Psicológica (2004), 25, 67-85.

Judgment frequency effects in generative and preventative causal learning

Andrés Catena*[*], José C. Perales**, and Antonio Maldonado*

*Universidad de Granada, **University College London

The frequency of judgment effect is a special case of Response Mode effect in human covariation and causal learning. Judgment adjustment -to ΔP-, depends on the trial type preceding that judgment, but that effect is restricted to situations in which participants are asked to make their judgments with a high frequency. Two experiments further demonstrated the reliability and the generality of this effect in positive and negative causal learning tasks. Experiment 1 yielded similar judgment frequency effects with a higher positive contingency (ΔP= 0.71) and a larger block size (n=16) than in previous studies. Experiment 2 showed that judgment frequency also modulates the detection of negative contingency (ΔP= -0.5), as far as judgment accuracy was shown to be a function of the type of trial just preceding that judgment in the high frequency group. Associative and statistical models of covariation learning cannot easily explain these results without incorporating relevant post-hoc assumptions. These findings add new-evidence to the growing body of data showing that human causal learning depends on the action of several mechanisms, as proposed by the Belief Revision Model.

The ability to detect causal relationships between environmental events is a major component of adaptive behaviour. Learning that one event is the cause of another is a basic psychological function, given the causal texture of our world (Tolman & Brunswick, 1935). For that reason, how people detect casual relationships has been a central topic in human learning research in the last two decades (see De Houwer and Beckers, 2002, for a review).

Causal learning allows humans and other organisms to know that two events are connected by some kind of link or mechanism, in such a way that the presence or the absence of the cause is consistently followed by the presence or the absence of the effect. Although not all covariation relationships are causal in nature, it can be affirmed that causality reveals itself by means of covariation. *Generative causes* have the power to produce an effect, which implies that the presence of the cause is consistently followed by

[*] Correspondence: Andrés Catena. Departamento de Psicología Experimental. Universidad de Granada. Facultad de Psicología. Campus de Cartuja s/n. 18071 Spain. e-mail: acatena@ugr.es Fax: 34-58-246239

the effect. For example, eating seafood makes some people feel sick, and smoking causes lung cancer. In other words, given a population of instances there is a certain positive correlation between the presence of the cause and the presence of the effect. On the other hand, *preventative causes* have the power to prevent certain outcomes to occur. In consequence, preventative causes are consistently followed by the absence of the outcomes they prevent. For example, people that have been injected a vaccine against flu have a lesser probability to suffer that disease than people who are not vaccinated. Or, in covariation terms, in a general population of instances the preventative cause and the outcome it prevents are inversely related.

In the simplest type of causal learning task, participants are asked to estimate the degree of covariation between two dichotomous variables whose values are given by the presence or absence of two discrete events, usually called cue (the event that appears in the first place in each trial) and *outcome* (the event that is presented in the second place). Combining the presence/absence values of both variables yields four trial types: in a type trials, both the cue and the outcome are present; in b type trials, only the cue appears; in c type trials, only the outcome appears, and, finally, in d type trials, neither the cue nor the outcome appear. After a series of trials, subjects are asked to estimate the strength of the correlation or causal relationship between the cue and the outcome. The objective relationship between cue and outcome, in relation to which the accuracy of judgments is assessed, is usually estimated by mean of ΔP. This statistic is defined as the difference between the probability of the outcome given that the cue has been presented P(O/C) and the probability of the outcome given that the cue has not been presented P(O/~C), and it can be estimated from the frequencies of the different types of trials that are presented during the task, according to the following equation:

$$\Delta P = P(O/C) - P(O/\tilde{C}) = \frac{a}{a+b} - \frac{c}{c+d} \quad (1)$$

where a, b, c and d stand for the frequencies of the different types of trial presented during the task.

Initially, two kinds of models emerged to explain causal and covariation learning. Statistical models –also called rule-based models- state that the mechanism underlying covariation and causal learning is a rule or heuristic, directly computed by the learner from the observed frequencies of trial types or conditioned probabilities, and whose output is directly mapped onto contingency / causality judgments. Although different covariation rules have been proposed, the most frequently cited one is ΔP. In general, it has been shown that human causal judgments are rather adjusted to programmed contingencies (Allan, 1993), which led to propose that humans act in this type of tasks as intuitive statisticians (Cheng and Novick, 1992, Prices and Yates, 1995).

The procedural similarities between some conditioning paradigms and standard causal learning tasks, and the similar results usually found by using both procedures (see Miller & Matute, 1996, Shanks & Dickinson, 1987) led some authors to propose the Rescorla-Wagner's (1972) associative rule as an alternative explanation of human causality learning (for example, Gluck and Bower, 1988). The associative mechanism is usually conceptualised as a single distributed two-layer network. According to the general rule, the increment of the link weight (Δw) between an input unit i (the node representing the cue) and an output unit j (the node representing the outcome) in a given trial is a function of the discrepancy between the expected and the current output in that trial:

$$\Delta w_{i,j}^{n} = k(\beta - \sum_{i,j} w_{i,j}^{n-1}) \quad (2)$$

where, $w_{i,j}$ is the matrix of weights from the input layer to the output layer, β represents the learning asymptote and, k is a learning rate parameter, normally related to the perceptual salience of the cue and the outcome.

According to associative models, the strength of the associative link between the cue and the outcome is directly mapped onto a causal judgment, denying or ignoring the possible existence of any further inductive or reasoning process. The importance of these models was increased by their success at accounting for several effects in causal and contingency judgment tasks that had been previously found in animal conditioning preparations, such as acquisition functions (Shanks, 1987), overshadowing, and cue competition (Chapman and Robbins, 1990, Miller & Matute, 1996 Prices & Yates, 1995 for reviews). These effects initially showed that an associative mechanism could account for covariation judgments and causal learning better than a statistical one. However, further modifications of both associative and rule-based models made their predictions virtually undistinguishable.

However, other experimental procedures, such as the manipulation of the way in which the information is provided (for example, summarized information in contingency Tables) raised doubts about the adequacy of associative models to fully explain human covariation and causal detection. Summarized presentations cannot activate an associative mechanism (Catena, Maldonado, López-Megías & Fresse, 2002, Price and Yates, 1995), which lead some authors to propose the existence of two independent mechanisms (Shanks, 1991) to account from the results obtained with summarized and trial-by-trial presentations. In accord with this proposal, judgments based on summarized presentations would rely on a statistical computation, whereas information presented in a trial-by-trial base would activate the associative mechanism. In fact, associative models appear to be designed to explain the causality/covariation judgments and learning only exclusively in conditioning-like (trial-by-trial) procedures.

In the last decade, findings such as retrospective revaluation, causal directionality effects, and judgment frequency effects (see De Houwer & Beckers, 2002) have added new evidence to the growing body of data showing that human causal learning does not depend exclusively on the operation of a single associative or statistical rule, but on the integrated operations of several hierarchically related mechanisms. Some of these effects are still controversial, and the discussion is now focused on the procedural conditions that make them appear. In the case of causal directionality effects (see Perales, Catena, and Maldonado, 2003), the demonstration that their appearance depend on factors like cognitive load, task demands, or inter-individual differences seems to reveal that causal learning mechanisms are more complex than considered to date (DeHouwer and Beckers, 2002).

Results like those previously mentioned make the whole corpus of data extremely complex, and somewhat contradictory. For that reason, some theorists have proposed that causal learning arises from the interaction between basic learning –data driven- mechanisms (either rule based or associative) and high-order (cognitively driven) mechanisms. Higher inductive mechanisms (Perales et al., 2003), information-integration mechanisms (Catena, Maldonado, and Cándido, 1998), and decisional strategies (Collins and Shanks, 2002) have been proposed to mediate the translation from 'raw' contingency to causal knowledge (Cheng, 1997, Waldman, 2000, Waldman and Martignon, 1998).

Judgment frequency effects represented a defy to single-mechanism models. Catena, Maldonado, and Cándido (1998) demonstrated that causal judgment accuracy -assessed by its adjustment to DP- depends on the frequency of judgment (see also Hastie and Pennington, 1995, Matute, Vegas, and De Marez, 2002, Vila, 2000, for recent replications of this effect). Moreover, it was also shown that judgments in high frequency conditions depends on the type of trial just preceding each judgment. In other words, when subjects are asked to make a judgment after each trial (maximum frequency), estimations made after an *a type* trial are higher than those made after a *b* or *c type* trial. Judgments after *d type* trials show a higher variability, but tend to be located in some point of the scale between a and b/c. This differential effect of the type of trial on participants' estimations was not initially predicted either by any statistical rule or any model based on the action of the Rescorla and Wagner rule (1972, see Catena et al, 1998, and Catena et al., 2002 and the general discussion in the present work for simulations of these models).

In order to provide an account of these two effects, Catena et al. (1998) suggested an anchoring-and-adjustment mechanism (Hogarth and Einhorn, 1992), based on the action of two serial algorithms. The first one is proposed to control judgment actualisation at trial n, being the judgment a function of the discrepancy between *New-Evidence* (the information on contingency accumulated since the last judgment) and the judgment at trial *n-k*, according to the following rule:

$$J^{n} = J^{n-k} + \beta(NewEvidence - J^{n-k}) \quad (3)$$

where J stands for the judgment at trial n (or n-k), k represents the number of trials presented since the last judgment, ß is a revision rate parameter, and New-Evidence refers to the amount of information presented between trial n-k and trial n.

The second algorithm is New-Evidence, which is computed by a statistical mechanism, according to the following equation (weighted ΔD):

$$NewEvidence = \frac{w_a a + w_b b + w_c c + w_d d}{a + b + c + d} \quad (4)$$

where a, b, c, and d stand for the frequencies of each type of trial, and w_i are the weights of each trial type, taking values restricted to the rule $w_a > w_b >= w_c > w_d$.

According to this model, the influence of the trial just preceding a judgment will increase as the frequency of judgment increases (see Catena et al, 1998, and Catena, Maldonado, López-Megías y Frese, 2002 for simulations), as it has been repeatedly shown in recent research. However, to the date, the effect of the last trial upon the subsequent judgment when increasing its frequency has been only shown with null or moderate positive contingencies (ΔP = 0 or 0.5) and with a little number of trials in each block (n=8). The following experiments were carried out to increase the generality of such effects (Experiment 1) and, specially, to study how judgment frequency modulated inhibitory (negative) causal learning in a preventative causal detection task (Experiment 2).

EXPERIMENT 1

The previously described Last-trial Effect[1] (the effect of the trial immediately preceding a judgment in high-frequency conditions) has been only demonstrated using blocks with a maximum length of 8 trials. This issue is relevant because the demonstration of the effect requires at least one trial of each of the four types of trials per block. Then, the positive contingency was limited to a maximum of 0.5, as measured by ΔP, being the block 8-trial long or shorter. Consequently, it might be argued that this effect is a special case of

[1] In order to simplify the notation, Henceforth, we will refer to the interaction between the type of trial preceding a judgment and the frequency of judgment as the Last-trial Effect. We will reserve the term Frequency of Judgment Effect to denote the main effect of the manipulation of the judgment frequency on the global adjustment of judgments to actual (objective) contingencies.

bias on the detection of null and low contingencies. If this were the case, the generality of the effect could be questioned and its theoretical relevance severely damaged. Therefore, in this experiment we replicate the last-trial effect using a higher positive contingency ($\Delta P = 0.71$) and a larger series of trials in each block (n=16).

METHOD

Participants. Thirty-two undergraduate University of Granada students from introductory courses of Psychology voluntarily participated in this experiment for course credits.

Apparatus and stimuli. All stimuli were presented on a high-resolution SVGA colour monitor controlled by a Pentium PC computer. Stimuli were verbal labels in different colours standing for a fictitious symptom (Atrofia) and a fictitious syndrome (Montero). In each trial, the participants received information about a fictitious patient. Thus, the combination of the presence/absence values of both events produced four patient types (trials): *a* type patients suffered both the symptom and the disease; *b* type patients presented the symptom alone; *c* type patient presented the disease alone, and *d* type patients presented neither the symptoms nor the disease. Covariation judgments (to what degree do the symptom and the disease tend to appear together?) were made by using a graded scale, ranged from -100 (Maximum Negative Relationship) to +100 (Maximum Positive Relationship), being 0 the value labelled as No-Relationship.

Design and procedure. Participants were randomly assigned to two groups. In the Low Frequency group (LF group), they were asked to make a judgment after every 16 trial-long block. In the High Frequency group (HF group), they made a judgment after each trial. 10 type *a* trials, 1 type *b* trial, 1 type *c* trial, and 4 type *d* trials were included in each training block. Within-block contingency was always fixed at 0.71, as measured by ΔP. The trial sequence in each block was randomised for each participant of each group (but yoked across groups), with the only restriction that the last trial of each block was fixed according to an incomplete counterbalancing sequence. Thus, the ultimate trial of each of the four blocks followed the sequence [a, b, d, c] in the first sub-group, and [b, c, a, d], [c, d, b, a], and [d, a, c, b], in the other three.

According to this procedure, two different designs can be defined over the same data. In both cases, the frequency of judgment (HF vs LF) was a between-subject factor. The within-subject factor was, in the first type of design, the order in which participant made their judgments across the task (after trials 16, 32, 48 and 64, being the type of trial counterbalanced). Hereafter, we will refer to this factorisation as the *Block Design*. Alternatively, we can reorder the judgments with regard to the last trial type (a, b, c and d) in each block, being the actual judgment order counterbalanced. We will refer to this design as the *Last-trial type Design*. The Block design is aimed at

ascertaining whether acquisition effects are observed across the task, as it tracks judgments down in the same order they are made, after each block of 16 trials. The Last-trial type Design counterbalances the potential effect of acquisition, and focuses on the effect of the trial type preceding a judgment on that same judgment. It is important to note that, when we focus on acquisition effects, the effect of the trial preceding a judgment is counterbalanced; and, when we focus on the effect of the trial preceding a judgment, the possible acquisition effect remains counterbalanced and unobservable. Therefore the effect of the two factors can not be studied simultaneously.

Before the onset of the task, each participant was seated at a distance of 60 cm from the monitor, where instructions and stimuli were presented. They were instructed to imaging being members of a research team interested in studying the relationship between a symptom and a disease. Once the instructions were presented and adequately understood, the information about each patient was presented in a trial-by-trial way. In each trial, a label indicating the presence or absence of the symptom was presented, and 500 ms later a second label indicated the presence or absence of the disease. The second label (or its absence) was presented besides the first label (or its absence) on the same screen during 1500 ms, which means that the whole trial lasted for 2 s. During the pre-training stage, eight practice trials (two of each type) with a different symptom and a different syndrome (Distonia and Cajal) were presented, followed by the response scale. After the experimenter was sure the participant had understood the task, he or she was told that the experimental task was going to begin. Their task was to estimate the strength of the relationship between the symptom and the disease after a given number of patients (16 in the LF group, and 1 in the HF group), whenever the response scale appeared. To this end, participants were instructed to move the cursor of the response scale to the point that best indicated their estimate of the strength of the relationship between the symptom and the disease, considering all the patients seen up to that moment. At the end of the session, participants were thanked for their participation, and informed about the aim of the study (on demand).

RESULTS AND DISCUSSION

The main results of this experiment are displayed in Table 1. Two different ANOVA's were performed according to the two experimental designs previously described. In the *Block Design*, the 2x4 ANOVA (Group x Block of trials) yielded only a significant main effect of Group, $\underline{F}$ (1 ,30) = 5.48, $\underline{MSE}$=2209, p<.05. Judgments were higher and more adjusted to the programmed contingency in the LF than in the HF group ($\underline{M}$=72.75, and 52.29, respectively, see Table 1). This result demonstrated a main Frequency of Judgment Effect (see note 1), as judgment adjustment to the objective contingency, as defined by ΔP, was an inverse function of the frequency of judgment. As far as there were no block effects, this design did not reveal any clear acquisition effect.

Nevertheless, the most interesting finding was revealed by the analysis of the *Last-trial Design* (see Table 1). The 2x4 ANOVA (Group x Last Trial Type in each block) yielded the same Group effect found in the previous design. However, in this second analysis the interaction between Groups and the Last trial type was also significant, $F(3,90)=3,68$, $MSE=585.09$, $p<.05$. This is the effect we have denoted as the Last-trial Effect (although, in fact, it is an interaction between the two manipulated factors (see footnote 1).

Table 1. Judgments for High (HF) and Low (LF) frequency groups according to the last trial of each block in experiments 1 and 2 and simulations of Markman associative model and the belief revision model (BRM).

Condition	**Type of Trial**					
	EXPERIMENT 1					
JUDGMENTS	a	b	c	d	mean	
LF +.70	70	77	70	73	73	
HF +.70	71	44	39	59	53	
MARKMAN	a	b	c	d	mean	RMSE
LF +.70	34	36	21	31	31	48.9
HF +.70	37	26	29	41	33	29.6
BRM	a	b	c	d	mean	RMSE
LF +.70	62	62	62	62	62	12.7
HF +.70	68	51	51	67	59	8.9
	EXPERIMENT 2					
JUDGMENTS	a	b	c	d	mean	
LF -.50	-27	-29	-30	-23	-27	
HF -.50	-13	-35	-41	-43	-33	
MARKMAN	a	b	c	d	mean	RMSE
LF -.50	-25	-21	-28	-14	-22	7.5
HF -.50	-8	-25	-30	-14	-20	19.2
BRM	a	b	c	d	mean	RMSE
LF -.50	-25	-25	-25	-25	-25	3.8
HF -.50	-6	-36	-36	-17	-24	15.9

Note: Each simulation was run sixteen times, trials sequences being the same used in the experiments and mean values rounded to the nearest integer. In the BRM model, the trials weights were the same of previous studies (Catena et al., 1998, 2002): $w_a = 100$; $w_b = -70$; $w_c = -70$; $w_d = 60$, being $\alpha = 0.2$ and $\alpha = 0.9$ for HF and LF condition, respectively. Learning rates in the Markman model were: HF $\beta = 0.2$; LF $\beta = 0.001$. for the cue and $\beta = 0.1$ for the context in both conditions. RMSE was the root mean square error of prediction.

Post-hoc simple effects analyses of this interaction showed significant differences among judgments only in the High Frequency group, $F(3,45)=3.23$, $MSE= 1083.82$, $p<.05$. LSD post-hoc tests showed that judgments after an a type trial in the High Frequency group were significantly higher than after any other type of trials. No other differences were significant. Secondly, post-hoc between-groups analyses revealed also that LF group judgments were higher than those of the HF group after b and c type trials, $F(1,30)= 6.77$, $MSE=1311.81$, $p<.05$, and $F(1,30)=4.61$, $MSE=1734.19$, $p<.05$, respectively. Judgments after *a* and *d* trials did not differ across groups.

In summary, causal judgments were rather adjusted to the programmed contingency and independent of the type of last trial in the low frequency group. In the high frequency group, judgments were significantly less accurate, especially after b and c type trials, being these judgments more sensible to the trial type preceding the judgment than to the actual contingency between events. These results replicated previous findings on the detection of lower positive causal relationships (Catena et al, 1998; Catena et al., 2002), and showed that the Last-trial effect is a general phenomenon. The next experiment aimed to study how the frequency of judgment modulated also the detection of negative causal relationships.

EXPERIMENT 2

In this experiment, two different procedures of negative contingency learning were used. In the single negative contingency procedure, one fictitious symptom (X) was negatively correlated with a fictitious disease. The second procedure was intended to be similar to the one commonly used to study inhibitory conditioning, as associative models maintain that inhibitory learning requires the simultaneous presence of an excitatory stimulus in the same context (Chapman and Robbins, 1990, Williams, 1996). Accordingly, two fictitious symptoms (X and Y) and the same fictitious disease were used in the inhibitory procedure. In order to establish the negative contingency level of the X cue, the four types of trial in this condition were: YX-syndrome (type a); YX-no syndrome (type b); Y-syndrome (type c) and Y-no syndrome (type d). Note that in this second inhibitory learning procedure Y is present in all trials, and, therefore, it plays the role of a contextual cue that has been made salient. The participants' task was to estimate the strength of the negative relationships established between symptom X and the disease. In both procedures, a high and a low frequency condition were used (HF and LF groups, as described in the first experiment) to study the effect of the judgment frequency on the detection of negative causal relationships.

METHOD

Subjects, stimuli and apparatus. Sixty-four undergraduate students from introductory courses of Psychology participated for course credits. Stimuli and apparatus were the same used in Experiment 1, with the addition of a new symptom (Y, Disforia) in the inhibitory procedure, as explained above.

Design and procedure. Instructions, practice trials, number of blocks, counterbalancing sequences, trial sequences and judgment recording, remained as described in Experiment 1. In addition, each participant was randomly assigned to one of the next four groups. Participants in the Low Frequency- Single Negative Contingency (LF-X) and Low Frequency-Inhibitory Learning (LF-YX) groups were asked to make a judgment after every block of sixteen trials. In the High Frequency-Negative Contingency (HF-X) and High Frequency-Inhibitory Learning (HF-YX) groups, they were asked to make a judgment after each trial. In the LF-YX and HF-YX groups the X symptom was negatively related with the disease, whereas Y-Outcome contingency was not calculable, as Y was presented in every trial during the task and the outcome appeared in half of these trials. Each block consisted of 2 *a* type trials, 6 *b* type trials , 6 *c* type trials , and 2 *d* type trials (in the two YX groups, the trials were : 2 YX-disease, 6 YX -no disease, 6 Y-disease, and 2 Y-no disease). Objective contingency for X was fixed at -0.5, as measured by ΔP (2/8-6/8) in all conditions. Only judgments on the target cue X were required.

RESULTS AND DISCUSSION

As in the previous experiment, two ANOVAs were performed according to the type of experimental design previously described. In the *Block Design*, the 2x2x4 ANOVA (Type of negative contingency procedure x Group x Block of trials) did not reveal any significant effect. Judgments were equally poorly adjusted to the objective contingency in both conditions (means of -27 and -33, in the LF and HF conditions, respectively). This result replicated previous findings (Maldonado, Catena, Cándido, and García, 1999) suggesting that the detection of negative contingencies is more difficult than the positive equivalent ones, when only one cause and one effect appear during the task[2].

---

[2] This poor detection of negative contingencies has been previously reported in several inter-event causal and contingency learning studies (see Maldonado et al., 1999). For example, in the study on conditioned inhibition by Chapman and Robbins (1999, Experiment 2) judgments about the inhibitory cue and the negatively correlated control cue (scores around -.70 and -.50 respectively) were less adjusted than those about the positive one (score higher than +90), although both cues were perfect (deterministic) inhibitory predictors of the outcome. However, there are also reports in which the adjustment of negative contingency judgments (or other measures of covariation learning) are as good as positive ones. In general terms, this happens in instrumental tasks where learners are asked to judge the relationship between a response (for example, pressing a key) and a given

According to the *Last-trial Design*, a second 2x2x4 ANOVA (Type of negative contingency procedure x Frequency of judgment x Type of last trial) was performed in order to replicate the Last-trial Effect found in the previous experiment. This analysis yielded only a significant effect for the first order interaction between the frequency of judgment and the type of trial, $F(3,180=2.68$, MSE=2049.01, $p<.05$. Simple effects analysis of this interaction showed, as in the previous experiment, significant differences among judgments only in the High Frequency groups, $F(3,93)=3.03$, MSE=1987.13, $p<.05$. Post hoc LSD tests revealed that judgments after *a type* trials were significantly higher than those after *b, c,* or *d* type of trials. No other difference was significant.

These results revealed, firstly, an absence of significant differences between the single negative contingency and the inhibitory procedures in the detection of causal negative relationships, contrarily to what has been reported by Chapman and Robbins (1990). It is possible that these differences could be attributed to procedural differences. We used the explicit symptom (Y) to simulate the possible role of the context and, consequently, it was present in every trial during the inhibitory procedure. Therefore, if the context accounts for the conditioned inhibition effect in a negative contingency task, it was expected that there were no differences between the two procedures used in this experiment. In any case, this remains an open question for future research (see Van Hamme and Wasserman, 1994, and Perales and Shanks, 2003, for discussions on the potential role of instructions in determining the role of the context in causal learning).

Secondly, there were no significant between-groups differences in the detection of negative contingencies, when using a Block Design, being judgment adjustment to the objective contingency poorer than in the positive contingency task. In other words, there was not a main Judgment Frequency effect equivalent to the one found in positive contingencies. These results are in agreement with previous findings about the asymmetrical detection of positive and negative contingencies in causal detection tasks (Maldonado et al, 1999).

Finally, the most important finding was the dependence of judgments on the last trial type in HF groups, but not in LF ones. This result replicates the Last-trial effect found in positive contingency learning tasks. However it is also important to note that in this case, the effect was exclusively due to the influence of *a* type trials, whereas in positive contingency tasks it was due to *b* and *c* ones (see LSD post-hoc comparisons). We will discuss the possible implications of this difference in the next section.

---

outcome (Chatlosh, Neunaber & Wasserman, 1985, Wasserman, 1990, Shanks & Dickinson, 1987, see Shanks, 1993, for a review). Our tentative hypothesis to account for this discrepancy is that these procedures makes the absence of the outcome more salient, in such a way that learners codify such absence as a positive effect.

GENERAL DISCUSSION

Two main new findings have been shown in this work. First, a Frequency of Judgment (HF vs LF) effect was found in generative causal learning tasks –in agreement with previous reports-, but not in preventative ones. In the detection of positive causal relationships, a higher response frequency led to a poorer adjustment of mean judgments to objective contingencies, whereas judgments in the low frequency condition were highly accurate (Experiment 1). This differential Judgment Frequency Effect disappeared when estimating negative causal relationships, as high- and low-frequency judgments were equally poorly adjusted to the objective contingency (Experiment 2). This set of results suggests an asymmetrical effect of positive and negative contingency detection, as also showed by previous research (Maldonado et al., 1999). This asymmetry is probably due to a general increase of task difficulty and cognitive load in the detection of negative relationships (Maldonado, Herrera, Jiménez, Perales & Catena, 2003).

The second and most noteworthy finding was the Last-trial effect, that is, the fact that judgment frequency modulated the effect of the type of trial just preceding a judgment in both preventative and generative HF causal learning tasks (according to the general restriction $w_a > w_b >= w_c > w_d$). In terms of adjustment to programmed contingencies, judgments were less accurate after *b and c* type trials than after *a and d* ones, when estimating high positive contingencies (Experiment 1). This effect was similar to the previously found with moderate positive contingencies (Catena et al., 1998). Conversely, judgments were less accurate after a type trials than after *b, c and d* ones, when estimating negative contingencies (Experiment 2), being these results much similar to those previously found in the detection of null contingencies. The sign of the influence of *d type* trials appears to be more similar to that of *a type* trials in positive and the null contingency cases, and to *b and c type* trials in the negative one, probably due to the greater difficulty for the subjects to adequately process the influence of these type of trials when estimating causal relationships (see Maldonado et al, 1999). However, it has been consistenly demonstrated that the interpretation of *d* type trials is not always clear for naïve learners, in such a way that some of them evaluate them as confirmatory, others as disconfirmatory, and even others do not take them into account when evaluating causal strength (White, 1998)

This pattern of results raises doubts about the adequacy of single-mechanism models to fully account for causal learning. On the one hand, the main problem for statistical (rule-based) models is that they do not consider any direct of interactive effect of response frequency manipulations, given that subjects are expected to compute the global contingency (independently of the rule they use) every time they are asked to make a judgment (see Catena et al., 2002, for simulations of such potential models).

On the other hand, traditional associative models as Rescorla and Wagner's (1972) are unable to explain why the effect of the trial type preceding a judgment differs in HF and LF groups (see simulations in Catena

et al, 1998). However, more recent models have suggested modifications of the Rescorla-Wagner rule to better account for learning on a cue in trials in which that cue is absent[3]. According to these proposals, learning about an absent cue can occur, if that cue is predicted by another one or expected in that context, for example, by influence of instructions (Markman, 1989, Van Hamme & Wasserman, 1994). Both Markman's and Van Hamme-Wasserman's models allow to negatively codify an absent-expected cue, just assigning a negative value to the salience parameter of that cue. The mathematical formulation of both models is in all other aspects equivalent to the LMS rule proposed by Rescorla and Wagner (1972). The only difference between the two modified algorithms is that Markman's formulation updates the weights of associative links using the input in the current trial, whereas van Hamme and Wasserman's model use the network weights of the previous trial for updating. This is important because, due to this difference, the assymptotic predictions of the two models differ. Moreover, in the case of Van Hamme and Wasserman's model, when more than one cue are included in the simulations (as it happens when considering the context) the obtained predictions can easily go beyond the theoretical assymptote, which means that the model is mathematically less usefull (Perales & Shanks, 2003).

For that reason, we have used the associative algorithm proposed by Markman to simulate the results of the two experiments (see Perales and Shanks, 2003, for a simplified procedure to obtain asymptotic predictions from Van Hamme and Wasserman's model). This algorithm can explain both the absence of differences between the single and the conditioning-like procedure (reported here) and its presence (reported in other works, Chapman and Robbins, 1990), depending on the assumptions made about how the influence of the context is to be taken into account. The context can be

---

[3] The R-W rules accounts for inhibitory learning in an indirect way. According to that model, the context recruits associative strength in those trials in which the target cue is absent, competing with it for the total amount of available strength. Therefore, in inhibitory learning procedures, if the context accumulates associative strength in those trials in which the target cue is absent, and, in a subsequent trial in which the target cue is present the outcome does not appear, the difference between the expected and the actual value of the outcome is negative, thus yielding a decrement of the associative strength attributed to the target cue. In the case of modifications of the RW rule like Markman's or Van-Hamme and Wasserman's, the mechanism works in a different manner: If a cue is expected in a given trial (due to its association to the context of other cues) it is *negatively* activated in that trial (the parameter associated to its salience is assigned a negative value). Thus, if in a given trial that cue is absent (expected) and the outcome is present, the associative strength attributed to that cue decreases. The main difference between R-W and its modified versions is the moment in which b and d trials affect causal learning. According to R-W, *b* and *d* trials do not have an immediate effect, but in subsequent trials in which the target cue is presented and compete with the context for the limited amount of associative strength. In other words, no learning about a cue occurs in trials in which that cue is absent. However, according to modified versions of R-W, the inhibitory effect on the associative strength attributed to a cue, in those trials in which the target cue is absent (*b* and *d*) is immediate. This can account, for example, for revaluation effects as backward blocking (Shanks, 1985).

codified as an extra cue, and thus recruit associative strength by itself, and/or can be considered as the reason why the absence of the target cue is expected in those trials in which it is actually absent. Given the observed results, we have assumed, first, that the context must be codified as an extra cue; and second, that it has the power to trigger the expectation of the target cue when it is absent (thus making possible its negative codification). Under these assumptions, an absent target cue will be similarly expected in both types of inhibitory procedures in this experiment, either by the presence of the context (X alone procedure) or by the explicit cue Y (YX design) given that this cue is always present. This assumptions makes the two procedures virtually equivalent.

How the influence of the context has to be codified remains a matter of debate in human and animal learning literature, and, probably, it is procedure-dependent. Then, it is possible that the procedural differences between Chapman and Robbins' (1990) experiments and this work could also explain the different obtained results. Among others differences, they used a within subjects design, making it possible a direct subjective comparison between the two types of inhibitory cues (accompanied or not by a predictive cue) and, at the same time, their procedure made the explicit cue and the context clearly differentiable.

Table 1 displays the product of simulating Markman's model, being the parameters also reported. Predictions from Markman's revised R-W rule qualitatively account for the Last-trial Effect both in the positive and the negative contingency task, although the quantitative fit is much better for the belief revision model, as previously reported (Catena et al., 2002). Moreover, Markman's model does not account for the main Response Mode effect found in the detection of positive relationships, as the predicted judgment means are similar in both frequency conditions. The belief revision model predicts the trend shown by Frequency of Judgment Effect in the detection of generative causal relationships, and the absence of this effect and the global lower accuracy observed in preventative conditions. The better adjustment of BRM is clear, although the actual difference observed between the HF and the LF group in the positive contingency task is larger than the one predicted by BRM. Note that the superiority of the proposed model appears under the self-restriction of using the same parameters that have been used in previous works.

Besides its empirical flaws, Markman's rule has also the theoretical limitation that it requires selecting different parameter sets for high- and low-frequency tasks, in order to be able to account for the observed data pattern. In this model, both qualitative and quantitative predictions are extremely sensible to the selected set of parameters. The salience parameter for the negative cue (Y) need to be close to 0 (β= 0.001) to account for the actual data pattern in the high frequency condition. This value theoretically implies almost the absence of learning about that cue, and that all the observed effects should be attributed to the context or the explicit cue influence. Finally, the quantitative fit for Markman's model predictions is always lower than for the belief revision model. The RMSE value (Root Mean Square Error) reported in Table

1 assesses the quantitative adjustment of each model to the observed data pattern. It is also important to note that that the set of parameters of the belief revision model is hold constant in all conditions, and is the same that has been used in previous studies (Catena et al., 1998, Catena et al., 2002).

Catena et al (1998) proposed that causal and covariation learning is based on the action of two serial mechanisms. Each time a judgment is required, the information-computing mechanism (New-Evidence, Equation 4) comes into play, and computes the degree of covariation between the cue and the outcome observed since the last judgment, by using the frequencies of the four trial types that have been stored in the working memory during the task. The information summarized by this mechanism then enters the information-integrating mechanism, which updates the last judgment in accord to that input (Equation 3), resetting the first mechanism and the working memory to compute the New-Evidence again. This last judgment is transferred to and updated in the reference memory, where it is stored to be retrieved when necessary. All this process is summarized in Figure 1.

The Belief Revision model can explain, not only the influence of the last trial in the high frequency condition, but also why this effect is dependent on the actual contingency, be it positive, null or negative. Table 1 shows the predictions of this model for the two experiments reported in this work. It is important to note that given the higher and positive weights always assigned to a type trials, this model is able to predict the lack of accuracy and higher difficulty of subjects detecting negative contingencies. However, in the case of the influence of d type trials, the predictions of the model are less accurate, perhaps due to the intrinsic variability usually found in the interpretation of this type of trials (White, 1998, see also Maldonado et al, 1999). In any case, the model predictions are always more accurate than the predictions from any other model (see RMSE in Table 1).

Concluding Remarks

The effects of judgment frequency manipulation on the detection of generative and preventative causal relationships emphasise the need of postulating several integrated mechanism to account for human causal learning (Waldman and Martignon, 1998). Associative models, including the recently proposed reformulated SOP model (Dickinson et al., 2001), need to make additional assumptions to account for these effects, such as the possibility of computation of an absent cue (Markman, 1989). Moreover, it is necessary to explain why the set of parameters changes from task to task. As shown by the simulations reported in Table 1 there is not a single set of parameters with which Markman's model is able to account for the whole pattern of data. To date, the change of parameters has not been based on theoretical grounds, which, eventually, multiplies the number of parameters actually assumed by the model. Conversely, the belief revision model is able to accurately account for the data obtained in this and previous works by using a single set of parameters.

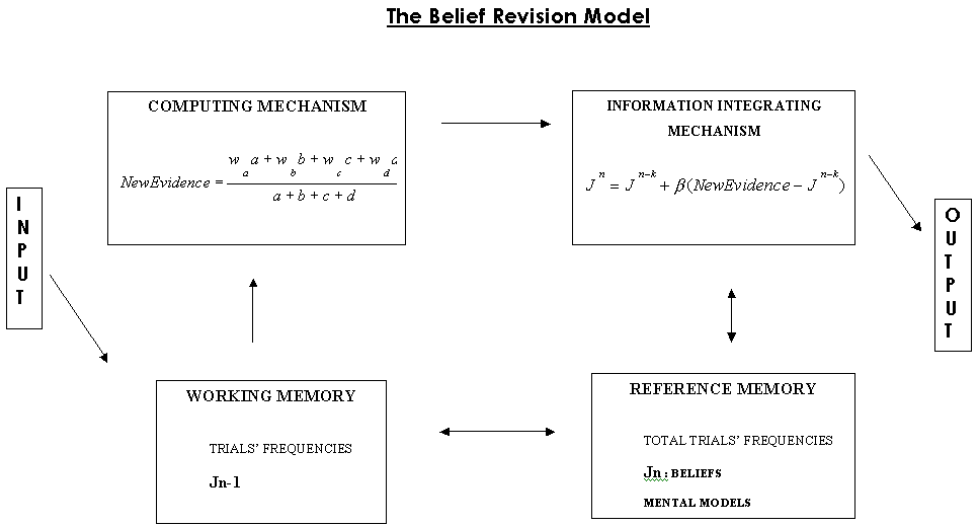

Figure 1. The serial two-mechanisms cognitive architecture proposed by the Belief Revision Model (Catena et al, 1998) to account for human causal and covariation learning.

The belief revision model (see Figure 1) proposes a cognitively coherent and parsimonious set of mechanisms to account for causal and covariation learning. The lower-level algorithm is proposed to compute partial covariation from the information stored in the working memory since the last estimation (New-Evidence). The higher-level algorithm updates previous estimations by using the information provided by the lower-level mechanism. It is also tentatively proposed that this high-order integration mechanism could have broader functions as a mechanism that makes possible the interaction among several sources of information such as instructions and experience with the task (Maldonado et al, 1999), causal mental models and previous beliefs (Perales, Catena, and Maldonado, 2003, Waldman and Martignon, 1998), attentional and executive demands (Maldonado et al, 2003), and emotional and motivational factors (for example, incentives value, see Reed, 1994).

RESUMEN

Efecto de la frecuencia de juicio en el aprendizaje de relaciones causales generativas y preventivas. El efecto de la frecuencia de juicio es un caso especial del modo de respuesta en el aprendizaje humano de relaciones causales y de covariación, con el que se muestra que el ajuste de los juicios -a ΔP-, depende del tipo de ensayo previo al juicio, aunque este efecto sólo ocurre cuando la frecuencia del juicio es alta. En dos experimentos se demostró la fiabilidad y generalidad de dicho efecto en tareas de aprendizaje de relaciones causales generativas y preventivas. El primer experimento demostró ese mismo efecto, con un mayor número de ensayo (16) y un mayor grado de contingencia positiva (ΔP= 0.71) que en estudios previos. El Experimento 2 demostró que la frecuencia del juicio modula también la detección de relaciones de contingencia negativas (ΔP= -0.5), dado que el ajuste a dicha contingencia dependía del tipo de ensayo precedente en condiciones de alta frecuencia del juicio. Los modelos basados en la acción de un único mecanismo, sea asociativo o estadístico, no pueden explicar fácilmente estos resultados sin incorporar nuevos supuestos. Por tanto, estos resultados aportan nueva evidencia experimental al cuerpo de datos que sugieren que el aprendizaje de relaciones causales en humanos depende de la acción integrada de diferentes mecanismos, como propone el modelo de revisión de creencias.

REFERENCES

Allan, L. (1993). Human contingency Judgments: Rule based or associative? *Psychological Bulletin, 114,* 435-448.

Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of the frequency of judgment and the type of trial on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 481-495.

Catena, A, Maldonado, A., Megías, J.L., & Frese, B (2002). The frequency of judgment effect: Belief revision and serial processing of causal information. *Quarterly Journal of Experimental Psychology, 55*(B), 267-281

Chapman, G.B. & Robbins, S.J. (1990). Cue interaction in human contingency judgment. *Memory and Cognition, 18,* 537-545.

Chatlosh, D.L., Neunaber, D.J., & Wasserman, E.A. (1985). Response-outcome contingency: behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning and Motivation, 16,* 1-34.

Cheng, P. (1997). From covariation to causation: A Causal Power Theory. *Psychological Review, 104* (2), 367-405.

Cheng, P. & Novick, L (1992). Covariation in natural causal induction. *Psychological Review, 99* (2), 365-382.

Collins, D. & Shanks, D.R. (2002). Momentary and integrative response strategies in causal judgment. *Memory and 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 Psychology, 55* (B), 289-310.

Dickinson, A. (2001) Causal learning: an associative analysis. *Quarterly Journal of Experimental Psychology, 54* (B), 3-25.

Gluck, M & Bower, G. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117* (3), 227-247.

Hastie, R., & Pennington, N. (1995). Cognitive approaches to judgment and decision making. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 32, pp. 1-31). San Diego, CA: Academic Press.

Hogarth, R.M & Einhorn, H.J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1-55.

Kao, S.F. & Wasserman, E.A. (1993). Assessment of an information integration account 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.

Maldonado, A., Catena, A., Cándido, A., & García, I. (1999). The belief revision model: Asymmetrical effects of no contingency on human covariation learning. *Animal Learning & Behavior, 27*, 168-180.

Maldonado, A., Herrera, A., Jiménez, I., Perales, J.C. & Catena, A. (2003). Inattentional blindness for negative relationships in human causal learning. Submitted.

Markman, A. (1989) LMS rules and the inverse base-rate effect: Comments on Gluck and Bower (1988) *Journal of Experimental Psychology: General, 118,* 417-421.

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

Miller, R. & Matute, H. (1996) Animal analogues of causal judgment. *The Psychology of Learning and Motivation, 34,* 133-166.

Perales, J.C., Catena, A., and Maldonado, A. (2003). Inferring non-observed correlations from causal scenarios: The role of causal knowledge. *Learning and Motivation.* (In press).

Perales, J.C. and Shanks, D.R. (2003). Normative and Descriptive accounts of the effect of power and contingency in causal judgment. *Quarterly Journal of Experimental Psychology: A, 56,* 977-1007.

Price, P.C. & Yates, J.F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning. Memory and Cognition, 21 (6),* 1639-1655.

Reed, P. (1994). Influence of the cost of responding on human causality Judgments. *Memory and Cognition, 22,* 243-248.

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 theory and research* (pp. 64-99 New York: Appleton-Century-Crofts.

Shanks, D.R. (1985). Forward and backward blocking in human contingency judgment. *The Quarterly Journal of Experimental Psychology, 37,* 1-21.

Shanks, D.R. (1987) Acquisition functions in contingency judgment. *Learning and Motivation, 18,* 147-66.

Shanks, D.R. (1991). On similarities between causal judgments in experienced and described situations. *Psychological Science, 2 (5),* 341-350.

Shanks, D.R. (1995) Is human learning rational? *Quarterly Journal of Experimental Psychology,48A,*257-79.

Shanks, D.R. & Dickinson, A. (1987). Associative accounts of causality judgment. In G.H. Bower (Ed.). *The Psyhology of Learning and Motivation* (Vol 21, pp. 229-261). San Diego: Accademic Press.

Tolman, E.C., & Brunswick, E. (1935). The organism and the causal texture of the environment. *Psychological Review, 42,* 43-77.

Van-Hamme, L. & Wasserman, E. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation, 25 (2),* 127-151.

Vila, J. (2000) Extinción e Inhibición en Juicios de Causalidad. *Psicológica, 21*, 257-273.

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

Waldmann, M.R. and Holyoak, K.J. (1992). Predictive and diagnostic learning with causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121,* 222-236.

Waldmann, M.R. and Martignon, L. (1998). A Bayesian network model of causal learning. *Proceedings of the 20th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: LEA.

Wasserman, E.A. (1990). Comparative cognition: Toward a general understanding of cognition in behavior. *Psychological Science, 4,* 156-161.

White, P. (1998). Causal Judgments: Use of different types of contingency information as confirmatory or disconfirmatory. *European Journal of Cognitive Psychology, 10,* 131-170.

(Manuscript received: 8 April 2003; accepted: 10 September 2003)