Interactive effects of the probability of the cue and the probability of the outcome on the overestimation of null contingency

Fernando Blanco\textsuperscript{a}, Helena Matute\textsuperscript{a}, and Miguel A. Vadillo\textsuperscript{b}

\textsuperscript{a} – Deusto University, Departamento de Fundamentos y Métodos de la Psicología. Bilbao, Spain

\textsuperscript{b} - University College London, Division of Psychology and Language Sciences. London, United Kingdom

Author note

Support for this research was provided by Dirección General de Investigación of the Spanish Government (Grant PSI2011-26965) and Departamento de Educación, Universidades e Investigación of the Basque Government (Grant IT363-10). We thank Samuel Hannah for his helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Fernando Blanco, University of Deusto, Departamento de Fundamentos y Métodos de la Psicología, Apdo. 1, 48080, Bilbao, Spain.
E-mail: fernandoblanco@deusto.es
Phone: +34 944133238
Abstract

Overestimations of null contingencies between a cue, C, and an outcome, O, are widely reported effects that can arise for multiple reasons. For instance, a high probability of the cue, P(C), and a high probability of the outcome, P(O), are conditions that promote the overestimation. In two experiments, the participants were asked to judge the contingency between a cue and an outcome. Both P(C) and P(O) were given extreme values (high and low) in a factorial design while maintaining the contingency between the two events at zero. Whilst we were able to observe main effects of the probability of each event, our studies showed that the cue- and outcome-density biases interacted such that a high probability of the two stimuli enhanced the overestimation beyond the effects observed when only one of the two events was frequent. This evidence can be used to better understand certain societal issues, such as belief in pseudoscience, that can be the result of overestimations of null contingencies in high P(C) or high P(O) situations.
It is a common assumption that certain cognitive abilities that human beings exhibit are grounded in the accurate perception of statistical contingency between variables. A celebrated example is causal learning and inference, which constitutes the major topic of this paper. The philosopher David Hume (1739/1888) stressed the relevance of contingency in causal inference as follows: other factors held constant, effects must be contingent on their causes. That is, the occurrence of the cause should alter the probability of the effect occurrence. Hence, this invariance can be exploited by human and nonhuman animals to infer causality from their exposure to contingency patterns. In line with this view, prominent theories and empirical reports (Rescorla, 1968; Wasserman, 1990) support the hypothesis that human and nonhuman animals are indeed sensitive to contingency manipulations and able to use contingency information as a major hint to infer causal relations. For instance, a scientist will conclude that a given drug produces a skin rash as a side effect if she observes that the rash appears often upon consumption of the drug but does not appear as frequently in the absence of the drug.

The typical paradigm aimed at studying human contingency and causal learning involves the exposure to a sequence of trials in which two stimuli, namely a cue (i.e., the potential cause, C) and an outcome (i.e., the potential effect, O), may or may not co-occur (e.g., Jenkins & Ward, 1965). The statistical contingency between these stimuli can be manipulated experimentally by determining their joint frequencies. Four combinations are possible with one binary cue and one binary outcome: either both the cue and the outcome co-occur (type a cell), or only the cue occurs (type b cell), or only the outcome occurs (type c cell), or neither the cue nor the outcome occurs (type d cell). After the training phase, participants provide a numerical judgment expressing
their estimation of the degree to which C and O were related to each other. This judgment can then be contrasted with the predictions of descriptive models and normative theories of contingency learning.

One of the most widely used normative standards for the measurement of contingency is the ΔP index (Allan, 1980; Jenkins & Ward, 1965). Considering the simplest case with only one cue and one outcome, each of which can be either present or absent, the ΔP rule amounts to the difference between the conditional probability of the outcome given that the cue is present, P(O|C), and the conditional probability of the outcome given that the cue is absent, P(O|¬C). Larger differences between the two conditional probabilities provide stronger support for the hypothesis that C and O are causally related. When ΔP equals 0, there is no statistical correlation between C and O and, thus, no apparent reason to believe that they are causally related.

A growing body of evidence demonstrates that although humans exhibit a remarkable ability to learn accurately from cue-outcome contingency patterns (Shanks & Dickinson, 1987; Wasserman, 1990), their cognitive systems are easily misled (Alloy & Abramson, 1979; López, Cobos, Caño, & Shanks, 1998). This often results in a biased judgment, such as attributing a side effect to a drug that is, in fact, innocuous. As discussed elsewhere, these biases have important implications for practical and societal issues, such as the popularity of pseumedicine (Matute, Yarritu, & Vadillo, 2011).

Among the factors that have been found to bias human contingency judgments, two are of interest for the purposes of the current report. Even when the contingency is held at zero, manipulations of the probability with which the outcome appears, P(O), bias the judgments such that higher values of P(O) lead to higher contingency
judgments. This phenomenon is sometimes referred to as the outcome-density effect (Allan & Jenkins, 1983; Allan, Siegel, & Tangen, 2005; Buehner, Cheng, & Clifford, 2003; Musca, Vadillo, Blanco, & Matute, 2010; Wasserman, Kao, Van Hamme, Katagari, & Young, 1996). In a similar vein, some studies have shown an analogous bias when the probability of the cue, P(C), is manipulated (this is known as the cue-density effect; Allan & Jenkins, 1983; Matute, et al., 2011; Perales, Catena, Shanks, & González, 2005; Vadillo, Musca, Blanco, & Matute, 2011; Wasserman et al., 1996). In these cases, the higher the probability of the cue, the higher the subjective judgment of the cue-outcome contingency.

In addition, these biases appear not only in observational situations in which participants are learning the relationship between two external events (cue and outcome) but also in assessing the relationship between their own actions and an outcome. In these latter settings, contingency judgments are also biased by the probability of the outcome (Alloy & Abramson, 1979; Hannah & Beneteau, 2009; Matute, 1995; Msetfi, Murphy, & Simpson, 2007; Msetfi, Murphy, Simpson, & Kornbrot, 2005; Vallée-Tourangeau, Murphy, & Baker, 2005) and the probability of the response (which plays the role of the cue in these paradigms; Blanco, Matute, & Vadillo, 2009, 2011, 2012; Hannah & Beneteau, 2009; Matute, 1996).

The cue- and outcome-density effects have been proposed as responsible for many real-life superstitions, illusions and irrational beliefs in which people perceive a causal relation between two events that do not covariate and are not causally related. More specifically, the combination of a high P(C) an a high P(O) has been described as the condition that promotes the stronger illusions of causality (e.g., Matute et al., 2011). Consider, for instance, the belief in bogus medicine such as homeopathy:
usually, this erroneous belief appears when treating mild diseases with high chances of spontaneous relief, i.e., high $P(O)$, and when the medicine has no side effects so that patients can self-administer the medicine very often, i.e., high $P(C)$. Nonetheless, an empirical gap renders unclear the question of how certain different combinations of $P(C)$ and $P(O)$ could affect the judgments. What would be the patients' judgments of the effectivity of a bogus medicine that is scarcely used (e.g. because it is too expensive, or hard to find) in treating a disease whose effects wear off frequently? [i.e., low $P(C)$, high $P(O)$]. Would these judgments be similar to the opposite situation in which a bogus treatment is used often but the disease rarely disappears? [i.e., high $P(C)$, low $P(O)$].

Thus, our current research question concerns the way in which the cue- and outcome-density biases combine when the actual contingency is zero. One possibility is that they simply combine additively. Then, the stronger contingency overestimation would be observed when the two probabilities are high (that is, when both the cue and the outcome take place frequently), and this is consistent with the observation of strong beliefs in the effectivity of actually ineffective treatments (Matute et al., 2011). A second possibility is that the two biases interact, even in an asymmetrical way. In fact, there are reasons to suspect that the two biases influence the contingency judgments unevenly. For instance, although the outcome-density bias seems pervasive and relatively easy to replicate, some authors have described the cue-density effect as elusive and relatively small (Blanco et al., 2011; Hannah & Beneteau, 2009; Perales & Shanks, 2007), suggesting the superiority of the former over the latter, at least under certain conditions.
Nonetheless, few studies have addressed the question of whether these two manipulations exert additive or interactive effects on contingency judgments. The few studies that have tested this specific question have typically presented the information to the participants summarized in a tabular format (White, 2004), rather than in the standard trial-by-trial procedure that is normally used in contingency learning tasks. This difference in the presentation format of the information renders unfair the comparison between the results derived from the two procedures, as extensively discussed by many researchers (e.g., Shanks, 1991; Vallée-Tourangeau, Payton, & Murphy, 2008). Focusing on the standard trial-by-trial procedure, studies reporting outcome-density effects in null contingency settings have generally kept P(C) constant at a medium level (e.g., Msetfi et al., 2005; 2007; Shanks & Dickinson, 1987). Conversely, studies reporting a cue-density effect have been conducted at a high (e.g., Blanco et al., 2009, 2011, 2012; Matute, 1996; Matute et al., 2011) or medium level of P(O) (e.g., Hannah & Beneteau, 2009; Perales et al., 2005). Therefore, it is unclear whether outcome- or cue-density effects may appear in low P(C) and low P(O) settings, respectively, in the standard contingency learning task. It could be the case that a medium/high probability of one of the two stimuli is a necessary condition to obtain a density bias by manipulating the probability of the counterpart. This would be consistent with a potential interaction of the two manipulations. This has not been tested in a standard contingency learning paradigm. The present experiments were designed to fill this empirical gap by manipulating orthogonally the P(C) and the P(O), with extreme values (high and low), in a noncontingent task with trial-by-trial presentation format. If potential interactions or asymmetries exist, they should be revealed under these conditions.
The aim of Experiment 1 was to clarify the way in which \( P(O) \) and \( P(C) \) affect contingency judgments. To do so, we manipulated both probabilities in a factorial 2x2 design while keeping the actual cue-outcome contingency set to zero, as measured by the \( \Delta P \) index. It is possible that both manipulations, \( P(C) \) and \( P(O) \), produce similar effects, showing an additive relationship when the two probabilities are high. Alternatively, they might interact such that a high, or at least a medium, level of \( P(O) \) is necessary to show a cue-density effect. This latter possibility has been implicitly assumed in studies that explored the effects of manipulating \( P(C) \) without a low \( P(O) \) condition (Blanco et al., 2009, 2011, 2012; Matute, 1996). Note that the manipulation of the two marginal probabilities in a factorial design while keeping a fixed contingency is possible in a zero contingency setting only. Experiment 1 made use of a causal scenario, which is the standard setting in contingency learning paradigms.

**Method**

*Participants and Apparatus*

One hundred and eight anonymous Internet users voluntarily participated in the experiment by accessing our virtual laboratory [http://www.labpsico.deusto.es]. The experiment was programmed in *JavaScript*, a web-based language that is interpretable by most browsers.

We did not collect demographic information of our sample in our web-based experiment. Nevertheless, and even though data collected through the Internet can indeed be noisier than data collected in the laboratory, we have previously shown that they are normally consistent with laboratory results and replicate the overestimation-of-contingency effect observed under controlled laboratory conditions (e.g., Matute,
Vadillo, Vegas, & Blanco, 2007). In any case, the basic findings of the present experiment will be replicated in the laboratory in Experiment 2.

The computer program randomly assigned participants to one of the four experimental groups. Two participants were identified as outliers because of their extreme judgments (both $|z| > 2.70$) within their respective groups and, therefore, were excluded from the analysis. The resulting sample of 106 participants was distributed across the groups as follows: 27 participants in the HighC-LowO group, 27 in the HighC-HighO group, 26 in the LowC-HighO group, and 26 in the LowC-LowO group. Group names refer to the probability (High or Low) of the occurrence of the cue, C, and the outcome, O, respectively.

**Procedure and Design**

The computer task is an adaptation of the standard allergy paradigm that has been extensively used in the literature of human contingency learning (e.g., Wasserman, 1990). In the current experiment, the participants were asked to imagine that they were medical doctors who had to elucidate whether a fictitious medicine, Batatrim, was effective in helping patients recover from a fictitious illness, the "Lindsay Syndrome". Thus, Batatrim was described as a potential cause of the patients’ recovery from the disease (see the full instructions in the Appendix).

During the training phase, participants were presented with a series of 100 medical records of fictitious patients suffering from the Lindsay Syndrome. In each trial, a message first showed whether the current patient had taken Batatrim, which played the role of the cue or potential cause, C. This information was conveyed using a picture of a pill bottle (C) or the same image crossed out in red ($\neg$C). Next, the participants predicted whether the patient was going to recover from the illness by
clicking on the corresponding button on the screen (labeled "Yes" or "No"). This predictive response was requested to ensure that the participants were reading the information on the screen, as is usually performed in this type of experiment (Vadillo, Miller, & Matute, 2005). Immediately after the response was recorded, information about the recovery was displayed. The patient either felt better (outcome-present trial, O) or did not (outcome-absent trial, ¬O), which was shown by the corresponding pictures (i.e., either a happy face or a sick face). The stimuli for the cues and the outcomes remained on the screen until the participant clicked on the button labeled "Continue" and proceeded to the next trial.

The general design of the experiment is shown in Table 1. In type a trials, both the cue and outcome occurred; in type b trials, the cue occurred but the outcome did not; in type c trials, the outcome occurred but the cue did not; finally, in type d trials, neither the cue nor the outcome occurred. The probability of taking the medicine, P(C), and the probability of the recovery, P(O), were orthogonally manipulated by varying the amount of a, b, c, and d trials during the training session. P(C) and P(O) were fixed to either of two levels, High (.80) and Low (.20), resulting in a 2x2 factorial design comprising four groups. The contingency between taking the medicine and recovering from the illness, as measured by ΔP, was held at zero for all groups.

After the training phase was completed, the participants rated the efficacy of the medicine by answering the question "To what extent do you think that the medicine Batatrim has been effective in healing the crises of the patients you have just seen?" Answers were provided by clicking on a 0-100 numerical scale, where 0 was labeled "It was not effective at all" and 100 was labeled "It was perfectly effective". A floating label appeared as soon as the participants moved their mouse over the scale,
indicating the number that was below the mouse pointer. Because the actual contingency between the medicine and the recoveries from the crises was set to zero in all groups, we assumed that a bias arose when the participant’s judgment departed significantly from zero, suggesting a failure to detect the absence of a relationship between the two stimuli.

**Results and Discussion**

The critical dependent variable of this experiment is the mean judgment of contingency after the training phase, as depicted in Figure 1. These results strongly suggest an interaction between P(C) and P(O) manipulations. Indeed, a 2x2 ANOVA yielded both a significant main effect of P(O), $F(1,102) = 67.80, p < .001, \eta^2_p = .40$, and of P(C), $F(1,102) = 4.70, p = .033, \eta^2_p = .04$, as well as the interaction between them, $F(1,102) = 4.50, p = .036, \eta^2_p = .04$. Further simple effect analyses showed that a high P(O) was a necessary condition to observe the cue-density bias. When P(O) was high, the group with a high P(C) showed higher –more biased– judgments than the Low P(C) group, $F(1, 102) = 9.20, p = .003, \eta^2_p = .08$. Consistent with the significant interaction, when P(O) was low, there was no difference between groups as a function of their P(C), $F(1, 102) < 0.01$. In contrast, the outcome-density bias was observed at both levels of P(C): when P(C) was high, $F(1, 102) = 54.65, p < .001, \eta^2_p = .35$, and when it was low, $F(1, 102) = 18.34, p < .001, \eta^2_p = .15$. The absence of a cue-density bias in the low P(O) condition suggests that the significant main effect of cue-density was entirely attributable to the high P(O) condition.

Finally, although the number of type a trials was identical in groups HighC-LowO and LowC-HighO (i.e., 16 trials), the judgments were significantly higher in the latter case, in which P(O) was high, $t(51) = 4.14, p < .001, d = 1.16$. This is also
indicative of an interaction between the two probabilities in which the outcome-density bias seems superior to the cue-density bias.

**Experiment 2**

Two concerns may be put forward against the asymmetry between P(C) and P(O) effects that was observed in Experiment 1. First, it is not possible to disentangle whether the differences between the cue-density and the outcome-density effects were, in fact, due to the role of these stimuli as cue and outcome or due to the semantic/perceptual difference between the medicine Batatrim and the recovery of the illness. Therefore, in Experiment 2, we used geometric figures rather than causes and effects, and an unfamiliar and neutral cover story that allowed the counterbalancing of the two stimuli in their role as cue and outcome. Second, the absence of differences between groups in the low P(O) condition might be due to a floor effect derived from the unidirectional response scale (from 0 to 100) that was used in Experiment 1. Participants did not have the opportunity to report a negative contingency and, consequently, may have chosen to give a rating of zero. Thus, if the participants answer on a bidirectional scale (from -100 to +100), allowing the report of a perceived negative contingency, then a cue-density bias in the low P(O) condition may be revealed. In addition, Experiment 2 was run in the laboratory, so as to make sure that the effects observed in Experiment 1 were not due to our use of Internet-based samples.

**Method**

*Participants and Apparatus*

One hundred and sixty-six students from Deusto University volunteered to participate in Experiment 2, which was conducted during four collective sessions in a
large computer room at the School of Psychology and Education. The participants were randomly assigned to one of four groups. One participant was excluded because he/she responded 'No' to the predictive question in every trial despite being assigned to the High P(O) condition, suggesting that he/she was not paying attention to the task. In addition, five participants exhibited extreme judgments (|z| > 2.70) within their groups, and thus, their data were excluded. The final sample consisted of 160 participants: 41 in the HighC-LowO group, 38 in the HighC-HighO group, 44 in the LowC-LowO group, and 37 in the LowC-HighO group.

Procedure and Design

The same design used in Experiment 1 was employed (see Table 1), but the medical cover story was replaced by a noncausal scenario (see the instructions in the Appendix) in which the participants were asked to learn the predictive relationship between two abstract stimuli (i.e., geometrical figures, namely, a triangle and a square). For roughly half of the participants, the roles of the triangle and the square as cue and outcome were interchanged. For the sake of simplicity, in the following explanation, we will only describe the procedure with the triangle in the role of the cue and the square in the role of the outcome.

On each of the 100 training trials, the occurrence of the cue was indicated by a picture of the triangle and the message "The triangle has appeared", whereas its absence was signaled by the same figure crossed out in red along with the message "The triangle has not appeared". Next, the participants were asked to predict whether the outcome (i.e., the square) was to occur or not. Upon providing the dichotomous (Yes/No) answer, feedback was displayed by showing either the picture of a square
(and the message "The square has appeared") or the square crossed out in red (and the message "The square has not appeared").

The frequency of each type of trial (a, b, c, and d) and the randomized sequence of trials were identical to those of Experiment 1. After the training phase, the participants were asked to rate the strength of the predictive relationship between the two figures by answering the question "To what extent do you think that the occurrence of the triangle was useful to predict the occurrence of the square?"

Importantly, ratings were provided on a bidirectional scale, ranging from -100 to +100. As in other experiments using bidirectional scales, the participants were instructed as follows: "A positive number means that the occurrence of the triangle predicts the occurrence of the square. A negative number means that the occurrence of the triangle predicts the absence of the square. A rating of zero means that the triangle cannot be used to predict the square". Numbers from -100 to 100 were shown as participants moved their mouse over the scale.

Results and Discussion

The critical measures are again the contingency judgments at test, shown in Figure 2. A visual inspection of this figure suggests a replication of the asymmetry that was found in Experiment 1. The P(O) manipulation seems to exert a greater impact on judgments than the P(C) manipulation. A 2x2 ANOVA, including P(C) and P(O) as factors, was conducted on the judgments. These manipulations did not yield equivalent results. Whereas a significant main effect of P(O) was observed, $F(1, 156) = 17.58, p < .001, \eta^2_p = .10$, the effect of P(C) was not significant, $F(1, 156) = 1.47, p = .23, \eta^2_p = .01$.

Furthermore, although small in size, the interaction between these two factors reached the significance level, $F(1, 156) = 4.01, p = .047, \eta^2_p = .025$. 
In addition, simple effect analyses revealed the cue-density effect in the high P(O) condition, $F(1, 156) = 4.86, p = .029, \eta^2_p = .03$, but not in the low P(O) condition, $F(1, 156) < 1$. The outcome-density effect was found in the high P(C) condition, $F(1, 156) = 19.02, p < .001, \eta^2_p = .11$, and was absent in the low P(C) condition, $F(1, 156) = 2.42, p = .12, \eta^2_p = .02$. As in Experiment 1, the contrast between the two groups with an identical number of type a trials, i.e., HighC-LowO and LowC-HighO, revealed that judgments were higher in the high P(O) setting, $t(76) = 1.97, p = .05, d = 0.45$ (although in Experiment 2 this difference was marginally significant).

Of note, the lack of a cue-density effect in the condition of low probability of the outcome was previously obtained in Experiment 1. The current replication suggests that this finding was not due to a floor effect artificially produced by the unidirectional response scale used in Experiment 1, as judgments were provided through a bidirectional scale in Experiment 2. This feature of Experiment 2 may also be at the basis of the overall small magnitude of the overestimations compared to Experiment 1: in Experiment 2, half of the screen was occupied by the positive values of the scale only, while in Experiment 1 the same space was occupied by positive and high (i.e., larger than 50) values.

In summary, the results of Experiment 2 showed that the cue- and outcome-density effects appeared prominently against a background of high P(O) and high P(C), respectively. The reported statistical interaction revealed that the biases were less likely to appear in the low P(C) and low P(O) conditions. Finally, while the main effect of outcome-density was reliable, the main effect of the P(C) manipulation failed to reach the significance level, suggesting that the impact of the latter manipulation is somewhat weaker. These conclusions align with those of Experiment 1 and support...
that the effects of manipulating P(C) and P(O) on contingency judgments are asymmetrical and interactive. Furthermore, a combined analysis of the two experiments, including the Experiment, P(C) and P(O) as factors and taking the z-scores of the judgments as the dependent variable, revealed similar results to those conducted on each experiment separately: we found a main effect of P(O), $F(1, 258) = 70.80, p < .001, \eta_p^2 = .22$, main effect of P(C), $F(1, 258) = 5.23, p = .023, \eta_p^2 = .02$, and their interaction, $F(1, 258) = 7.81, p = .006, \eta_p^2 = .03$, indicating that the cue-density bias was observable under high P(O) conditions only, $F(1, 258) = 12.61, p < .001, \eta_p^2 = .05$. In addition, a two-way P(O) x Experiment interaction showed that the main outcome-density bias was more pronounced in Experiment 1 than in Experiment 2, consistent with the abovementioned observation about the rating scale.

Several additional controls and improvements over Experiment 1 were introduced in Experiment 2. First, the semantic/perceptual features of cues and outcomes were controlled (by counterbalancing them and using a noncausal, abstract cover story involving geometrical figures). Second, a bidirectional scale was utilized to ensure that no floor effects artificially prevented the cue-density bias in the low P(O) condition. Third, whereas the Experiment 1 data were collected through the Internet, the Experiment 2 data were collected in a laboratory. Finally, Experiment 2 allows the generalization of the findings of Experiment 1 to a noncausal scenario in which there is little or no room for prior knowledge about the way in which cues and outcomes should relate to each other.

**General discussion**

The current two studies provide further evidence that human contingency judgments are easily biased by covariational manipulations, often departing from the
objective contingency values provided by normative indexes such as $\Delta P$ (see, e.g., Allan & Jenkins, 1983; Alloy & Abramson, 1979; López et al., 1998). Both the frequency of the cue and of the outcome occurrences were able to produce overestimations of contingency (i.e., cue- and outcome-density biases). Moreover, these two manipulations seem to interact. Particularly, we have shown that the overestimation of contingency is magnified when both the cue and the outcome are presented very frequently, and is weak or disappears when one of the two stimuli is presented scarcely.

In fact, as noted in the Introduction, many of the previous experimental studies that reported cue- and outcome-density biases used a high $P(O)$ condition (e.g., Blanco et al., 2011, 2012; Hannah & Beneteau, 2009; Matute, 1996; Matute et al., 2011; Perales et al., 2005) or a medium/high $P(C)$ condition (e.g., Msetfi et al., 2005, 2007), respectively. Our results complement these studies because they provide evidence for the weakening of the biases in backgrounds of low $P(O)$ and low $P(C)$.

In addition, taking together the results of Experiment 1 and 2, we can suggest that the outcome-density bias is less vulnerable to this interactive combination compared to the cue-density bias, because the main effect of the $P(C)$ manipulation was obtained only in Experiment 1 and, as discussed above, it can be completely attributed to the High $P(O)$ condition, which boosted the cue-density bias. By contrast, the main effect of the $P(O)$ manipulation was readily obtained and exhibited larger effects in the two experiments, although it was, again, stronger in a background of high $P(C)$.

The demonstration of an interactive effect of the $P(C)$ and $P(O)$ manipulations is not only an empirical question that deserves being studied, but it also entails some
implications for the theoretical debate. Models of contingency learning should be able to accommodate the present results. Much has been discussed on the mechanism at the basis of the biases in contingency judgments. Some have argued (e.g., Allan, 1980) that these biases are led by contiguity, or the number of cue-outcome co-occurrences (i.e., type \(a\) cells). Then, because the manipulations of \(P(C)\) and \(P(O)\) are mutually independent, both should contribute evenly to the probability of co-occurrences. In other words, if the biases were produced by the number of type \(a\) cells, one would expect an additive, rather than interactive, combination of the two manipulations, \(P(C)\) and \(P(O)\). In fact, our finding of interactive, asymmetrical effects of these manipulations on the judgments suggests that their contributions to the overestimation of contingency are uneven. In addition, the overestimation cannot be fully explained by the number of type \(a\) cells because, when we compared the two groups in which the number of this type of trials were identical, we found stronger illusions when \(P(O)\) was higher, in both experiments.

A more refined account of these biases does not rely on type \(a\) trials solely. Indeed, many researchers argue that people weight differently each piece of information when judging contingencies (e.g., Perales & Shanks, 2007; White, 2004). Interestingly, the ranking of cell weights that has been usually reported in previous research (i.e., \(a > b > c > d\); see Kao & Wasserman, 1993; Wasserman, Dorner, & Kao, 1990; Wasserman et al., 1996) could be used to predict our current results. This ranking can be incorporated in several contingency learning models (e.g., the EI model, Perales & Shanks, 2007; the Rescorla-Wagner model, Rescorla & Wagner, 1972). Although our current experiments are, admittedly, unable to discriminate between specific theories conclusively, they do suggest that any model that does not
incorporate, either explicitly (e.g., White, 2004) or implicitly (e.g., Rescorla & Wagner, 1972), unequal cell weights, will face problems to offer a complete account of the cue- and outcome-density effects.

Importantly, the relevance of cue- and outcome-density biases is not restricted to the academic debate. As discussed elsewhere, both biases have important implications for the study of everyday causal illusions and superstitions (e.g., Matute et al., 2011) as well as stereotype formation (e.g., Murphy, Schmeer, Vallée-Tourangeau, Mondragón, & Hilton, 2011). The knowledge obtained from studying overestimations of null contingency should, therefore, be of great value in the development of strategies to reduce the impact of irrationality in modern society (see, e.g., Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Lilienfeld, Ammirati, & Lanfield, 2009; Matute et al., 2011).
References


Table 1.

*Design Summary of Experiments 1 and 2*

<table>
<thead>
<tr>
<th>Group</th>
<th>P(C)</th>
<th>P(O)</th>
<th>Cell frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>LowC-LowO</td>
<td>Low (.20)</td>
<td>Low (.20)</td>
<td>4 a, 16 b, 16 c, 64 d</td>
</tr>
<tr>
<td>LowC-HighO</td>
<td>Low (.20)</td>
<td>High (.80)</td>
<td>16 a, 4 b, 64 c, 16 d</td>
</tr>
<tr>
<td>HighC-LowO</td>
<td>High (.80)</td>
<td>Low (.20)</td>
<td>16 a, 64 b, 4 c, 16 d</td>
</tr>
<tr>
<td>HighC-HighO</td>
<td>High (.80)</td>
<td>High (.80)</td>
<td>64 a, 16 b, 16 c, 4 d</td>
</tr>
</tbody>
</table>

Group names refer to the two probabilities that were manipulated, P(C) and P(O), whose values are shown in the second and third columns, respectively. The names of the cells (i.e., a, b, c, and d) refer to the standard 2x2 contingency matrix (see main text).
Figure 1. Mean judgments in Experiment 1, as a function of the probability of the cue, P(C), and the probability of the outcome, P(O). Error bars depict 95% confidence intervals for the mean.
Figure 2. Mean judgments in Experiment 2 as a function of the probability of the cue, P(C), and the probability of the outcome, P(O). Error bars depict 95% confidence intervals for the mean.
Appendix

Full instructions of Experiment 1 (Translated from the original in Spanish)

Imagine that you are a doctor who works at the Hospital Emergency Department. You are specialist in a rare and dangerous disease called "the Lindsay Syndrome", which must be treated quickly in the emergency room.

Crisis induced by this illness may be stopped immediately by using a medicine called "Batatrim", but this medicine is still in its testing stage, therefore its reliability has not been yet proven. In addition, you should know that it produces several side effects, which can be severe, in every patient who takes it, so it cannot always be used.

When you click on the button "Start", you will be presented with a series of medical records of patients suffering from the Lindsay Syndrome. In each record, you will first know whether or not the patient was given Batatrim, and you will be asked to indicate whether you think that the patient will recover from the crisis.

Then, you will know whether or not the patient actually recovered from the crisis. Try to find out the extent to which Batatrim is actually effective. Once you see a number of patients, you will be asked a few questions.

Full instructions of Experiment 2 (Translated from the original in Spanish)

In this experiment, two geometrical figures (a square and a triangle) will appear on your computer screen.

In each trial, you will first know whether or not the square [the triangle] appears (when it is does not appear, it will be crossed out in red). You will be asked to predict if
the other figure, the triangle [the square], will appear, by clicking on the corresponding button.

Once you make your prediction, you will know if the triangle [square] actually appeared. Your goal is to try to find out the extent to which the square [triangle] can help you predict when the triangle [square] will appear.