

# Contrasting cue-density effects in causal and prediction judgments

Miguel A. Vadillo · Serban C. Musca ·  
Fernando Blanco · Helena Matute

Published online: 12 November 2010  
© Psychonomic Society, Inc. 2010

**Abstract** Many theories of contingency learning assume (either explicitly or implicitly) that predicting whether an outcome will occur should be easier than making a causal judgment. Previous research suggests that outcome predictions would depart from normative standards less often than causal judgments, which is consistent with the idea that the latter are based on more numerous and complex processes. However, only indirect evidence exists for this view. The experiment presented here specifically addresses this issue by allowing for a fair comparison of causal judgments and outcome predictions, both collected at the same stage with identical rating scales. Cue density, a parameter known to affect judgments, is manipulated in a contingency learning paradigm. The results show that, if anything, the cue-density bias is stronger in outcome predictions than in causal judgments. These results contradict key assumptions of many influential theories of contingency learning.

**Keywords** Contingency learning · Causal learning · Cue-density bias · Causal judgment · Prediction judgment

The ability to acquire causal knowledge is essential for humans' survival and well-being. It allows predicting future

events on the basis of present ones and to plan actions in order to achieve desired goals. Therefore, this ability has been extensively studied by psychologists with the aim of understanding the cognitive processes underlying it. One of the most widely used paradigms to study causal knowledge is a very simple one where information on the presence or absence of a cue (C) and on the presence or absence of an outcome (O) is given to the participants on a trial-by-trial basis (Jenkins & Ward, 1965). That is, in each trial the cue is either present (C) or absent ( $\sim$ C) and the outcome either occurs (O) or does not occur ( $\sim$ O). If the cue is a cause of the occurrence of the outcome, the outcome should occur more often in the presence than in the absence of the cue, other things being equal. Based on this reasoning, the  $\Delta p$  index was proposed by Jenkins & Ward (1965; see also Allan, 1980; Cheng & Novick, 1992) as a normative measure of causality:

$$\Delta p = p(O|C) - p(O|\sim C) \quad (1)$$

This index has positive values when the cue is a generative cause of the outcome. Negative values of  $\Delta p$ , on the other hand, correspond to cases where the cue is a preventive cause of the outcome. Finally, a  $\Delta p$  of zero is obtained when the outcome occurs as frequently in the presence of the cue as in its absence, that is, when the outcome occurs independently of the cue.<sup>1</sup>

Importantly, when the joint frequencies of C and O are manipulated so that the value of  $\Delta p$  remains unchanged, normative analyses predict that people's estimations should still be based on this  $\Delta p$  value, regardless of that manipulation. While not expected from a normative viewpoint, a number of manipulations that do not affect the

---

M. A. Vadillo (✉) · H. Matute  
Departamento de Fundamentos y Métodos de la Psicología,  
Universidad de Deusto,  
Apdo. 1,  
48080, Bilbao, Spain  
e-mail: mvadillo@deusto.es

S. C. Musca  
CRPCC (EA 1285), Université Rennes 2,  
F-35000, Rennes, France

F. Blanco  
Katholieke Universiteit Leuven,  
Leuven, Belgium

---

<sup>1</sup> Cheng (1997) proposed another statistical index,  $p$ , as an alternative normative referent for causal judgments. However, since the cue density effect explained below affects both  $\Delta p$  and  $p$  in a similar way, for the sake of simplicity, we focus on the simpler,  $\Delta p$  rule.

objective  $\Delta p$  value nevertheless have an impact on people's estimations of causality. The *outcome-density bias* consists of the fact that, for a fixed  $\Delta p$ , ratings of contingency increase with the overall probability of the outcome,  $p(O)$  (e.g., Allan & Jenkins, 1983; Allan, Siegel, & Tangen, 2005; Alloy & Abramson, 1979; Buehner, Cheng, & Clifford, 2003; Matute, 1995; Musca, Vadillo, Blanco, & Matute, 2010; Wasserman, Kao, van Hamme, Katagiri, & Young, 1996). Likewise, some researchers have found that, for a fixed  $\Delta p$ , ratings of contingency increase with the overall probability of the cue,  $p(C)$ , an effect known as *cue-density bias* (e.g. Allan & Jenkins, 1983; Matute, Yarritu, & Vadillo, 2010; Perales, Catena, Shanks, & González, 2005; Wasserman et al., 1996). In spite of the formal parallelism between both density effects, the available evidence strongly suggests that the cue-density effect is smaller and less robust than the outcome-density effect (e.g., Hannah & Beneteau, 2009; Perales & Shanks, 2007).

To understand how participants acquire and use the covariational information within these experiments, many types of dependent variables have been measured. In a standard contingency learning task, a predictive question is generally used in each trial, just after the presentation of the cue and before the presentation of the corresponding outcome; as its name specifies, participants have to indicate by means of a *yes/no* response whether they think that the outcome will occur given the presence/absence of the cue in that trial. Upon completion of the learning phase, there is usually a final test phase where the participants have to provide a causal judgment by rating on a numerical scale the perceived strength of the causal link between the cue and the outcome.

Interestingly, both cue- and outcome-density effects have been found in causal judgments assessed after the learning phase but not in the outcome predictions requested on a trial-by-trial basis during the learning phase (e.g., Allan et al., 2005; Perales et al., 2005). In light of this evidence, some authors argue that the deviations from the normative value that occur in participants' causal judgments are due to additional processes intervening in causal estimations as compared to those intervening in predictions. The processes underlying causal estimations would thus be more numerous and complex (and consequently, more prone to errors) than those underlying predictions.<sup>2</sup> For instance, Allan et al.

(2005) proposed that trial-by-trial predictions reflect participant's sensitivity to the objective cue-outcome contingency.<sup>3</sup> However, according to Allan et al., causal judgments would involve not only participant's knowledge of the cue-outcome relationship, but also a decision process that can give rise to biases such as the outcome-density bias.

Nevertheless, this view is based on a comparison that can be misleading because it disregards the fact that causal judgments and predictive responses that appear to be dissociated in those experiments are not collected in a comparable way. Indeed, causal judgments are collected after completion of the learning phase and by means of a numerical rating, while predictive responses are collected on a trial-by-trial basis during the learning phase and by means of yes/no responses. Thus, these dependent variables differ not only in their predictive or causal status, but also in a number of procedural details. A few studies have already explored whether predictions, causal judgments and other subjective ratings of covariation are sensitive to the same information (Blanco, Matute, & Vadillo, 2010; De Houwer, Beckers, & Vandorpe, 2007; Vadillo & Matute, 2007; Vadillo, Miller, & Matute, 2005), but to our best knowledge, no experiment has manipulated cue or outcome density and collected both causal and predictive judgments at the same time and with the same rating scale so as to allow an unbiased comparison of human's prediction and causation abilities. In the experiment presented here, we offer such a comparison of cue-density effects in causal and predictive judgments, by collecting them at the same stage of the experiment with identical rating scales.

The reason why we are testing the cue-density effect is that the covariational manipulation must be one that neither affects the normatively expected causal judgment nor the normatively expected prediction judgment. An outcome-density manipulation does not satisfy this condition: With such a manipulation, one would expect, from a normative point of view, different prediction judgments as a function of the outcome density (i.e., participants' normatively expected predictions of the outcome are to be higher if the outcome occurs frequently than if it occurs with a low probability), but no effect on the causal judgments. Thus, because the normatively expected impact of outcome density on predictive and causal judgments is distinct, a direct comparison of the outcome-density effect on these two types of judgments is unfair, at best. By contrast, a cue-density manipulation

<sup>2</sup> Note that this interpretation is arguable: If a given judgment is based on more complex processes, it is natural that this judgment shows more random variance, but it is unclear why this variance should result in any systematic bias, such as the cue- and outcome-density effects. However, given that the assumption made by these authors is reasonable and plausible, we will test the predictions that can be drawn from it, without questioning its validity. In any case, the assumption that outcome predictions should be easier to make than causal judgments is also implicit in the rationale behind the computation of the  $\Delta p$  rule: Contingency is computed on the basis of conditional probabilities (which are useful to make predictions), and not the opposite.

<sup>3</sup> The dependent variable in Allan et al. (2005),  $\Delta p_{\text{PRED}}$ , is computed as the difference between the proportion of "yes" responses in cue-present trials and the proportion of "yes" responses in cue-absent trials. Participants perceiving a positive cue-outcome contingency are expected to give more "yes" responses in cue-present than in cue-absent trials. Thus, this index is assumed to be an indirect measure of participants' perception of contingency.

should, from a normative viewpoint, affect neither the prediction of the outcome nor the causal judgments, so that participants' predictive and causal judgments can be compared straightforwardly. Therefore, cue density was manipulated in the following experiment to check whether it does more readily induce a bias in causal or in prediction judgments, even though it is well-known that the cue-density effect is usually weak and elusive, and that, in this sense, it might be a suboptimal manipulation to induce a systematic bias on judgments (see Hannah & Beneteau, 2009; Perales & Shanks, 2007).

## Method

### Participants and apparatus

One hundred and forty-four anonymous Internet users voluntarily took part and were randomly assigned to each of two groups. This resulted in 71 participants in the *High Cue Density* (hereafter, High) group and 73 participants in the *Low Cue Density* (hereafter, Low) group. The experimental program used is an adaptation of the allergy task that has been extensively used in contingency learning experiments (e.g., Wasserman, 1990). The experiment was run on the Internet, implemented as an HTML document dynamically modified with JavaScript that any computer connected to the World Wide Web with a standard Internet browser can run. Previous experiments conducted with this task showed that the results obtained over the Internet are virtually identical to those obtained under traditional laboratory conditions (e.g., Vadillo & Matute, 2007; Vadillo et al., 2005).

### Design and procedure

In the current version of the allergy task participants were asked to imagine that a space alien from Mars was offered carrots, which it ate (C) or did not eat ( $\sim$ C), and then the Martian felt sick (O) or did not feel sick ( $\sim$ O). Each trial started with the presentation of the phrase "The Martian ate / did not eat carrots". The participant had to click a "Click when ready" rectangle located below that phrase in order to continue. On click, with the phrase still present, the rectangle was substituted by a predictive question, which read "Do you think the Martian will be sick?" and the participants had to choose between a "Yes" and a "No" answer. Once the "Yes" or "No" rectangle was clicked, the question disappeared from the screen. With the information on the cue still present on the upper part of the screen, a pre-programmed outcome was displayed on the lower part of the screen. It

consisted of the phrase "The Martian is OK/sick", of a happy/sad smiley and of a "Click to continue" rectangle that once clicked triggered the next trial.

Upon completion of the training trials, participants were presented with the test phase. This consisted of a prediction question and a causal question, with presentation position (upper/lower half of the same screen) counterbalanced between participants. The questions read: "If the Martian ate carrots, how likely is it that it will be sick?" (prediction judgment), and "To what extent do the carrots have the power to make the Martian feel sick?" (causal judgment). Below each question a 101-point scale ranging from 0 to 100 was displayed. For the prediction judgment, 0 was labeled as "Very unlikely" and 100 as "Very likely". For the causal judgment, 0 was labeled as "Definitely it is not the cause" and 100 as "Definitely it is the cause". Participants were able to answer the questions in the order they preferred, through a click on the corresponding scale. On click, the value corresponding to their answer was displayed and remained visible. Participants had the opportunity to correct their answers as many times as they wanted. Although causal judgments are sometimes collected by means of a bidirectional scale (from  $-100$  to  $100$ ), we decided to request causal judgments on a unidirectional scale ( $0-100$ ) in order to improve their comparability with predictive judgments (which can only take positive values).

Cue density was manipulated between participants. For group High, cue density was of .80, with 38 trials in which C and O co-occurred (type *a* trials), 26 trials in which C was present but O was absent (type *b* trials), four trials in which C was absent but O was present (type *c* trials), and 12 trials in which neither C nor O were present (type *d* trials). For group Low, cue density was of .20, with 13 *a*, 3 *b*, 29 *c*, and 35 *d* trials. With these frequencies of each trial type, the overall density of the outcome,  $p(O)$ , was .525 in both groups. Contingency, as measured by  $\Delta p$ , was about .35 in both groups (specifically, .344 in group High and .359 in group Low). The probability of the outcome in the presence of the cue,  $p(O|C)$ , was .59 in Group High and .81 in group Low. These differences in the  $p(O|C)$  are an unavoidable consequence of manipulating the density of the cue while keeping contingency and  $p(O)$  constant, with a nonzero positive contingency. However, they were carefully chosen so that, if anything, they worked against the observation of a cue-density effect. The sequence of trials was randomized for each participant.

## Results

Upon application of the studentized deleted residuals (SDR) outliers-detection method proposed by McClelland (2000),

data from three participants were eliminated from further analyses because their ratings showed extreme between-judgments difference (all SDRs > 3), which would have compromised the homoscedasticity assumed in the following mixed analysis of variance. We also removed the data from one additional participant who responded “yes” in almost all cue-absent trials during training (60 out of 64,  $|z| > 3.50$ ), which indicates that this participant was not paying attention to the experiment. The following analyses were conducted with the remaining participants: 70 in group High and 70 in Low.

The pattern of results depicted in Fig. 1 shows a cue-density effect in the prediction judgments, whereas no such effect was found in causal judgments. A 2 (Group: High vs. Low)  $\times$  2 (Judgment: Prediction vs. Causal) mixed ANOVA yielded a significant interaction between both factors,  $F(1, 138) = 5.01, p < 0.05$ . Planned comparisons showed that there were no significant differences in causal judgments,  $t(138) < 1$ . However, there was a significant cue-density effect in predictive judgments,  $t(138) = 2.02, p < 0.05$ . Thus, the pattern of results observed here provides no support for the hypothesis that causal judgments are more biased than predictive ones. If anything, the significant interaction between cue-density and type of judgment supports the opposite conclusion.

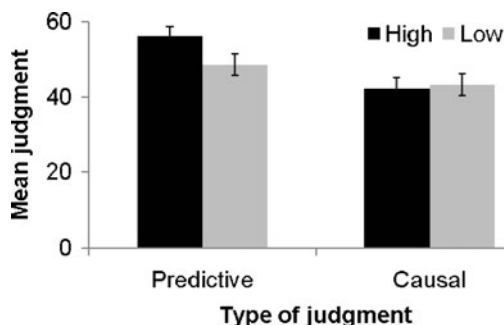
## Discussion

One quite popular view, common to authors that champion diverse and even conflicting theories, is that causal judgments are based on more numerous and complex processes than predictive responses, which in turn would explain why people depart from what is normatively expected less often when they say what is going to happen than when they say whether an event is or is not the cause of another (e.g. Allan et al., 2005; Perales et al., 2005). However, when looking closer at the available evidence, it becomes clear that such

interpretation is based on an unwarranted comparison between yes/no predictive responses collected during training, on the one hand, and numerical causal judgments collected after training, on the other hand. It does not seem reasonable to compare prediction and causation dependent variables that lie on different measuring scales and are collected at different points in time during the experiment. Indeed, it has long been shown that the moment when a judgment is collected and the frequency with which it is collected can change dramatically the participants’ responses (e.g., Catena, Maldonado, & Candido, 1998; Collins & Shanks, 2002; Matute, Vegas, & De Marez, 2002).

Even more important, causal judgments may seem more difficult than predictive responses in previous studies not because of the causal/predictive distinction, but simply because a numerical judgment is more complex than a yes/no response. Because of their dichotomous nature, discrete predictive responses bear greater similarity to the structure of the cues and outcomes: Just as cues and outcomes are either present or absent, a predictive response is either positive or negative. On the other hand, a numerical judgment somehow requires that the participants ignore the dichotomous nature of the cue and outcome information they have been presented with during the training phase and engage in a probabilistic evaluation, the result of which is to be expressed numerically.

If one accepts the idea that cue- and outcome-density biases can be taken as indicative of the relative complexity of the processes involved in several types of judgment (Allan et al., 2005; Perales et al., 2005), then the present results suggest, if anything, that predicting the outcome is more complex than inferring causality. Although this conclusion is at odds with the prevailing framework described in the Introduction, the idea that causal judgments might be more automatic or intuitive than predictions based on conditional probabilities is consistent with a recent tendency to see causal structure judgments as more primary and fundamental than probabilistic judgments (for a review, see Lagnado, Waldmann, Hagmayer, & Sloman, 2007). From this point of view, people would first infer causal structure on the basis not only of contingency information, but also on the basis of several additional cues, such as the timing of events, interventions, or previous knowledge. The assessment of the strength of the links connecting causes and effects in this structure (i.e., the parameters or weights) would take place only subsequently. Given that this updating of the weights would be necessary to make accurate probabilistic predictions, this would explain why outcome predictions seem to be more complex and open to biases than causal judgments. In other words, inferring the causal structure is a necessary, but not sufficient, step towards making accurate predictions.



**Fig. 1** Mean predictive and causal judgments given by participants. Whiskers stand for standard error of the mean.

The idea that outcome predictions might require more cognitive processing than causal judgments is also consistent with our previous attempts to explain the divergences between both types of judgments from an associative point of view. For example, Vadillo and Matute (2007) showed that manipulating the order in which trials *c* and *d* were presented to participants biased causal judgments towards recency without any significant effect on outcome predictions. We proposed that this pattern of results could be easily accommodated by the influential Rescorla and Wagner (1972) associative model by assuming that causal judgments might be a relatively direct expression of the association between the target cue and the outcome, whereas outcome predictions might arise from a combination of the information contained in several associations (e.g., context-outcome association and cue-outcome association). Although the particular details of the model proposed therein are not well suited to account for the specific pattern of data found in this experiment (e.g., the asymptotic predictions of the model are insensitive to cue-density effects), the general idea behind the model (i.e., that causal judgments are directly based on single associations, while outcome predictions require a combination of associative strengths) is perfectly consistent with the present results.

**Authors Note** Support for this research was provided by Dirección General de Investigación of the Spanish Government (grant SEJ2007-63691/PSIC) and Dirección General de Investigación, Tecnología y Empresa de the Junta de Andalucía (grant SEJ-406). Correspondence concerning this article should be addressed to Miguel A. Vadillo (Departamento de Fundamentos y Métodos de la Psicología, Universidad de Deusto, Apartado 1, 48080 Bilbao, Spain). E-mail: mvadillo@deusto.es

## References

- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgement tasks. *Bulletin of the Psychonomic Society*, *15*, 147–149.
- Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation*, *14*, 381–405. doi:10.1016/0023-9690(83)90024-3
- Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior*, *33*, 250–263.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgements of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*, 441–485. doi:10.1037/0096-3445.108.4.441
- Blanco, F., Matute, H., & Vadillo, M. A. (2010). Contingency is used to prepare for outcomes: Implications for a functional analysis of learning. *Psychonomic Bulletin & Review*, *17*, 117–121. doi:10.3758/PBR.17.1.117
- 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*, 1119–1140. doi:10.1037/0278-7393.29.6.1119
- Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of the frequency of judgment and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 481–495. doi:10.1037//0096-1523.24.2.481
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405. doi:10.1037//0033-295X.104.2.367
- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365–382. doi:10.1037//0033-295X.99.2.365
- Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition*, *30*, 1138–1147.
- De Houwer, J., Vandorpe, S., & Beckers, T. (2007). Statistical contingency has a different impact on preparation judgments than on causal judgments. *Quarterly Journal of Experimental Psychology*, *60*, 418–432. doi:10.1080/17470210601001084
- 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*, 59–73. doi:10.1037/a0013403
- Jenkins, H. M., & Ward, W. C. (1965). Judgement of contingency between responses and outcomes. *Psychological Monographs*, *79*, 1–17.
- Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195176803.003.0011
- Matute, H. (1995). Human reactions to uncontrollable outcomes: Further evidence for superstitions rather than helplessness. *Quarterly Journal of Experimental Psychology*, *48B*, 142–157.
- Matute, H., Vegas, S., & De Marez, P. J. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 714–725. doi:10.1037/0278-7393.28.4.714
- Matute, H., Yarritu, I., & Vadillo, M. A. (2010). Illusions of causality at the heart of pseudoscience. *British Journal of Psychology*. doi:10.1348/000712610X532210
- McClelland, G. H. (2000). Nasty data: Unruly, ill-mannered observations can ruin your analysis. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social Psychology* (pp. 393–411). Cambridge, UK: Cambridge University Press.
- Musca, S. C., Vadillo, M. A., Blanco, F., & Matute, H. (2010). The role of cue information in the outcome-density effect: Evidence from neural network simulations and a causal learning experiment. *Connection Science*, *22*, 177–192. doi:10.1080/09540091003623797
- Perales, J. C., Catena, A., Shanks, D. R., & González, J. A. (2005). Dissociation between judgments and outcome expectancy measures in covariation learning: A signal detection theory approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1105–1120. doi:10.1037/0278-7393.31.5.1105
- Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review*, *14*, 577–596.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

- Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *Quarterly Journal of Experimental Psychology*, *60*, 433–447. doi:10.1080/17470210601002520
- Vadillo, M. A., Miller, R. R., & Matute, H. (2005). Causal and predictive-value judgments, but not predictions, are based on cue-outcome contingency. *Learning & Behavior*, *33*, 172–183.
- Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science*, *1*, 298–302. doi:10.1111/j.1467-9280.1990.tb00221.x
- Wasserman, E. A., Kao, S. F., Van Hamme, L. J., Katagari, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 207–264). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(08)60562-9