Cognitive Biases in Human Causal Learning

Antonio Maldonado, Andrés Catena, José César Perales, and Antonio Cándido
Universidad de Granada

The main aim of this work was to look for cognitive biases in human inference of causal relationships in order to emphasize the psychological processes that modulate causal learning. From the effect of the judgment frequency, this work presents subsequent research on cue competition (overshadowing, blocking, and super-conditioning effects) showing that the strength of prior beliefs and new evidence based upon covariation computation contributes additively to predict causal judgments, whereas the balance between the reliability of both, beliefs and covariation knowledge, modulates their relative weight. New findings also showed "inattentional blindness" for negative or preventative causal relationships but not for positive or generative ones, due to failure in codifying and retrieving the necessary information for its computation. Overall results unveil the need of three hierarchical levels of a whole architecture for human causal learning: the lower one, responsible for codifying the events during the task; the second one, computing the retrieved information; finally, the higher level, integrating this evidence with previous causal knowledge. In summary, whereas current theoretical frameworks on causal inference and decision-making usually focused either on causal beliefs or covariation information, the present work shows how both are required to be able to explain the complexity and flexibility involved in human causal learning.
Keywords: causal learning, causal inference, cognitive bias, belief

El objetivo de este trabajo fue la búsqueda de sesgos cognitivos en la inferencia de relaciones causales para descubrir qué procesos psicológicos modulan el aprendizaje causal. A partir del efecto de la frecuencia de juicio, este trabajo presenta investigación consecuente sobre competición entre claves (ensombrecimiento, bloqueo o súper-condicionamiento) para demostrar cómo la fuerza de las creencias previas y la evidencia sobre la covariación de cada causa contribuyen aditivamente en los juicios causales y en la toma de decisiones, siendo su fuerza relativa modulada por la fiabilidad otorgada a cada tipo de información. Nuevos datos muestran también la incapacidad para detectar relaciones causales incidentales preventivas, pero no generativas. Esta "ceguera inatencional" parece deberse a un fallo en la codificación o recuperación de la información. Todos estos datos revelan que una arquitectura cognitiva del aprendizaje causal debe basarse en tres niveles. El primer nivel sería responsable de la codificación de los eventos en cada ensayo. El segundo nivel computaría la nueva evidencia a partir de la información recibida del primer nivel. En el tercer nivel, el individuo debe interpretar e integrar toda esta información con su conocimiento causal previo. En suma, los modelos sobre juicios de causalidad y toma de decisiones normalmente se han centrado en el efecto exclusivo de las "creencias y conocimiento causal" o de la "experiencia directa y covariación" entre causas y efectos. Este trabajo demuestra que ambos tipos de información se requieren e interactúan cuando se trata de explicar la complejidad y flexibilidad que implica el aprendizaje y la inferencia de relaciones causales en humanos.
Palabras clave: aprendizaje causal, inferencia causal, sesgo cognitivo, creencia

Causal learning has been an emergent topic of research during the last decades (De Houwer & Becker, 2002; Perales & Catena, 2006) and is one of the most basic psychological components of human cognition, motivation, emotion, and behavior. In many everyday situations, we make causal attributions either consciously or automatically. When we have a headache, the flu, or stomach acidity, we usually take a specific pill under the assumption or causal belief that this pill will make us well. How have we arrived at this causal belief? It seems that at least two factors determine our action. First, we probably have been told by our doctor or by some friend that these pills are good for these illnesses. However, we also need to confirm these beliefs by our own experience, for example, avoiding side-effects and being sure that this pill also has the desired effect when we take it. Therefore, causal attributions seem to depend on "previous beliefs," but also on the "direct experience" we received when we tested them. It is important to acknowledge that scientific reasoning shares some features of causal reasoning. Scientists always have a previous "theory" (i.e., "beliefs") when they undertake new experiments. Whether or not they change this theory will depend on the data obtained from the experiments, confirming or disproving the theoretical predictions. However, the scenario is not that easy and we are not always "rational." In this work, we will try to determine the psychological mechanisms underlying causal inferences to illustrate how "previous beliefs or theories" and "direct empirical evidence" interact to generate causal attributions, but also how they give rise to specific cognitive biases when we finally make a causal judgment or we develop a new theory.

In the search for the psychological processes involved in human causal judgments, the main question is: What do people need to make a causal attribution? A growing body of data have questioned that any independent single mechanism can explain human causal detection and attribution (Catena, Maldonado, Megías, & Fresse, 2002; Fugelsang & Thompson, 2003) and the chief aim of this work was to demonstrate how a cognitive architecture of causal learning should have at least three hierarchical levels (Perales, Catena, & Maldonado, 2002).

It seems obvious that the lower mechanism should be able to accurately detect and codify the relevant stimuli and events to make a causal inference. This mechanism is assumed to codify frequencies of cause-effect type of events, probably in an episodic temporal order: (a) cause + effect, (b) cause alone, (c) effect alone, and (d) none (Maldonado, Herrera, Jiménez, Perales, & Catena, 2006). The second mechanism should compute statistical cause-effect associations, based upon the information retrieved from the first one. Most theories and research on causal learning have focused on describing this mechanism. The debate about whether it is associative or statistical has been extensive, and no final resolution has been reached to date (De Houwer & Becker, 2002; Perales & Catena,

2006), because predictions are often equivalent in the asymptote. In any case, it is clear that causal links do not correspond to mere correlations but to mechanisms giving rise to those correlations (Cheng, 1997; Perales et al., 2002, among others). Then, once completed the "computation process," the third mechanism should be able to integrate the information retrieved from the computational mechanism (i.e., the direct empirical evidence) with previous information or beliefs, based upon the reasoners' causal mental model. Such a mental construct encodes assumptions about the causal status of the events (causes vs. effects or causal directionality) and about hypothetical causal relationships in this causal scenario (independent vs. chained causes, common causes, or effects, and so on) and it allows reasoners to distinguish genuine from spurious causes and to look for the boundary conditions under which they can properly estimate the power of a given cause to produce an effect (Fugelsang & Thompson, 2003; Lien & Cheng, 2000; Perales, Catena, & Maldonado, 2004).

According to the above proposals, causal judgments depend on the action of these three mechanisms. The main objective of this work was firstly, to show the existence of specific cognitive biases at each processing level unaccountable by any single mechanism—be it associative or statistical. The final aim was to propose a cognitive architecture underlying causal learning, able to explain at which processing stage these specific biases happen and how they modulate the causal inference process, in order to accurately predict human causal judgments (see Figure 3). The main procedure always implies a "dissociation technique," showing how the same contingency gives rise to different causal judgments as a function of subjective or psychological factors, such as attention, previous beliefs, or memory loads, which will further demonstrate the need of cognitive mechanisms beyond the mere computation of covariation in order to be able to explain human causal inference.

Judgment Frequency: Insufficiency of any Single Mechanism Model in Causal Learning

The frequency of judgment effect (Catena, Maldonado, & Cándido, 1998; Catena, Perales, & Maldonado, 2004), an effect that has been replicated in different settings (see also Collins & Shanks, 2002; Matute, Vegas, & Marez, 2002), showed how a higher judgment frequency induced not only less accurate contingency judgments—as shown by previous social research on jurors' decision accuracy (Pennington & Hastie, 1992)—but also a differential influence of the last type of trials upon subsequent causal judgments, according to its usual weight (trials type "a" > "b" $\lessgtr$ "c" > "d", see Kao & Wasserman, 1993; Maldonado, Catena, Cándido, & Garcia, 1999).

Figure 1 shows the usual design and findings when a positive contingency is at work. Two groups of participants were exposed to same contingency ($\Delta P = 0.5$), everything being identical except the frequency of judgment (after block of 8 trials or after every single trial, Figure 1, top panel). The results show how people are quite accurate when estimating the causal strength after a sufficient number of trials (block of 8 trials); but also, how they are inaccurate and very sensitive to the last information, when they make the same estimations after every trial (Figure 1, bottom panel).

The original belief revision model (Catena et al., 1998) explaining this effect, proposed an anchoring and adjustment algorithm based upon the action of two serial mechanisms during the process of learning the causal relationships between a single causal cue and an effect (see also Appendix).
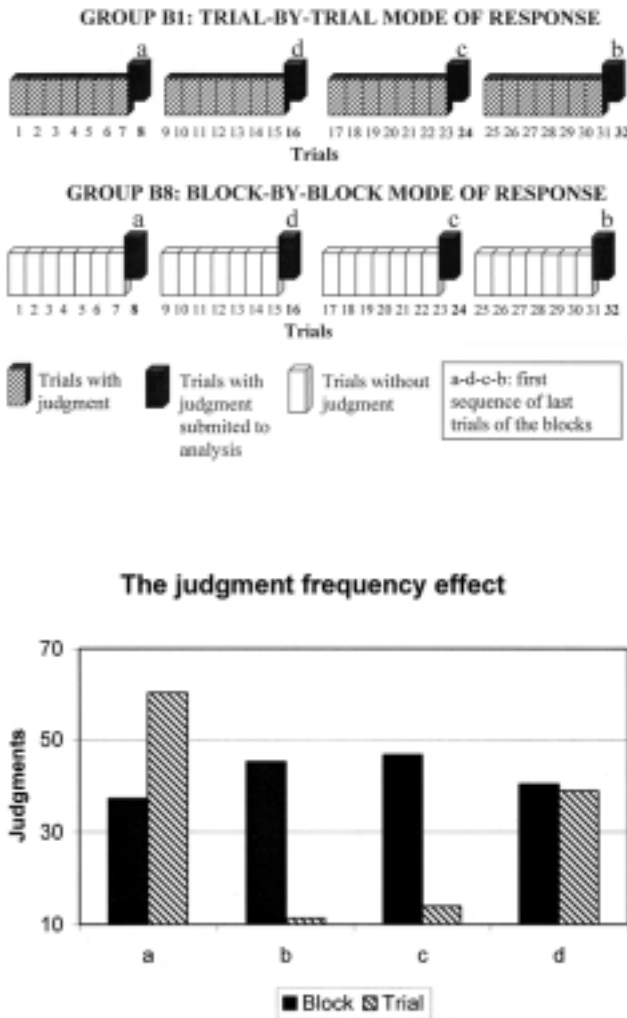

The judgment frequency effect

Figure 1. The judgment frequency effect: design and results.
Note. A, B, C, and D stand for the type of trial (Unpublished data, see text for details).

The first mechanism computed the new evidence and is a key feature of the belief revision model that frequency information should be computed as a weighted ΔD:

$$NewEvidence = \frac{w_1\, a + w_2\, b + w_3\, c + w_4\, d}{a + b + c + d} \quad (1)$$

where $\underline{a}$, $\underline{b}$, $\underline{c}$, and $\underline{d}$ are the number of each type of trial, and w_j is the weight of each trial type, and always $a > b \lessgtr c > d$, as shown by independent research (Kao & Wasserman, 1993; Maldonado et al., 1999). It is important to point out that later research has also shown how this statistical rule better predicts human causal judgments than any other statistical or associative rule (Perales & Shanks, 2003). At this level of processing (i.e., when computing the cause-effect contingency), recent work has also shown the existence of specific biases, such as how people give more weight to confirming or positive (i.e., "a" type of trials) than to negative or disconfirming evidence (i.e., "b" and "c" type), and how they show frequency biases and asymmetrical detection when estimating the contingency between a given cause and its effect (Maldonado et al., 1999; Perales & Shanks, 2003).

The second information-integration mechanism controlled judgment updating at trial *n*. This judgment was a function of the discrepancy between the new evidence and the causal judgment at trial *n - k,* according to which:

$$J^n = J^{n-k} + \beta(NewEvidence - J^{n-k}) \quad (2)$$

where J stands for causal judgment at trial *n* (or *n - k*), *k* is the number of trials since the last judgment, β is the revision rate parameter, and new evidence refers to the amount of new information presented since the last judgment was made, that is, between trial *n - k* and trial *n* in a trial-by-trial procedure. This second mechanism is believed to hold operative control over the computation mechanism, resetting it after each judgment (J^n), which defines the amount of information that will be summarized. In addition, this mechanism should, in theory, be able to combine the information from different sources—such as instructions and previous beliefs—to be integrated with the information received by the statistical mechanism, via the ß parameter. The next section will try to develop its implications for causal learning.

The Role of Causal Beliefs and Reliability on Causal Cue Competition

The belief revision model was initially developed to explain causal judgments between one cue and one effect, in which case, the computation is always unconditional. Recently however, De Houwer and Becker (2002) pointed out that this model needed further developments to be able to deal with cue competition effects and also to show the

role of causal beliefs upon causal judgments. The tendency of most of the research on causal learning to date has been to look for the influence of each one of these two factors, beliefs or contingency, as competing explanations (Fugelsang & Thompson, 2003). Recent results have shown not only the need of both sources of information, but also their relative weight to derive causal judgments in scenarios with multiple potential causal candidates.

In a causal scenario with multiple cues, people should compute the new evidence about each one of the cues and the composite of each causal candidate before making a causal judgment. Recent research has shown that the default assumption is an independent causal mental model and the computing rule could be a variation of the noisy-or rule:

$$J(E/A \cap B) = J(E/A) + J(E/B) \qquad (3)$$

where the strength of the causal relationships (causal judgment, J) of the conjunction of both A and B causes to produce the effect (E) will be an additive function of the causal power of each one of the two A and B causes alone. In any case, it is important to remember that computation of contingency is always based upon a weighted ΔD, as proposed above.

Different studies have pointed out that the process of estimating the causal power of any single candidate in such complex scenarios also depends on at least two sources of information: the actual covariation, as assumed by any single-mechanism model, but also the causal mental model, including previous beliefs about the nature of the causes (Waldman, 2000). However, it seems important to acknowledge that different kinds of scenarios are possible in a situation where two possible causal candidates predict the same effect, also giving rise to different phenomena. Imagine, for example, that we have a given disease or physical reaction such as dermatitis, after eating two different meals or after taking two different medicines (for example, an antihistamine or vitamin and an antibiotic).

The first causal scenario, when people have only information about the composite of the two causal candidates, is a situation analogous to an "overshadowing" paradigm in animal conditioning. This occurs if we are taking two medicines at the same time and we have an allergic reaction. In this case, without any previous knowledge about the nature of these medicines, people seem to first estimate the contingency of the conjunction, based upon a "weighted ΔD." Then, given the high ambiguity of the situation, when asking for the causal power of each causal candidate alone, the simpler heuristic is to divide this causal power between the two possible causes, as proposed in Equation 3 (see Table 1, the NO beliefs—overshadowing group). These results are identical to those usually observed in overshadowing paradigms, either in human causal learning or animal conditioning (Waldman, 2000). It seems obvious that previous beliefs about the propensity of each cause to produce the effect should have a strong influence upon the final causal judgments: the bigger the propensity, the higher the judgment. This is exactly what happened when inducing the belief that participants were taking a vitamin and an antibiotic. In this case, almost all causal power of the conjunction was attributed to the cause the most likely to produce the effect (the antibiotic) and very few to the less likely one (the vitamin, see Table 1, the PB [previous beliefs]—overshadowing group). According to the belief revision model, this effect happens via the second mechanism, by the modification of the anchoring judgment (J^n), showing the role and influence of previous beliefs.

The next causal scenarios include information not only about the conjunction, but also about one of the possible candidates, which is now an analogue of blocking—when this single cause is also a generative one—or super-conditioning—when the opposite occurs, and the single cue is a preventative cause of the effect. In this case, the strategy was to estimate both, the contingency of the conjunction and the contingency of the known cause alone, and then derive the causal power of the other one. Table 1 displays the results of the BRM model predictions, which are very similar to the participants' judgments (see Catena,

Table 1

Previous Beliefs and Cue Competition Effects: Results of the Simulation of the Belief Revision Model (BRM)

Cause Group	BRM-Predicted Causal Judgments		
	A	B	AB
NO-Overshadowing	37	36	56
PB-Overshadowing	51	21	48
NO-Blocking	31	44	64
PB-Blocking	36	39	56
NO-Super-Conditioning	37	–21	42
PB-Super-Conditioning	31	–35	56

Note. PB = Previous Beliefs. PB and NO stand for induction or non-induction of previous beliefs, respectively. A, B, and AB stand for the causes; A and B are the cause with the higher or lower propensity, respectively, to produce the effect on the PB groups. See text and Appendix for more details (based on data from Catena et al., 2007).

Maldonado, Perales, Candido, & Herrera, 2007, for details). This shows how participants were quite accurate in their judgments about the similar generative causal power of the conjunction, independently of the causal scenario and also rather independent of previous beliefs. However, when the single cue was also generative, more casual power was assigned to this cause than to the other one; but when the single cue was preventative, as accurately detected by the negative judgments (predicted only by a weighted ΔD), all causal power was assigned to the other cue (Table 1) which shows a blocking-like and a super-conditioning effect, respectively. It is important to note that covariation information was now sufficient to know the causal power of each one of the two causes in both conditions. Therefore, these effects were rather independent of the previous beliefs about the nature of the cues.

A recent investigation about the influence of previous beliefs and previous experience and their interaction both on decision-making and causal judgments have clarified the relative value of covariation and belief information on causal learning (García-Retamero et al., 2007). Participants had to decide who of two patients had a higher degree of allergic dermatitis (the outcome) in the light of different causal cues. The causal scenario now included four possible causal candidates looking for the differential effect of actual contingency evidence or previous beliefs about the nature of the cues. To this end and via instructions, participants were told that a possible causal mechanism could explain the influence of two of the possible cues, the causal ones, whereas the other two cues have no such mechanism underlying their action, the neutral ones. For example, the patients could have been bitten by a spider or ingested a certain prescription drug that produced dermatitis as a side-effect (when the cues were causal) or they had touched a white butterfly or taken a vitamin without any known side-effect (when the cues were neutral). At the same time, one causal and one neutral cue had a higher contingency (.70) than the other neutral or causal one (.30). After training, they were asked to estimate the strength of the causal relationships between each cause and the effect (dermatitis).

The results of a group without any previous experience showed that previous beliefs had indeed a stronger influence than actual contingency upon causal judgments (see Table 2

for the predicted judgments from the simulation of the BRM model, which are very similar to the causal judgments of the participants; see García-Retamero et al., 2007, for details) and also upon decisions, being the results similar in both measures. This effect was even more evident when participants were given pre-training in the presence of the causal cues alone (Table 2, Causal pre-training group). In both groups, causal cues were estimated higher than neutral cues, independently of their actual contingency, bearing out models proposing that causal knowledge would be the main factor guiding whether covariation is interpreted as causation or as a mere spurious relationship. According to the belief revision model, these effects seem to occur via the information-integration mechanism, and they can be predicted by the modification of the anchoring judgment (J^0), as also occurred in the overshadowing paradigm explained above (see Table 1).

The most important and novel results occurred as a function of previous training with the neutral cues alone. Participants were now able to estimate the causal power of each cue, based on actual contingencies rather than on previous beliefs (see Table 2, Neutral pre-training group). Given the previous experience with the neutral cues, the influence of the actual cues' contingency was enhanced, whereas previous beliefs had lesser influence both on decisions and judgments. Thus, participants were able to accurately detect the actual contingency between cues and outcomes independently of their previous beliefs about the nature of the cues, which also modifies the pattern of causal judgments in this group regarding the other two groups. These results are less easily accounted for in cognitive models emphasizing the role of beliefs upon causal judgments (Fugelsang & Thompson, 2003). However, according to the belief revision model, this effect also occurs via the second mechanism, but in this case, by the modification of the reliability (the β parameter) of the information (see Appendix).

Overall results and the simulations of the belief revision model suggest that causal judgments depend on the integration of causal beliefs and empirical evidence. The tendency of most of the models on causal inference to date has been to look for the influence of beliefs or direct experience based on covariation, as competing explanations (Fugelsang & Thompson, 2003; Lien & Cheng, 2000; Sloman & Hagmayer 2006). On the contrary, the belief

Table 2

The Effects of Previous Beliefs, Previous Training, and Actual Evidence as Predicted by the Belief Revision Model (BRM)

Cause Group	BRM-Predicted Causal Judgments			
	Causal High	Causal Low	Neutral High	Neutral Low
Naïve	76	65	20	9
Causal Pre-training	73	55	20	9
Neutral Pre-training	59	23	64	29

Note. See text and Appendix for more details (based on data from García-Retamero et al., 2007).

revision model proposes that previous beliefs knowledge does not work as an absolute filter of further covariation knowledge—letting through only information that is compatible with previous beliefs—but as a framework to interpret covariation information. Such information will be considered in the light and strength of previous beliefs—J^{n-k}—and its capacity to change previous beliefs depends on its reliability—β (Perales, Catena, Maldonado, & Cándido, 2007).

When new evidence is perceived as inconclusive or less reliable, previous causal beliefs impose a certain model that makes this evidence interpretable. As a consequence, the reliability of the new information is considered very low, as occurs in the naïve and causal-pre-training groups. From the belief revision model, the higher influence of previous beliefs over covariation information was predicted by increasing its initial value (J^{n-k}) and at the same time, decreasing the β parameter. Thus it is easy to predict and understand the "confirmation bias" usually found in scientific and causal reasoning, as scientists easily accept results confirming their initials predictions, being reluctant to change their theories when they find disconfirming results (Fugelsang, Stein, Green, & Dunbar, 2004).

However, when causal claims can be derived from the empirical evidence, which implies an increase in their reliability (an increment of the β parameter), the influence of the evidence increases and it can even modify the influence of the causal beliefs when they are in conflict. Thus, it is possible to predict and understand what happened in the neutral-pretraining group. Only by increasing the reliability (β) of the empirical evidence, was it also possible to predict the influence of such previous training with valid neutral cues. This is what probably happens when scientists replicate results contrary to their predictions, which finally lead them to change their theories (Fugelsang et al., 2004). These results suggest that causal and covariation knowledge share some representational basis allowing the modification of their relative influence on causal judgments, and this is another key feature of the belief revision model.

In short, most studies have favored either previous knowledge or covariation based on direct experience as the key piece to understand human causal judgment. The present results however, point out that both types of information are required and emphasize the need of new models and new research to entirely explain the complexity and flexibility involved in causal inference processes.

The Role of Attentional Resources in Causal Relationship Detection

A new step in our understanding of human causal learning comes from the knowledge of our processing system. Our limited capacity to process information from our environment is well known. Accordingly, since the pioneering work of Mackintosh (1975), the attentional mechanism has been incorporated into models of causal learning (Kruschke, 2001). Recent research has shown how the attentional mechanism modulates causal learning (Maldonado et al., 2006). As before, the paradigm involved the presentation of two cues with the same contingency, However, participants were induced to attend to one of them (the attentional cue) by instructions and task demands— they were asked to estimate its causal relationships with the effect every eight trials—while another cue appeared in some of the trials (the incidental cue).

Figure 2 shows judgments to the attended as well as to the incidental cues. It is evident that these causal judgments were quite accurate and sensitive to the objective contingencies when the cue was attended, although they seem more accurate for positive than for negatives contingencies, bearing out previous research on the asymmetrical detection of positive and negative contingencies (Maldonado et al., 1999, 2006). The most intriguing results were that, although participants were able to accurately detect positive or generative causal relationships between an incidental cause and its effect, they were "blind" and unable to detect similar causal relationships when they were negative or preventative.
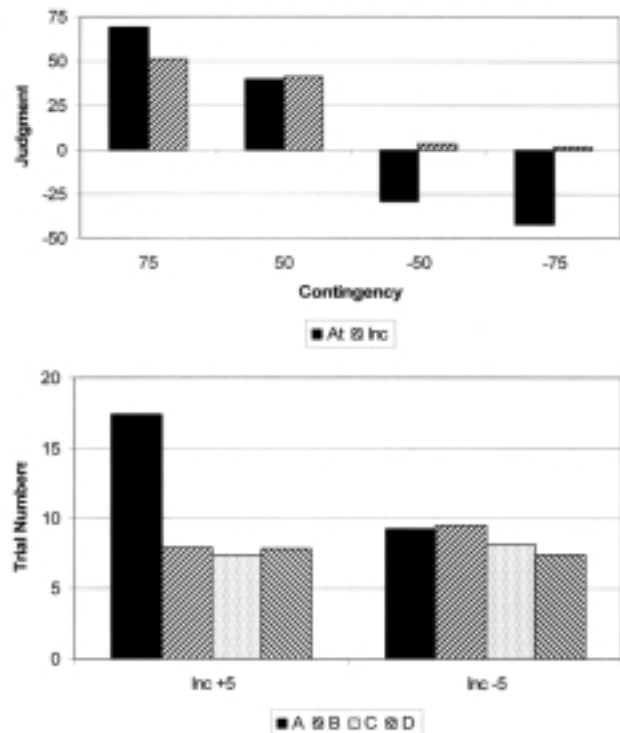

Figure 2. The "inattentional blindness" effect on causal learning.
 Note. At = Attended cues; Inc = Incidental cues. In the bottom panel, +5 and –5 represent a contingency of +50 and –50, respectively (adapted from Maldonado et al., 2006).

This lack of detection of incidental negative contingencies could be due either to a contingency computation failure (independently of the type of rule used by the computing mechanism) or to a failure in coding or retrieving the information necessary for such computation. Subsequent results when asking participants to remember the number of each type of trials after a positive and a negative experience point out to the latter factor as responsible for the effect. The estimation of frequencies differed from positive and negative "incidental" cues (see Figure 2, bottom panel), whereas it was always similarly accurate when the cues were "attended." More important, when the cue was "incidental," participants accurately remembered that the number of "a" trials was higher than "b" or "c" type trials when the contingency was positive, allowing its accurate computation, despite the incorrect recall of the number of "d" type of trials (Figure 2, bottom panel). However, they were unable to remember any difference in the number of each type of trials when the contingency was negative (Figure 2, bottom panel), preventing the correct computation of the relationships between the incidental cause and its effect, as shown by their inaccurate causal judgment (Figure 2, top panel). Our findings reveal the existence of a particular case of "inattentional blindness" for the detection of negative relationships in causal learning, rather than for the cues themselves (Maldonado et al., 2006). These results suggest that inattentional blindness for unexpected or incidental events depends on the simultaneous presence of other —attended—events as well as their relationships with the same output.

In summary, the effect described in this section points to the central role of attention allocation strategies in causal learning and adds new evidence to the growing body of data showing that research needs to be expanded beyond covariation and associative or statistical rules. Our results reveal the role of attentional resources in coding the information necessary for causal contingency computation and, indirectly, the importance of summoning up a whole cognitive architecture for causal learning.

Conclusions

According to the evidence presented above, such a cognitive architecture should have at least three hierarchical levels (see Figure 3). The lowest level—the "detection mechanism"—is responsible for coding the cues and their combinations in the trials presented during the task. At this point, attention allocation does not seem to play an important role in determining the representation of present cues, but it clearly affects the efficiency with which absent cues are represented. The second level is the "computing mechanism," which is constrained by the information received from the lower level. This mechanism could either be rule-based or associative, and, in fact, it could be argued that humans possess both types of mechanisms (Sloman & Hagmayer,

2006). If we assume the default mechanism is an associative one, we would also need to assume that whether or not an absent cue is coded depends on whether or not attentional resources are focused on that cue. On the other hand, if the proposed mechanism is rule-based and, thus, dependent on trial type frequencies, the ability to retrieve these frequencies would depend on the ability to code them. Finally, at the third level, contingency information needs to be interpreted and integrated with previous knowledge. This previous knowledge includes previous beliefs updated in accordance with the new evidence, and knowledge about the causal structures in which the newly learnt contingency is to be integrated (Perales et al., 2007, among others).
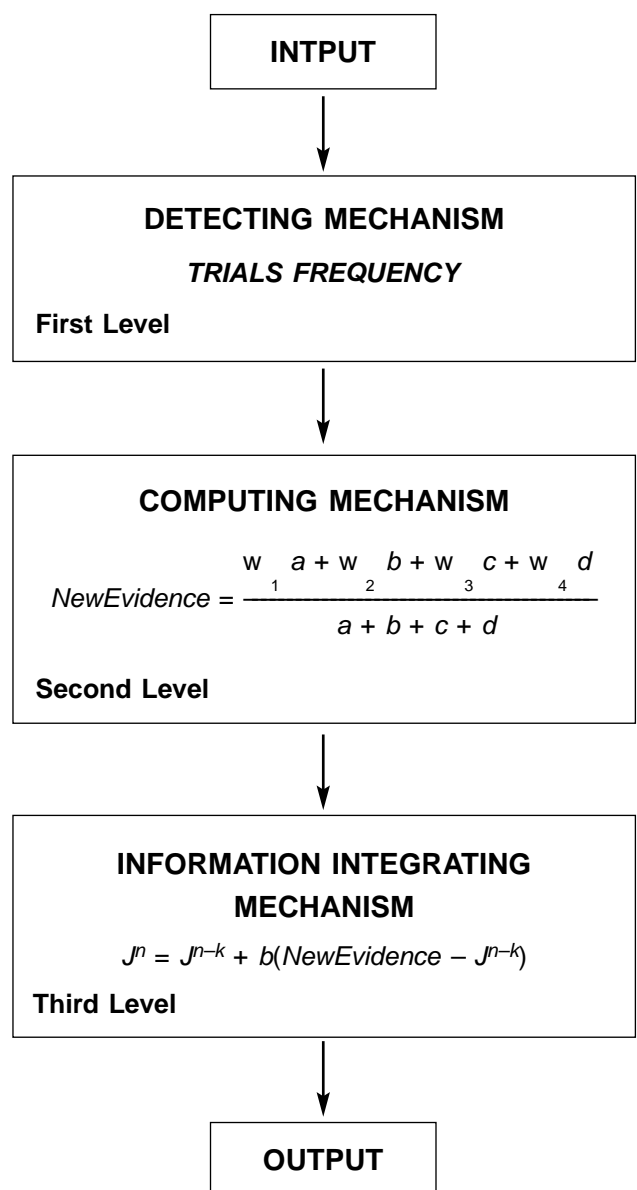
INTPUT

↓

DETECTING MECHANISM

TRIALS FREQUENCY

First Level

↓

COMPUTING MECHANISM

$$NewEvidence = \frac{w_1 a + w_2 b + w_3 c + w_4 d}{a + b + c + d}$$

Second Level

↓

INFORMATION INTEGRATING MECHANISM

$$J^n = J^{n-k} + b(NewEvidence - J^{n-k})$$

Third Level

↓

OUTPUT

Figure 3. The cognitive architecture underlying the belief revision model (BRM).

Summing up, the main point of this work is that mechanism-based beliefs and covariation-based knowledge can portray similar information about causal efficiency, independently of their origins. This commonality allows reasoners to combine them to make integrative causal judgments. Prior evidence has shown that this integrative process requires a significant amount of attentional resources. Finally, causal efficiency knowledge of mechanism-based and in covariation-based information share a common representation and can be assessed in two dimensions—reliability and strength (or magnitude). Although usually, the reliability of the information on which our prior belief is based indicates the existence of an efficient causal link, reliability is not necessarily related to the existence of a known mechanism, but can also be drawn from contingency information, if such information is based on a large enough sample of reliable observations. From this perspective, the belief revision model (see Figure 3) proposes a cognitively coherent and parsimonious set of mechanisms and algorithms to account for causal learning and the present work was aimed at opening new lines of research to better understand the complexity and flexibility of human causal learning and why people are not always objective when estimating causal relationships.

References

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.

Catena, A., Maldonado, A, Megias, J., & Frese, B. (2002). Judgement frequency, belief revision and serial processing of causal information. *Quarterly Journal of Experimental Psychology*, 55B, 267-281.

Catena, A., Perales, J.C., & Maldonado, A. (2004). Previous noncontingency and frequency effects in generative and preventative causal learning. *Psicológica, 25,* 67-85.

Catena, A., Maldonado, A., Perales, J.C., Candido, A., & Herrera, A. (2007). *Previous beliefs and cue predictive value interactions in covariation-based causal induction*. Manuscript submitted for publication.

Cheng, P.W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104, 367-405.

Collins, D.J., & Shanks, D.R (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition, 30,* 1138-1147.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental*, 55B, 289-310.

Fugelsang, J.A., Stein, C.B., Green, A.E., & Dunbar, K.N. (2004). Theory and data interactions in the scientific mind: Evidence from molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology*, 58, 86-95.

Fugelsang, J.A., & Thompson, V.A. (2003). A dual-process model of belief and evidence interactions in causal reasoning. *Memory & Cognition*, 31, 800-815.

García-Retamero, R., Maldonado, A., Catena, A., Hoffrage, U., Herrera, A., & Candido, A. (2007). *The influence of causal beliefs and empirical evidence on decision making and causal attribution*. Manuscript submitted for publication.

Kao, S.F., & Wasserman, E.A. (1993). Assessment of information integration of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1363-1386.

Kruschke, J. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45,* 812-863.

Lien, Y., & Cheng, P.W. (2000). Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology*, 40, 87-137.

Mackintosh, N. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276-298.

Maldonado, A, Catena, A., Cándido, A., & Garcia, I. (1999). The belief revision model: Asymmetrical effects of noncontingency on human covariation learning. *Learning and Behavior, 27,* 168-180.

Maldonado, A., Herrera, A., Jiménez, I., Perales, J.C., & Catena, A. (2006). Inattentional blindness for negative relationships in human causal learning. *Quarterly Journal of Experimental*, 59, 457-470.

Matute, H., Vegas, S., & Marez, P. (2002). Flexible use of recent information in predictive and causal judgments. *Journal of Experimental Psychology: Learning, Memory and cognition, 28,* 714-725.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology, 62,* 189-206.

Perales, J.C., & Catena, A. (2006). Human causal induction: A glimpse at the whole picture. *The European Journal of Cognitive Psychology, 18*, 277-320.

Perales, J.C., Catena, A., & Maldonado, A. (2002). Aprendizaje de relaciones de causalidad: hacia un análisis integral del aprendizaje desde una perspectiva computacional. *Cognitiva, 14,* 15-41.

Perales, J.C., Catena, A., & Maldonado, A. (2004). Inferring non-observed correlations from causal scenarios: The role of causal knowledge. *Learning and Motivation, 35,* 115-135.

Perales, J.C., Catena, A., Maldonado, A., & Cándido, A. (2007). The role of mechanism and covariation information in causal belief updating. *Cognition (*available online).

Perales, J.C., & Shanks, D. (2003). Normative and descriptive accounts of the influence of power and contingency on causal judgment. *Quarterly Journal of Experimental Psychology*, 56A, 977-1007.

Sloman, S.A., & Hagmayer, Y. (2006). The causal psycho-logic of choice. *Trends in Cognitive Science, 10,* 407-412.

Waldmann, M.R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 53-76.

Appendix

The belief revision model predictions of the frequency of judgment effect (Table 3) were obtained by running each simulation sixteen times, and the trials were the same sequences as those used in the experiment (see Catena et al., 1998, for details). According to model restrictions, trial weights were $w_a = 1$, $w_b = -.7$, $w_c = -.7$, and $w_d = .6$ for this simulation, as well as for the effect of previous belief and pre-training (Tables 2 and 3), being slightly modified, $w_a = 1$, $w_b = -.6$, $w_c = -.6$, and $w_d = .5$, for cue-competition effects (Table 1).

Table 3
The Belief Revision Model (BRM) Predictions of the Frequency of Judgment Effect

Trial Type Group	BRM-Predicted Causal Judgments			
	A	B	C	D
Low Frequency: block	46	46	46	46
High Frequency: trial	57	14	14	49

Note. A, B, C and D stand for the type of trial (See text and Figure 1 for details; based on data from Catena et al., 1998).

In the frequency of judgment effect (Table 3), the revision parameter was fixed at $\beta = 0.2$ and $\beta = 0.9$ for HF- (high frequency) and LF- (low frequency) conditions respectively, according to the assumed differential reliability of the new evidence in each case. In the cue-competition effects (Table 1), β was fixed at .50, whereas J_{n-1} was allowed to vary as a free parameter to capture the initial value of previous beliefs. It is noteworthy that the anchoring judgment (J^{n-k}) was close to zero (+ .03) when no prior information was provided (NO-groups), but when previous beliefs were induced (PB- groups), it was positive for A (+ .33), the cue favored by the cover story, and negative for B (–.26), the cue less likely to produce the effect, whereas for the conjunction of AB, it was also positive (+ .28).

Finally, regarding the effect of previous beliefs and pre-training (Table 2), the β parameter was fixed at .30 for all conditions at the beginning of training. The influence of causal beliefs was coded as the anchoring judgments (J^{n-k}), with an initial value of .80 and .00 for the causal and neutral cues, respectively. After pre-training with a given cue, the strength of the previous beliefs (J^{n-k}) at the beginning of the subsequent phase was the predicted strength of the model, taking into account the actual cue-outcome covariations during pre-training. However, at the beginning of the second phase of the neutral-pretraining group, β was increased to .80 and the initial J^{n-k} value of the causal cues was reduced to 0.30, under the assumption that this pre-training should lead to increased attention and reliability of the empirical evidence, and attenuated influence of previous causal beliefs. It is important to note that comparing predictions and results (see, Catena et al., 1998, 2007, and García-Retamero et al., 2007, for details), the overall pattern of causal judgments was fully predicted by the belief revision model.