causes other than the medicine (Figure 1). A planned contrast comparing the Medicine category to the other three categories was significant, $F(1,136)=7.29$, $p=.008$). However, further orthogonal contrasts within the other three categories did not reach significance, both $Fs<1$.

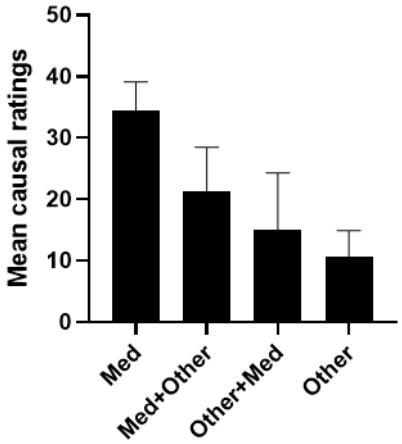

Figure 1: Causal ratings as a function of attributions for cell A outcomes, based on participants' explanations for both cells A and C, in Experiment 1

Experiment 2

Experiment 1 showed that participants are able to generate plausible explanations for cell C outcomes, such as immunity or spontaneous recovery. Despite this, more than half of the participants failed to mention these causes as potential explanations for cell A outcomes, which they instead appeared to attribute solely to the drug. Furthermore, several of the commonly mentioned causes for cell C outcomes, such as less severe illness or stronger immune systems, implied that participants thought these patients differed from cell A patients.

However, the data were derived from separate questions probing each cell. Therefore, in Experiment 2 we developed forced choice response options and we directly asked participants whether their inferred causes for cell C outcomes also applied to cell A outcomes. We also manipulated cue density to see whether greater experience with cue absent trials (cells C and D) would provide more opportunity to recognize their potential relevance to cue present trials (cells A and B). We hypothesized that participants who inferred a cause for cell C outcomes that was not also present on cell A trials would provide stronger causal ratings for the target cue compared to participants who recognized that the cause of cell C outcomes could also explain cell A outcomes.

Method

Participants Two hundred participants (69 female, 124 male, 7 other) were recruited on the Prolific platform. They were randomly allocated to a 75% Cue Density group and a 50% Cue Density group. After exclusions there were 86 participants in the 75% Cue Density group and 97 in the 50% Cue Density group.

Apparatus and Stimuli The experiment used the same scenario and testing method as in Experiment 1.

Procedure The trial procedure was the same as in Experiment 1, except that no probe questions were administered. The 75% Cue Density group was given the same trial structure as in Experiment 1, with trial frequencies for cells A/B/C/D of 9/3/3/1. The 50% Cue Density group was instead given trials with frequencies of 6/2/6/2. Both groups received 75% outcome density, as in Experiment 1.

In this experiment the open-ended post-experimental questions were replaced with two forced choice questions (see Table 3). The first question referred to cell C, and the options included the two most common open-ended explanations from Experiment 1 (spontaneous recovery and immunity) as well as an open-ended option and an option to say they hadn't thought about this issue during the experiment. On the basis of Experiment 1, we didn't expect substantial differences in causal ratings between participants who selected these options. Rather, the purpose of the first question was to encourage participants to reflect on how they interpreted the cell C trials, so that we could refer to this explanation in the second question, to assess whether participants inferred the same cause for cell A. Here the options included the medicine, the explanation they had chosen for cell C, a mixture of the two, and again an option that they hadn't thought about this issue. Our planned contrasts compared causal ratings for 1) those who chose the medicine option with those who chose either of the options that included their cell C explanation, and 2) those who chose the cell C cause alone with those who chose the medicine plus cell C mixture.

Table 3: Sample size and mean causal rating for each forced choice option for cells C and A in Experiment 2.

	75% Cue Density		50% Cue Density	
Forced choice question	n	Rating	n	Rating
1. (cell C) Some patients **recovered** from Linda Syndrome **after no treatment**. Did you think this was:				
because of the natural course of the illness (spontaneous recovery)	37	39.72	49	21.33
because these patients had stronger immune systems	35	37.83	37	24.08
because of something else (please describe below) <text box>	8	31.88	8	22.13
I didn't think about this issue during the experiment	6		3	
2. (cell A) Some patients **recovered** from Linda Syndrome **after taking Cloveritol.** Did you think this was:				
mostly because of the medicine	21	52.10	20	40.20
mostly for the same reason you chose for the previous question	12	12.83	32	0.09
partly because of the medicine and partly for the previous reason	51	36.92	45	29.82
I didn't think about this issue during the experiment	2		0	

Table 3 shows that participants' causal ratings varied substantially as a function of the option they chose for the cell A forced choice question. Those who chose the medicine option gave significantly higher mean causal ratings than those who chose the options that included the cause they nominated for cell C, $F(1, 175)=15.79$, $p<.001$. Furthermore, those who chose the mixture option gave higher ratings than those who chose the cell C cause alone, $F(1,175)=14.44$, $p<.001$. Neither of these contrasts interacted with group, both $Fs<1$. Accordingly, Figure 2 shows the mean causal ratings for the first three cell A options, collapsed over groups, illustrating the same pattern as in Experiment 1 (Figure 1).

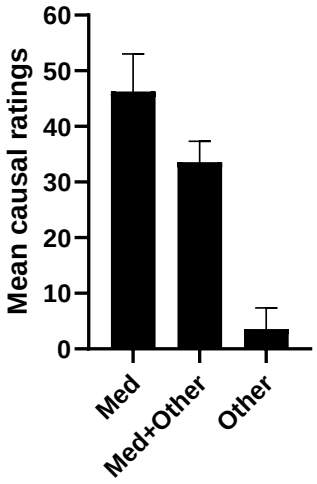

Figure 2: Causal ratings as a function of attributions for cell A outcomes in Experiment 2

General Discussion

Both experiments showed a robust causal illusion at the group level, and Experiment 2 replicated the well-established finding that the illusion is stronger at a higher cue density. Experiment 1 indicated that participants spontaneously nominated a variety of hidden causes of cell C (cue absent) outcomes, including higher immunity, lower severity, and spontaneous recovery. These explanations were not associated with substantial differences in participants' causal ratings for the medicine cue.

However, both experiments also showed differences between participants in their causal attributions for cell A outcomes, and these attributions were systematically associated with their causal ratings for the medicine. In both experiments, participants who only nominated the medicine as the cause of recovery in patients given the medicine showed the strongest causal illusions. Conversely, participants who attributed cell A recoveries to the same factor as cell C recoveries (immunity, lower severity etc.) gave causal ratings closer to the expected normative value of zero. Those who nominated both the medicine and a cell C cause gave intermediate causal ratings.

The conventional way to interpret the high causal ratings in participants who gave a causal role to the medicine is as a processing bias. These participants have failed to appreciate that the factor that caused recovery in patients not given the medicine also applied to those who were given the medicine, and since the recovery rate was the same, the medicine was in fact completely ineffective. This is the rationale behind considering delta P as a normative standard.

However, this logic depends critically on an assumption that participants might not share, namely that the cause of cell C outcomes is also present on cell A trials. In a true clinical trial, this assumption is strengthened by the process of randomization of participants to the drug or control conditions, which matches participant variables such as severity of illness. However, like previous causal illusion research using the medical scenario, we did not instruct participants that administration of the medicine was randomized. Furthermore, randomization does not preclude subsequent differences in causal exposure. For example, patients who are not improving may seek a different treatment (and some participants mentioned this possibility). In real world causal reasoning, it seems even less likely that causal factors would be perfectly matched so as to allow isolation of the target cause.

Thus, the present data suggest that we should be cautious in assuming that participants necessarily consider the causes of cell C outcomes also apply to cell A outcomes. A substantial proportion of participants nominated different causes for these outcomes, such as medicine for cell A and spontaneous recovery for cell C. It could be argued that participants who nominated the medicine as the cause of cell A recoveries in Experiment 1 were being lazy or failing to fully report their causal beliefs. However, we saw the same pattern in Experiment 2 where they rejected equally easy forced choice options that included other causal factors. Hagmayer and Waldmann (2007) similarly found that participants tended not to infer the presence of an unobserved cause when an observable cause was already present.

If we take seriously the possibility that some participants attribute cell A and C outcomes to independent causes, this provides a plausible alternative account for the high causal ratings that have previously been labelled causal illusions. These participants have inferred a second, hidden, causal factor that is capable of generating the same outcome as the target cue but is not present on trials that include the target cue. Such a causal structure is quite coherent; the existence of two separate causes for a given outcome is an entirely commonplace occurrence. Yet it undermines the logic of the delta P standard if the second causal factor is not taken into account. In other words, delta P is only an appropriate normative standard when we can be confident that the causes of cell C outcomes are also present on cell A trials.

Note we are not arguing that participants who nominate a second cause that is only present on cell C trials *are* in fact acting normatively. It is entirely possible, and indeed likely, that people systematically underestimate the existence of alternative causes or their relevance to explaining outcomes when a target cause is present (e.g., see Fernbach, Darlow & Sloman, 2010). Rather, our main point is that we can't say definitively that these participants are *not* acting normatively.

A simple one-cue, one-outcome laboratory task is too ambiguous to adjudicate this issue (cf. Luhmann & Ahn, 2011).

Applied to real-world false causal beliefs such as the efficacy of pseudo-medicines, the present data support previous debiasing approaches that highlight the relevance of the base rate to undermine the perceived effectiveness of the target cause (e.g., MacFarlane, Hurlstone & Ecker, 2018). More specifically, our data suggest it may be valuable to directly challenge idiosyncratic causal beliefs that people invoke to explain (away) outcomes that occur in the absence of the target cause.

In summary, we have shown that when participants are exposed to a null contingency between a single cue and an outcome, with no other causal information available, they explain cell C (cue absent) outcomes in a variety of ways. Those who invoke the same cause for cell C and cell A outcomes tend to give causal ratings for the cue that are close to zero. Those who invoke different causes tend to give much higher causal ratings. Both of these ways of understanding the situation are defensible and can be seen as potentially normative, rather than the latter being seen as an illusion or error. The task is intrinsically indeterminate, and the conventional view that the correct causal rating is zero can only be justified if all other causes are assumed to be absent or held constant. This assumption is typically not made clear in laboratory "causal illusion" tasks, and is even less likely to hold in real world situations.

Acknowledgments

This study was funded by an Australian Research Council Discovery Project Grant [DP190103738] to Peter Lovibond. Jessica Lee was supported by a Discovery Early Career Researcher Award from the Australian Research Council [DE210100292]. We thank four anonymous referees for very helpful comments on this paper.

References

Barberia, I., Vadillo, M.A. & Rodriguez-Ferreiro, J. (2019). Persistence of causal illusions after extensive training. *Frontiers in Psychology*, *10*, 24.

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*, 333-340.

Bott, F.M., Kellen, D. & Klauer, K.C. (2021). Normative accounts of illusory correlations. *Psychological Review*, *128*, 856-878.

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*, 1-12.

Fernbach, P.M., Darlow, A. & Sloman, S.A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychological Science*, *21*, 329-336.

Fiedler, K., Kutzner, F. & Vogel, T. (2013). Pseudocontingencies: Logically unwarranted but smart inferences. *Current Directions in Psychological Science*, *22*, 324-329.

Goldwater, M.B. (2020). Health medication and causal illusions. *Meeting of the Cognitive Science Society*.

Hagmayer, Y. & Waldmann, M. (2007). Inferences about unobserved causes in human contingency learning. *Quarterly Journal of Experimental Psychology*, 60, 330–55.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*, 392-407.

Lange, K., Kühn, S., & Filevich, E. (2015). " Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS One*, *10*, e0130834.

Lassiter G.D. (2002). Illusory causation in the courtroom. *Current Directions in Psychological Science, 11(6)*, 204-208.

Luhmann, C.C. & Ahn, W-K. (2011). Expectations and interpretations during causal learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *37*, 568 –587.

MacFarlane, D., Hurlstone, M.J., & Ecker, U.K.H. (2018). Reducing demand for ineffective health remedies: overcoming the illusion of causality. *Psychology & Health*, *33*, 1472-1489.

Matute, H., Blanco, F. & Díaz-Lago, M. (2019). Learning mechanisms underlying accurate and biased contingency judgments. *Journal of Experimental Psychology: Animal Learning & Cognition*, *45*, 373-389.

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. British Journal of Psychology, 102, 392-405.

Rescorla, R.A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II*. NY: Appleton-Century-Crofts.

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.

Torres, M. N., Barberia, I., & Rodriguez-Ferreiro, J. (2020). Causal illusion as a cognitive basis of pseudoscientific beliefs. *British Journal of Psychology*, *111*, 840-852.

Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, *19*, 231.

White, P.A. (2003). Making causal judgments from the proportion of confirming instances: The pCI rule. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *29*, 710-727.