Use of prior beliefs in the assignment of causal roles: Causal powers versus regularity-based accounts

PETER A. WHITE

University of Wales College of Cardiff, Cardiff, Wales

There is a tradition of models of causal judgment in which causes and other causal roles are defined and identified in terms of empirical patterns of association with effects. In the present experiments, results conflicting with the predictions of such models were obtained. In one experiment, subjects judged that an interpretation in which a factor constantly present was identified as the cause was more likely than was an interpretation in which a perfect positive covariate was identified as the cause. In a second experiment, possible effects of prior beliefs about covariation were controlled and similar findings were obtained in two out of three scenarios. These results favor the idea that people make causal judgments by applying preexisting beliefs framed in terms of causal concepts, such as causal powers, and in ways that cannot be accounted for by models in the empiricist tradition.

Many theories and models of causal judgment and attribution have been based on the idea of regularity of association between cause and effect. In these models, regularity may be represented as covariation or contingency (Forsterling, 1989; Kelley, 1967, 1973; Schustack & Sternberg, 1981), as conditional relations (Hewstone & Jaspars, 1987; Hilton & Slugoski, 1986), or as probabilistic association (Cheng & Novick, 1990, 1991, 1992; Shanks, 1993; Shanks & Dickinson, 1987). Despite these differences, the various models all agree in the fundamental claim that a thing that exhibits regularity of association with an effect is more likely to be identified as a cause of that effect than is a thing that does not exhibit regularity of association with the effect.

This claim is usefully illustrated in a study by Wasserman (1990). In this study, subjects were asked to imagine that they were trying to discover the cause of an allergic reaction in a patient. They were presented with data regarding meals eaten by the patient over a period of time and whether or not the allergic reaction occurred after each meal. One of the data sets used by Wasserman is shown in Table 1.

In this table, it can be seen that the combination of shrimp and strawberries is always followed by an allergic reaction, and the combination of shrimp and peanuts never is. The three foods in the table have three different patterns of association with the outcome. The strawberries are a *positive covariate* (in this case, a perfect posi-

tive covariate, present whenever the outcome occurs and absent whenever it does not occur); the shrimps are a *constant factor*, present in all meals regardless of the occurrence of the outcome; and the peanuts are a *negative covariate* (in this case, a perfect negative covariate). The empirical relations are also unidirectional, meaning that the effect is not temporally prior to any possible cause. For the data set shown, subjects gave high ratings of causal efficacy to the positive covariate and very low ratings of causal efficacy to the constant factor and the negative covariate. In other data sets, the ratings of causal efficacy of the three components converged toward a moderate value as the distinctiveness of association of the positive covariate with the outcome decreased.

The main aim of that study was to show that ratings of the causal efficacy of the constant factor varied depending on the strength of association between the positive covariate and the outcome: here, Wasserman drew a parallel with results found in studies of animal learning (e.g., Wagner, Logan, Haberlandt, & Price, 1968; Wasserman, 1974). In general, though, the results are consistent with the fundamental claim of regularity-based¹ accounts of causal judgment: for the data set shown, all such models predict higher judgments of causal efficacy for the positive covariate than for the constant factor or the negative covariate, because the positive covariate has the strongest empirical association with the occurrence or nonoccurrence of the outcome.

Wasserman's study, however, only allowed subjects to say whether a given factor was a cause of the outcome or not. It did not allow them to assign different causal roles to the various factors. To see what sorts of causal roles there might be, consider a theoretical approach to causal judgment not based on the use of regularity information: the causal powers theory (White, 1989, 1992, 1993; see also Harré & Madden, 1975). This theory postulates that

I am happy to express my gratitude to Valerie Thompson and an anonymous reviewer for their most perspicacious and helpful comments on earlier versions of this article. The discussion of Experiment 1 in particular owes a great deal to their contribution. I am also grateful to Alan Milne for giving me confidence in my arithmetic. Correspondence should be addressed to P. A. White, P.O. Box 901, Cardiff CF1 3YG, U.K.

Table 1
Example Data Set Used by Wasserman (1990, Condition 1.00)

Meal	Allergic Reaction	No Allergic Reaction
Shrimp and strawberries	8	0
Shrimp and peanuts	0	8

people possess a framework of concepts that underpins and gives meaning to particular causal beliefs. These concepts can be illustrated with the example of a hammer smashing a plate. Under the causal powers theory, people understand this as a specific *causal power* of the hammer acting to *produce* or *generate* (Shultz, 1982) the effect of the plate smashing under the *releasing condition* of forcible contact between them. Reference may also be made to the *liability* or *resistance* of the plate, meaning its capacity to be smashed by the hammer. A causal power, then, is a (usually) stable property of a thing, a capacity to produce a certain sort of effect; the causal relation is understood as generative or productive;² and the power of a thing operates to produce its effect under some suitable condition or set of conditions. The role of the releasing condition is important. Although the hammer possesses the power to smash plates at all times, it is not constantly doing so: the power only operates when a suitable condition (such as a certain kind of forcible contact) is met. One more concept may be added to this: if a pad of some shock-absorbing substance is placed on top of the plate, then the hammer may not succeed in smashing the plate. This illustrates the notion of a condition that *prevents* or *inhibits* the production of an effect by the operation of a causal power. Under the causal powers theory, people acquire an extensive range of particular causal beliefs that have meaning as causal (as opposed to noncausal) beliefs because they draw on this basic conceptual repertoire. Causal judgment is largely a matter of ascertaining the applicability of a causal belief to an instance, although this is complicated by such factors as information availability and the prevailing practical concerns of the judge (White, 1989, 1992, 1993).

Under this theory, for the data set shown in Table 1, people may still identify the positive covariate as the cause of the effect, but at least two other causal interpretations are possible. In one, the constant factor is identified as the cause of the outcome and the positive covariate is identified as a releasing (or enabling) condition under which the power of the constant factor operates. In the other, the constant factor is identified as the cause of the outcome and the negative covariate is identified as a factor that prevents or inhibits the operation of the causal power of the constant factor. Note that (1) in both of these interpretations the constant factor and not the positive covariate is identified as the cause, and (2) choice of interpretation is governed not by the pattern of empirical association but by the particular beliefs held by the subject about the things in question. This implies that Wasserman's findings hold only when subjects are not allowed to assign causal roles to the

elements in compound stimuli or when the particular causal beliefs they hold lead them to favor an interpretation in which the positive covariate is the cause. These implications are tested in Experiment 1 by allowing subjects to choose between different interpretations of a data set corresponding to those listed above and by manipulating the content of the scenario without altering the pattern of empirical association in the data set.

The kind of data set used by Wasserman is most useful for the proposed test because (1) it is one for which all regularity-based models generate the same prediction, and (2) the presence of multiple elements is necessary to test any proposition that people assign different causal roles to different elements. Essentially, for the data sets to be used in the first experiment, regularity-based models predict that subjects will preferentially identify the positive covariate as the cause, regardless of scenario content. Under the causal powers theory, exact predictions depend on details of scenario content: the aim is to show that, for some data sets, an interpretation in which the constant factor is identified as the cause will be preferred to one in which a perfect positive covariate is identified as the cause. Specifically, the aim is to design scenarios in which the preferred interpretation identifies the constant factor as the cause and the negative covariate as an inhibiting factor. Because such an interpretation assigns no causal role to the positive covariate, a preference for that interpretation would be contrary to the predictions of regularity-based models.

EXPERIMENT 1

Method

Subjects. The subjects were 50 paid first-year undergraduate students of subjects other than psychology at University of Wales College of Cardiff, recruited by means of advertisements around the campus. There were 26 women and 24 men, with a mean age of 21.8 years (range 18–31).

Materials. Three scenarios were used. Each scenario began with a brief cover story, and then presented a data set with a pattern of empirical association similar to that shown in Table 1. The data sets were presented in the form of a 2×2 table, which is one of the two presentation methods used by Wasserman (1990).³ The subjects judged the likelihood of each of three possible interpretations of the data set: (1) identifying the positive covariate as the cause (labeled for convenience the *simple-cause* interpretation), (2) identifying the constant factor as the cause and the positive covariate as an enabling condition (the *enabling* interpretation),⁴ and (3) identifying the constant factor as the cause and the negative covariate as an inhibiting condition (the *inhibiting* interpretation). The subjects were instructed to assign to each interpretation a score out of 100, where 0 meant *definitely not the right interpretation* and 100 meant *definitely the right interpretation*. The more likely they thought it was that a given interpretation was the right one, the higher the number they should put beside it. The subjects were not required to make their ratings sum to 100.

Each scenario existed in two versions, labeled *high-belief covariate* and *low-belief covariate*. In the high-belief covariate version, the perfect positive covariate was something that is commonly believed to have the causal power to produce the effect in question. In the low-belief covariate version, the perfect positive covariate was something that is not commonly believed to have the

causal power to produce the effect in question. This manipulation was achieved by changing roles for the elements in the scenario: the element that was the perfect positive covariate in one version was the constant factor in the other, and vice versa. This manipulation holds scenario content constant.

For illustrative purposes, one scenario will be presented in full. This is the "dog" scenario. In this scenario, the introductory paragraph read:

Imagine you are a psychiatrist who is trying to determine the cause of a fear reaction in a patient who is afraid of some dogs but not others. You keep a record of dogs the patient has encountered and whether the fear reaction occurred or not. After a while you find you have gathered the following observations:

Table 2 depicts the data sets used for this scenario. As the table shows, in the low-belief covariate condition, having a black coat is the positive covariate and being big is the constant factor. In the high-belief covariate condition being big is the positive covariate and having a black coat is the constant factor. Everything else is identical. The three possible interpretations offered to subjects in the low-belief covariate condition were:

1. *The fact that the dog has a black coat causes the fear reaction.* This is the simple-cause interpretation, identifying the positive covariate as the cause.

2. *The fact that the dog is big causes the fear reaction, but it only does so when the dog also has a black coat.* This is the enabling interpretation, identifying the constant factor as the cause and the positive covariate as an enabling or releasing condition.

3. *The fact that the dog is big causes the fear reaction, but the fear reaction is prevented from occurring by the fact that the dog is muzzled.* This is the inhibiting interpretation, identifying the constant factor as the cause and the negative covariate as an inhibiting condition.

The three alternatives for the high-belief covariate condition were similar except that the black coat and big size were interchanged. Brief details of the other scenarios now follow.

In the "freeze" scenario, the cover story concerned observations of conditions under which the surface of a river freezes or does not freeze. In the low-belief covariate condition, a cloudless sky was the positive covariate, sub-zero temperature was the constant factor, and a turbulent current was the negative covariate. In the high-belief covariate condition, sub-zero temperature was the positive covariate and cloudless sky the constant factor.

In the "upset-stomach" scenario, the cover story concerned observations of meals eaten following which a patient suffered an upset stomach or did not. In the low-belief covariate condition, a pint of lager was the positive covariate, a hot curry was the constant factor, and indigestion tablets were the negative covariate. In the high-belief covariate condition, the curry was the positive covariate and the lager was the constant factor.

Procedure. The materials for this experiment were included among a set of questionnaire materials on unrelated topics. The three scenarios were separated from each other by materials on other topics. The subjects were run in small groups of 2 or 3 each,

supervised by an experimenter. They were seated apart so that none could see what the others were doing. The experimenter was available both to supervise the subjects and to answer queries (however, there were no queries about the materials for this experiment).

Predictions. In each case, scenario content was chosen in accordance with commonly held beliefs. It is widely believed that people tend to fear big dogs more than they fear small ones, but there is no common belief relating fear to the color of a dog's coat. Most people know that water is made to freeze by sub-zero temperature, not by a cloudless sky (independently of temperature). There is (at least in Britain) a stereotype (false, in the author's experience) of hot curries as causes of upset stomachs, but no common belief about the effect of a single pint of lager. In addition, the negative covariate in each scenario was chosen to reflect a common belief about what can block or inhibit the causal power exploited in the scenarios. Thus, people believe that people are less likely to fear big dogs with muzzles (because the muzzle reduces the actual threat posed by the dog), a turbulent current can prevent water from freezing when the air temperature is below zero (because of the mixing of water of different temperatures), and indigestion tablets can help to alleviate an upset stomach.

Given these common beliefs, the causal powers theory predicts that, in each scenario, the inhibiting interpretation should be judged more likely in the low-belief covariate condition than in the high-belief covariate condition. It also predicts that the inhibiting interpretation should be judged more likely than the simple-cause interpretation within the low-belief covariate condition. Regularity-based models predict that the interpretation identifying the positive covariate as the cause (i.e., the simple-cause interpretation) should be preferred in both conditions of all three scenarios.

Results

One copy of the materials for the high-belief covariate condition of the upset-stomach scenario was unintentionally omitted from the set of materials for one subject. For this condition, $n = 24$. For all other conditions, $n = 25$. Each scenario was analyzed separately. In each case data were analyzed with a mixed-design two-way analysis of variance (ANOVA), the between-subject factor being conditions and the within-subject factor being interpretations. Means of causal judgments for all scenarios are reported in Table 3.

Tests of predictions. Under the causal powers theory, it is predicted that the inhibiting interpretation should be judged more likely in the low-belief covariate condition than in the high-belief covariate condition. This prediction was tested with simple effects analysis⁵ and, in each scenario, the results supported the prediction [for the dog scenario, $F(1,143) = 4.25$, $MS_e = 742.83$, $p < .05$; for the freeze scenario, $F(1,143) = 31.39$, $MS_e = 759.72$, $p < .001$; for the upset-stomach scenario, $F(1,139) = 13.33$, $MS_e = 660.70$, $p < .001$].

Under the causal powers theory, it is also predicted that the inhibiting interpretation should be judged more likely than the simple-cause interpretation within the low-belief covariate condition. This prediction was tested with paired comparisons carried out using a one-way ANOVA with repeated measures. In each scenario, the results again supported the prediction [for the dog scenario, $F(1,24) = 9.90$, $MS_e = 1,266.95$, $p < .01$; for the freeze scenario, $F(1,24) = 241.72$, $MS_e = 288.40$,

Table 2
Data Sets Used in Experiment 1, Dog Scenario

	Fear Reaction	
	Yes	No
Low-Belief Covariate Condition		
Dog is big and has black coat	6	0
Dog is big and is muzzled	0	6
High-Belief Covariate Condition		
Dog has black coat and is big	6	0
Dog has black coat and is muzzled	0	6

Table 3
Mean Likelihood Judgments for Each Interpretation
in Experiment 1

Condition	Interpretation		
	Simple Cause	Enabling	Inhibiting
Dog Scenario			
High-belief covariate	63.48	41.24	59.24
Low-belief covariate	43.40	37.20	75.08
Freeze Scenario			
High-belief covariate	86.20	57.60	46.84
Low-belief covariate	15.84	49.24	90.52
Upset-Stomach Scenario			
High-belief covariate	65.00	55.00	45.58
Low-belief covariate	38.80	54.68	72.40

Note—Scores may range from 0 (*definitely not the right interpretation*) to 100 (*definitely the right interpretation*).

$p < .001$; for the upset-stomach scenario, $F(1,24) = 16.16$, $MS_e = 873.46$, $p < .001$]. These results run counter to the prediction of regularity-based models that the interpretation identifying the positive covariate as the cause (i.e., the simple-cause interpretation) should be preferred.

Additional findings. In addition to the differences predicted, there were other instances in which the constant factor was judged significantly more likely to be the cause of the effect than the perfect positive covariate. In the low-belief covariate condition of the freeze scenario, the enabling condition was judged more likely than the simple-cause interpretation [$F(1,24) = 18.05$, $MS_e = 772.54$, $p < .001$]. The same pattern was found in the low-belief covariate condition of the upset-stomach scenario [$F(1,24) = 4.57$, $MS_e = 690.10$, $p < .05$].

In the dog scenario, there was a significant main effect of interpretation [$F(2,96) = 13.37$, $MS_e = 729.94$, $p < .01$]. Paired comparisons with the t test for related means revealed significant differences in all cases, with the inhibiting interpretation receiving higher likelihood judgments than the other two interpretations ($p < .05$, for the comparison with the cause interpretation, and $p < .01$ for that with the enabling interpretation) and the cause interpretations receiving higher likelihood judgments than the enabling interpretation ($p < .01$). There was also a main effect of interpretation in the freeze scenario [$F(2,96) = 6.47$, $p < .01$]. Paired comparisons with the t test for related means revealed that the inhibiting interpretation received higher judgments of likelihood than either of the other two interpretations ($p < .01$, in both comparisons), but there was no significant difference between the cause and enabling interpretations. There was no effect of interpretation for the upset-stomach scenario, although there was a nonsignificant preference for the inhibiting interpretation over the simple-cause interpretation.

Finally, comparisons between the high-belief covariate condition and the low-belief covariate condition, in which both the form of interpretation and the empirical

pattern in the data were controlled, revealed significant effects of scenario content in addition to those predicted. In each scenario, the simple-cause interpretation was judged more likely in the high-belief covariate condition than in the low-belief covariate condition [for the dog scenario, $F(1,143) = 6.56$, $MS_e = 742.83$, $p < .05$; for the freeze scenario, $F(1,143) = 81.45$, $MS_e = 759.72$, $p < .001$; for the upset-stomach scenario, $F(1,139) = 12.72$, $MS_e = 660.70$, $p < .001$]. There were no significant differences between the high-belief covariate condition and the low-belief covariate condition for the enabling interpretation.

Discussion

The results are consistent with the predictions based on the causal powers theory. In each scenario, the inhibiting interpretation was judged more likely in the low-belief covariate condition than in the high-belief covariate condition. Within the low-belief covariate condition of each scenario, the inhibiting interpretation was judged more likely than the cause interpretation. These results disconfirm the predictions of regularity-based models. Such models predict that a perfect positive covariate of an effect should be identified as the cause of that effect in preference to a factor constantly present. The inhibiting interpretation assigns no causal role to the perfect covariate at all.

The main effects of interpretation reported under additional findings show that, in two out of three scenarios, the inhibiting interpretation was significantly preferred over the simple-cause interpretation. The *content* of the cause in this comparison is controlled, so the only substantive difference between the two is the presence or absence of the negative covariate in the interpretation as an inhibiting factor. This therefore shows that the negative covariate plays a significant role in the causal interpretation of the empirical data. This supports the case for the claim, under the causal powers theory, that people adopt interpretations of effects in which different factors or events have different causal roles. It shows also that the combination of constant factor and negative covariate can be preferred over a positive covariate as a causal interpretation of an effect even when the identities of the positive covariate and constant factor are controlled by counterbalancing. This further weakens the case for the basic claim of regularity-based models that perfect positive covariates should dominate causal interpretations.

There is one regularity-based model, however, that may be able to account for these results. This is the probabilistic contrast model (Cheng & Novick, 1990, 1991, 1992), in which causal roles are defined empirically. Cheng and Novick defined a cause as a factor, the presence of which (relative to its absence) increases the likelihood of the effect. For a cause involving a single factor, the probabilistic contrast is defined by a simple unidirectional contingency rule (Alloy & Abramson, 1979):

$$\Delta p_i = p_i - p_{-i},$$

where i denotes some possible cause, p_i denotes the proportion of cases on which the effect occurs in the presence of i , and p_{-i} denotes the proportion of cases on which the effect occurs in the absence of i (Cheng & Novick, 1992, p. 367). A positive contrast specifies a cause, whereas a negative contrast specifies an inhibitory factor.

Probabilistic contrasts can also be calculated for possible causes involving interactive factors. Cheng and Novick (1992) state:

A two-way interaction contrast, Δp_{ij} , involving potential causal factors i and j , is defined as follows:

$$\Delta p_{ij} = (p_{ij} - p_{-ij}) - (p_{i-j} - p_{-i-j}),$$

where p , as before, denotes the proportion of cases in which the effect occurs when a potential contributing factor is either present or absent, as denoted by its subscripts. (p. 368)

Probabilistic contrasts are calculated for events within a selected set, called a *focal set*. A focal set need not include all events that could be deemed relevant to the problem in question, but it may be defined in different ways on different occasions or by different people. It follows that different people using the probabilistic contrast model may identify different causes for a given effect, if they have defined different focal sets. It is also possible that an event identified as a cause by the model under one focal set may be identified as an enabling condition by the model under a different focal set.

Cheng and Novick (1992) made some definite statements about the interpretation of constant factors. They stated that "covariation is a necessary condition for causal induction" (p. 367), which implies that a constant factor—which of course is not a covariate—will not be identified as a cause, an implication on which all regularity-based models are in agreement. They argued that a constant factor may be identified as either an enabling condition or a causally irrelevant factor, depending on its presence in other possible focal sets. They further stated: "Our model predicts that a potential causal factor that covaries with the effect in the focal set will be considered a cause and will be distinguished from necessary factors that are constantly present in that set" (Cheng & Novick, 1992, p. 378). By this reasoning, for the data set in Table 2, the positive covariate should be identified as the cause because it is distinctively and positively probabilistically associated with the outcome, and the constant factor should not be identified as the cause, not only because it is constantly present but also because it is less strongly associated with the outcome than the positive covariate is. On the face of it, then, the results of Experiment 1 are contrary to the predictions of the probabilistic contrast model, as they are to other regularity-based models. There is one way out for the model, however.

To compute a probabilistic contrast requires information about the occurrence of the effect when a given cause is both present and absent. The problem is that the

data sets presented in Experiment 1 are not sufficiently informative to allow this. Taking the dog scenario as an example (see Table 2), the main effect contrast for size cannot be calculated because no information is presented about cases in which the dog is not big. The main effect contrast for having a black coat can be calculated on the assumption that the dog does not have a black coat when the color of the coat is not mentioned, but there is no guarantee that subjects actually make this assumption. The same can be said of the main effect contrast for being muzzled. Finally, the interactive contrast for being big and muzzled also cannot be calculated because of the lack of information about dogs that are not big.

Faced with this, someone wishing to calculate probabilistic contrasts could only do so by making assumptions about the missing information. That is to say, they may draw on prior beliefs to fill in the missing values and then use those in combination with the information given in the data set to calculate probabilistic contrasts. In the language of the model, they are defining a focal set that includes not only the data given but also relevant prior beliefs, and they compute a probabilistic contrast for that defined set. Consider first having a black coat. Suppose a judge has the prior belief that non-black dogs cause fear reactions as often as black dogs. This means, in effect, that the judged probability of fear is the same for black dogs as for non-black dogs, which on its own gives a main effect contrast of zero. If this is integrated with the information given in the data set, then the main effect contrast for having a black coat will be marginally above zero, depending on how the prior beliefs are weighted against the information given.

Now consider size. In the data set given, size is a constant factor, and the fear reaction occurs for half of the cases and not for the other half. If a judge has the prior belief that fear reactions occur more frequently for big dogs than for small dogs, and integrates this with the information given in the data set, then the main effect contrast for size will be positive. How large it is will depend on how it is weighted relative to the information given in the data set and on the proportions present in the prior belief. For example, suppose the judge has the belief that fear occurs in the presence of 80 out of 100 big dogs and 20 out of 100 small dogs. Integrating this with the information given in the data set results in a table in which the fear reaction occurs for 86 big dogs and does not occur for 26 big dogs, and the fear reaction occurs for 20 small dogs and does not occur for 80 small dogs. Clearly, with these numbers, the occurrence of fear is strongly associated with size, and there is a strong positive main effect contrast for size. As for the muzzle, a judge need only assume that the fear reaction occurs in some cases when the dog is not muzzled and never when it is muzzled. This would result in a strong negative main effect contrast for the muzzle, which, according to Cheng and Novick (1992), should lead subjects to identify it as an inhibitory factor.

If subjects follow this procedure, resulting in a strong positive contrast for size, a strong negative contrast for

being muzzled, and a contrast close to zero for color, this would account for the observed preference for the inhibiting interpretation over the simple-cause interpretation in the low-belief covariate condition. Note the form that the prior beliefs take: they are beliefs about *empirical association* (specifically unidirectional contingency), not causal beliefs as such. This is how the probabilistic contrast model works: it generates causal judgments from information about contingency, whether that information be given to subjects or already held by them as prior beliefs, or some combination of the two.

The assumptions used in the example are quite plausible: most people would agree (1) that fear is more likely for big dogs than for small ones, regardless of their color, (2) that fear is not much, if at all, more likely for black dogs than for dogs of other colors, regardless of their size, and (3) that fear is comparatively unlikely to occur when the dog is muzzled. If subjects have prior beliefs like these, and draw on these beliefs to make their causal judgments, then the probabilistic contrast model can account for the results of the dog scenario. An equivalent argument can be mounted for each of the other scenarios used in Experiment 1.

Since the probabilistic contrast model and the causal powers theory can both account for effects of prior beliefs on judgment, it is necessary to distinguish them by referring to the kind of belief they employ. The probabilistic contrast model employs prior beliefs about contingency or covariation; the causal powers theory employs prior beliefs about causal powers and other causally relevant properties of things. It is therefore necessary to show that causal judgments are influenced by one kind of belief as distinct from the other. This was the strategy adopted in Experiment 2. For each of three scenarios, there were two possible causes, one shown as a perfect positive covariate of an effect, and one as a weak positive covariate. (To be precise, $\Delta p = +.33$ for the main effect contrast for the weak covariate in each scenario.) Prior beliefs about the degree of correlation between each possible cause and the effect in question were assessed before subjects saw the data sets for the causal judgment task. To ensure that the probabilistic contrast model predicts a preference for the perfect positive covariate, it must be the case that subjects believe that possible cause to be correlated with the effect at least as strongly as the possible cause in the role of weak positive covariate. To ensure this, any subject who judged that the possible cause that would be the weak covariate was more highly correlated with the effect than the possible cause that would be the perfect covariate was excluded.

The aim of this strategy was to distinguish the possible effects of prior beliefs about contingency from beliefs about causal powers. The stimulus materials posed a contest between a factor that covaried relatively strongly with an effect, both in the data set given and in prior belief, but lacked the causal power to produce the effect, and a factor that had the causal power to produce the effect but covaried relatively weakly with it, both in the

data set given and in prior belief. According to the probabilistic contrast model, the former should have received higher causal ratings than the latter. According to the causal powers theory, the opposite should have been the case.

EXPERIMENT 2

Method

Subjects. The subjects were 51 first-year undergraduate students of psychology at University of Wales College of Cardiff, participating as part of a course requirement. There were 35 women and 16 men, with a mean age of 21.9 years (range 18–49).

Covariation judgment. The subjects were presented with a list of pairs of things and written instructions stating that they were to judge how often the things in each pair went together. The instructions reminded the subjects that A and B might not go together, not only because A might occur in the absence of B but also because B might occur in the absence of A.⁶ The subjects were also told that they were not judging whether one thing caused the other or not. Each pair was accompanied by a rating scale from 1 (*never*) to 7 (*always*). The subjects were instructed to circle whichever number on the scale best expressed their judgment of how often the two things go together.

There followed 18 items, each consisting of a pair of things. Six items were the prior belief items for the three scenarios used in the later, causal judgment phase of the study. These were:

Furniture-shop scenario:

- Being in a furniture shop and buying furniture.
- Needing new furniture and buying furniture.

Car scenario:

- In a car, speeding up and the accelerator being pressed.
- In a car, speeding up and the engine note rising.

Bus scenario:

- Standing at a bus-stop and catching a bus.
- Wanting to visit a friend and catching a bus.

At this stage, the subjects were unaware that these particular items related to any subsequent materials. These were randomly distributed among the 18 items. The remaining 12 items were fillers and were on unrelated topics.

Causal judgment task. Three scenarios were used. As in Experiment 1, there was a brief cover story followed by presentation of covariation information. The judgment task was the same as in Experiment 1, except that only two interpretations were offered for judgment—the simple-cause and inhibiting interpretations. For illustrative purposes, one scenario will be described in full. This is the “furniture-shop” scenario. The introductory paragraph read:

Imagine you are a social studies researcher investigating people's buying habits. You take a sample of people and see whether they buy furniture or not, and after a while you find you have gathered the following observations:

The data set for this scenario is presented in Table 4. After the instructions for the causal judgment task, the following two possible interpretations were given:

1. *People buy new furniture because they are in a furniture shop.* This is the simple-cause interpretation.

2. *People buy new furniture because they need new furniture, but not having money prevents them from doing so.* This is the inhibiting interpretation.

Δp for each element is as follows (see Appendix for details of the calculations): being in a furniture shop, $\Delta p_i = +1.0$; needing new furniture, $\Delta p_j = +0.33$; not having money, $\Delta p_k = -0.5$, where the negative sign indicates that not having money is an inhibitory factor (see Appendix for further explanation).

Table 4
Data Set Used in the Furniture-Shop Scenario
in Experiment 2

	Buy Furniture	
	Yes	No
1. People who are in a furniture shop, need new furniture, and have money	10	0
2. People who are not in a furniture shop, need new furniture, and don't have money	0	20
3. People who are not in a furniture shop, don't need new furniture, and do have money	0	10

The value of Δp_{ik} for the combination of needing new furniture and not having money cannot be calculated, because no information is presented concerning one of the terms in the formula (occurrences of the effect for people who don't need new furniture and don't have money). For present purposes, this is not a problem. We need only be sure that the degree of covariation between the combination (of needing new furniture and not having money) and buying furniture cannot be greater than that between being in a furniture shop and buying furniture, and this is guaranteed by the fact that there is perfect covariation between being in a furniture shop and buying furniture. There could only be perfect covariation between the combination (of needing new furniture and not having money) and buying furniture if all the subjects assume that people never buy furniture when they don't need it and don't have money. This is possible, but people do sometimes buy things they don't need and, moreover, often buy such things as furniture on credit, so it is perhaps unlikely that subjects would all make such an extreme assumption. In any case, there can be no assumption under which the covariation could be better than perfect, and this is sufficient to ensure the internal validity of the test. Briefer details of the other scenarios now follow.

In the "car" scenario, the cover story asked the subject to imagine he/she was a Martian engineer who has never visited Earth before, making observations to discover what makes cars speed up. The aim of this story was to provide extra incentive for the subjects to disregard their beliefs (on the grounds that a Martian engineer would probably not share them) and to judge from the data in a theory-neutral manner. The data set followed a similar pattern to that for the previous study, with the three elements being: the engine note rising, $\Delta p_i = +1.0$; the accelerator being pressed, $\Delta p_j = +0.33$; the car not being in gear, $\Delta p_k = -0.5$.

As already explained, the value of Δp_{ik} for the combination of the accelerator being pressed and the car not being in gear cannot be calculated. In any case, the degree of covariation cannot be better than perfect.

In the "bus" scenario, following a cover story about observations of a person catching a bus or not on a range of occasions, the data set conformed to the same pattern, with the three elements as follows: standing at a bus stop, $\Delta p_i = +1.0$; wanting to visit a friend, $\Delta p_j = +0.33$; knowing that the friend has gone away on holiday, $\Delta p_k = -0.5$.

Again, the value of Δp_{ik} for the combination of wanting to visit a friend and knowing that the friend has gone on holiday cannot be calculated, but the degree of covariation cannot be better than perfect.

As in Experiment 1, scenario content was designed with an eye to commonly held beliefs. In the car scenario, the perfect positive covariate is something that would be regarded by most people as a noncausal covariate of the mechanical operation of the car. Cheng and Novick (1992, p. 367) explicitly omitted noncausal covariates from their account, and it therefore seemed appropriate to include a scenario that illustrates the problems posed by noncausal covariates for regularity-based models (see also Bunge, 1963, and

White, 1990). In the other two scenarios, the perfect positive covariate is not noncausal but instead exemplifies a different kind of causal role, an opportunity. This is a condition that stands in relation to action much as an enabling condition stands in relation to causation (Harré & Secord, 1972). The aim is to show that even a causally relevant factor that has a stronger probabilistic contrast than any other factor or combination of factors will not necessarily be preferred as the cause.

Procedure. The materials for this experiment were included among a set of questionnaire materials on unrelated topics. The materials for the covariation judgment task were presented first. There followed a set of questionnaires on other topics, which typically took approximately 30 min to complete. The materials for the causal judgment task came after these. The subjects were run in small groups of 2 or 3 each, supervised by an experimenter. They were seated apart so that none could see what the others were doing. The experimenter was available both to supervise the subjects and to answer queries (however, there were no queries about the materials for this experiment).

Results and Discussion

Prior to analysis of the causal judgment task, subjects were excluded from the furniture-shop scenario if they gave a higher covariation judgment for needing new furniture and buying furniture than for being in a furniture shop and buying furniture. They were excluded from the car scenario if they gave a higher covariation judgment for speeding up and the accelerator being pressed than for speeding up and the engine note rising. Subjects were excluded from the bus scenario if they gave a higher covariation judgment for wanting to visit a friend and catching a bus than for standing at a bus stop and catching a bus. Following exclusion of subjects by these rules, $n = 30$ for the furniture-shop scenario, $n = 25$ for the car scenario, and $n = 49$ for the bus scenario.

With this selection procedure, Δp for the weak covariate must be less than Δp for the perfect positive covariate, even if prior beliefs are integrated with the data set given, for subjects included in analysis of the causal judgment task. The reason for this is that Δp for the weak covariate is less than Δp for the perfect positive covariate in the data set given, and Δp for the weak covariate is also no greater than Δp for the perfect positive covariate in the prior beliefs of the subjects included in the analysis of causal judgments, as assessed by the covariation judgment. The selection procedure also ensures that Δp for the interactive contrast for the two elements mentioned in the inhibiting interpretation cannot be higher than Δp for the element mentioned in the simple-cause interpretation, for subjects included in analysis of the causal judgment task.

If prior beliefs show that the perfect positive covariate and the weak covariate are judged to covary equally with the effect and if the number of cases involved is large, then Δp for the perfect positive covariate might be only marginally higher than Δp for the weak covariate. In this case, a failure to find a difference in preference for the two interpretations might not be fatal for the probabilistic contrast model. In fact, prior beliefs for subjects included in analysis of causal judgments, as analyzed with a one-way ANOVA with repeated measures, do not show

equal covariation. For the furniture-shop scenario, judged covariation of being in a furniture shop and buying furniture was significantly higher ($M = 4.30$) than was judged covariation of needing new furniture and buying furniture ($M = 3.60$) [$F(1,29) = 17.54$, $MS_e = 0.42$, $p < .001$]. For the car scenario, judged covariation of the engine note rising and the car speeding up was significantly higher ($M = 6.52$) than was judged covariation of the accelerator being pressed and the car speeding up ($M = 6.00$) [$F(1,24) = 8.89$, $MS_e = 0.38$, $p < .01$]. For the bus scenario, judged covariation of standing at a bus stop and catching a bus was significantly higher ($M = 5.57$) than was judged covariation of wanting to visit a friend and catching a bus ($M = 3.45$) [$F(1,48) = 106.74$, $MS_e = 1.03$, $p < .001$]. This means that Δp for the perfect positive covariate must be higher than Δp for the weak covariate and Δp for the two elements mentioned in the inhibiting interpretation if subjects are integrating prior beliefs with the data set presented to them. (And if subjects are not integrating prior beliefs with the data set, then Δp for the perfect positive covariate is higher than Δp for the weak covariate and at least as high as Δp for the two elements mentioned in the inhibiting interpretation, because that is the pattern in the data sets.) In each case, therefore, the probabilistic contrast model predicts that the simple-cause interpretation will be judged more likely than the inhibiting interpretation.

Mean causal judgments for each scenario, for the selected samples of subjects, are shown in Table 5. For each scenario, judgments of likelihood were analyzed with a one-way ANOVA with repeated measures. For the furniture-shop scenario, the inhibiting interpretation was judged significantly more likely than the simple-cause interpretation [$F(1,29) = 39.75$, $MS_e = 942.22$, $p < .001$]. For the car scenario, the inhibiting interpretation was judged significantly more likely than the simple-cause interpretation [$F(1,24) = 55.98$, $MS_e = 714.34$, $p < .001$]. For the bus scenario, no significant difference was found [$F(1,48) = 0.25$, $MS_e = 1,924.40$].

For two scenarios, the furniture-shop and car scenarios, the results disconfirm the predictions of the probabilistic contrast model. For the bus scenario, the results are inconclusive. There is a nonsignificant trend in the direction opposite to that predicted by the probabilistic contrast model. The probabilistic contrast model predicts a significant preference for the simple-cause interpretation, and this prediction is not supported; however, no inferences can be drawn from a nonsignificant result with any confidence.

There may be one final escape route for the probabilistic contrast model. Taking the furniture-shop scenario as our example, the inhibiting interpretation mentions two things: needing furniture and not having money. Prior beliefs were assessed for needing furniture, but not for not having money. Suppose that subjects have a prior belief that there is very high covariation between having or not having money and buying or not buying furniture. In this case, they could increase the degree of covaria-

Table 5
Mean Causal Judgments for Each Scenario in Experiment 2

Scenario	Interpretation	
	Simple Cause	Inhibiting
Furniture shop	30.20	76.87
Car	33.04	89.60
Bus	52.65	57.10

Note—Scores may range from 0 (*definitely not the right interpretation*) to 100 (*definitely the right interpretation*).

tion between the needing furniture/not having money combination and buying furniture by integrating this prior belief with the information given.

To check on this possibility, we have to assess how high this judged covariation would have to be in order to make the inhibiting interpretation as likely as the simple-cause interpretation. Recall that the subjects were only included if they judged at least as much covariation between being in a furniture shop and buying furniture as they judged between needing furniture and buying furniture. In fact, the mean for being in a furniture shop and buying furniture was substantially and significantly higher than that for needing furniture and buying furniture. To overcome the effect of this difference, the judged covariation between having or not having money and buying or not buying furniture would have to be substantially higher than that for being in a furniture shop and buying furniture. This is not impossible because the mean judged covariation for being in a furniture shop and buying furniture was not very high (4.30 on a 1–7 scale). But it still would not be enough, for two reasons: (1) It ignores the information given in the data set. The more weight subjects give to the information in the data set, the higher their judged covariation between having or not having money and buying or not buying furniture would have to be. Since we do not know how much weight the subjects gave to this, it is impossible to calculate how high the judged covariation would have to be to overcome it. (2) The judged covariation between having or not having money and buying or not buying furniture would have to be substantially higher than that for being in a furniture shop and buying furniture merely to make the causal rating for the inhibiting interpretation *equal* to that for the simple-cause interpretation. But, in fact, the two ratings were not equal: the mean for the inhibiting interpretation was 76.87, and that for the simple-cause interpretation was only 30.20. It is doubtful whether even a prior belief of perfect covariation between the two would be sufficient to bridge the gap between these two means.

In summary, to account for the results, judged covariation between having or not having money and buying or not buying furniture would have to be substantially higher than that for being in a furniture shop and buying furniture, it would have to be even higher than that to overcome whatever weight subjects give to the information in the data set, and it would have to be even higher than that to account for the observed difference between means (76.87 and 30.20). It does not seem possible that

it could be high enough to overcome all of these things. Furthermore, it is very unlikely that subjects would judge high covariation between having or not having money and buying or not buying furniture. This would mean a belief that, among other things, people buy furniture *whenever* they have money, and this does not seem plausible. In conclusion, it is most unlikely that the model could be saved by the incorporation of prior beliefs about covariation between having or not having money and buying or not buying furniture.

If there is some residual doubt in the case of the furniture-shop scenario, there is none in the case of the car scenario. Here, the factor that is equivalent to having money or not having it is the car being in gear or not being in gear. To account for the results in this scenario, the judged covariation between speeding up and being in gear must be substantially higher than that for speeding up and the engine note rising. The problem is that the latter is very close to the extreme of the scale (6.52 on a 1–7 scale). The judged covariation between speeding up and being in gear cannot possibly be sufficiently higher than this to account for the observed results. For the car scenario, then, prior beliefs about the inhibitory factor cannot save the day for the probabilistic contrast model.

GENERAL DISCUSSION

These two experiments have been concerned with two different theoretical traditions in the study of causal judgment: the regularity-based tradition, represented by such models as the probabilistic contrast model, and the concept-based tradition, represented by the causal powers theory. These two traditions have their philosophical roots in, respectively, regularity theories, such as Hume's (1739/1978) radical empiricist theory, and singularist theories, such as the theory of powerful particulars (Harré & Madden, 1975). The fundamental contrast between these two theoretical traditions concerns not so much the existence of different causal roles as the way in which causal roles are defined.

In the tradition of regularity-based theories, causal roles are defined *empirically*. Different models have different ideas of the sort of empirical pattern that serves to identify not only a cause but other particular causal roles (e.g., enabling conditions); however, all of them share the assumption that *some* definite empirical pattern is the defining mark of each causal role. Under the causal powers theory, causal roles are defined *conceptually*. The kind of concept involved is essentially that of the particular function played by a thing or event in bringing about some effect. The causal role of an enabling condition, for example, is (roughly) to be that by virtue of which a thing has the causal power to produce a given effect. These conceptual definitions do not specify a fixed empirical pattern. *A* may be an enabling condition for *X* to have the power to produce a given effect, and *B* may be an enabling condition for *Y* to produce some other effect, but *A* and *B* are not thereby required to stand in the same empirical relation to the respective effects of the

powers they enable. Likewise, *X* may be a cause of effect *e*₁, and *Y* may be a cause of effect *e*₂, but *X* and *Y* are not thereby required to stand in the same empirical relation to their respective effects.

It is by virtue of this that experiments can discriminate between regularity-based models and the causal powers theory. The two experiments reported here have set up factors that would be identified as causes under models in which causal roles are defined empirically. These factors were then judged less likely to be causes of certain effects than were other combinations of factors that would not be identified as causes by empirical definitions. To be specific, in Experiment 1, in one condition within each of three scenarios, an interpretation identifying a constant factor as the cause and a negative covariate as an inhibiting or a blocking condition was significantly preferred to an interpretation identifying a perfect positive covariate as the cause. Furthermore, in two of those conditions, the interpretation identifying the constant factor as the cause and the positive covariate as a condition was also preferred to the interpretation identifying the positive covariate as the cause. The most extreme result was found in the freeze scenario. In the cloud covariate condition, the mean judged likelihood of the positive covariate being the cause of the effect was only 15.84. Under regularity-based models, a figure approaching 100 would be predicted. This contrasts with the mean judged likelihood of the inhibiting interpretation, 90.52. In this interpretation, the positive covariate is not even mentioned and has no causal role. In Experiment 2, the possibility that people might be integrating covariation beliefs (as opposed to causal beliefs) with the data set given and judging likelihood by computing probabilistic contrasts was controlled; in two of three scenarios, a strongly significant trend disconfirming the predictions of regularity-based models—in particular, the probabilistic contrast model—was found. These results tell decisively against regularity-based models. Despite the elegance and appeal these models can have, and despite the weight of supportive evidence from past research, they are fundamentally wrong.

This is not meant to imply that perfect positive covariates will never be identified as causes under the causal powers theory. In many cases, the function of a cause will be held by something that happens to show strong covariation with an effect. What is implied by these results is that, when a positive covariate of an effect is identified as its cause, it is so identified not because of its empirical association with the effect but because of the function that is attributed to it. Many experiments have shown subjects behaving in roughly the way that regularity-based models predict (e.g., Cheng & Novick, 1990, 1992; Shanks & Dickinson, 1987), although the predictions do not usually account for a great deal of the variance in judgments (Iacobucci & McGill, 1990). Such experiments do not address the contest between regularity-based models and concept-based models, however, because the experimenters did not assess whether these kinds of models would make

different predictions for the stimuli used: they were not discriminative tests. Part of the aim of the present research was to set up suitable discriminative tests, and the results have favored the concept-based approach over the regularity-based approach. Indeed, the effects of scenario content found in Experiment 1 suggest that the support found for regularity-based models in other studies may depend to a considerable extent on the particular scenarios and content employed in those studies.

The causal powers theory takes the notion of the generative relation as its starting point (see Note 2) and shows how an organized set of causal concepts is constructed around it. This approach is supported by the findings of several studies showing that the cue of generative relations is primary in causal inference (Shultz, 1982; Shultz, Fisher, Pratt, & Rulf, 1986). Generative relations cues are used by children as young as 2 years old (Shultz, 1982) and by people from a non-Western cultural background (Shultz, 1982). They are used to an equal degree by young children and adults (Shultz, 1982). That research included a number of discriminative tests in which generative relations cues were compared with covariation cues (and cues of other kinds). Results consistently showed a strong preference for generative relations cues over covariation cues (Shultz, 1982; Shultz et al., 1986); indeed, Shultz et al. proposed that cues other than generative relations cues would be used only when generative relations cues were not available, and their findings were consistent with this claim. These findings lend support to one of the central tenets of the causal powers theory, concerning the kind of understanding of the causal relation that people possess, although they have been contested by Cheng (1993). It is beyond the scope of this discussion to explore the causal powers theory in much depth (see White, 1989, 1992, 1993). To what extent, though, do the present findings support the theory? There are two main points to be made.

First, under the causal powers theory, causal judgment is made mainly by the application of beliefs to instances, unless generative relations cues are available. People develop and possess an extensive range of causal beliefs that they can use in application to instances. It is the basic causal concepts of the generative relation, coupled with causal power, releasing condition, and liability or resistance, that serve to identify a belief as causal and to distinguish it from all noncausal beliefs. Under this theory, therefore, information content is preferred over empirical cues as a basis for causal judgment, insofar as that content relates to existing causal beliefs. The theory therefore predicts strong effects of content where causal judgments can be made by application of beliefs. This prediction is supported by the results of Experiment 1, in which strong effects of content were found in each scenario. Experiment 2 showed that effects of content cannot be accounted for in terms of prior beliefs about covariation between the things involved.

Second, the basic concepts of the theory distinguish several roles that events and factors may play in causal

relations: power, releasing condition, liability, and inhibiting condition. Any given instance of a causal relation may involve events and factors taking on any or all of these roles. It is therefore important in experiments to allow people to assign multiple causal roles (i.e., to identify different factors and events as playing different parts in the production of an effect). Most tests of regularity-based theories have not permitted people to make causal judgments that reflect their assignment of multiple causal roles in the interpretation of events, but they have allowed them to say only whether one or another candidate was a cause of a given effect or that some unspecified kind of interaction between candidates was the cause. Findings of research on causal judgment cannot be deemed to have much validity if subjects are not permitted to make and express multiple causal role interpretations. This is the general force of the critical evaluation of Wasserman's (1990) study in the introduction. The results of Experiment 1 support this line of reasoning by showing that subjects do make interpretations involving the assignment of different causal roles to different elements. Perhaps most relevant, two of the interpretations offered in each condition involved identifying the constant factor as the cause of an effect, but the subjects showed significant preferences for one over the other (the particular preference shown varied with scenario content), depending on the conditional role assigned to some other factor. This clearly shows that people interpret causal relations as involving not just a cause and an effect but also conditions relating to the operation of the causal factor, as would be predicted under the causal powers theory.

The causal powers theory is unlikely to be the only possible theory under which such judgmental phenomena would be predicted, although it seems to be the only one currently in the literature. There are other theoretical approaches to causal judgment not based on the idea of regularity, such as that by Einhorn and Hogarth (1986), but these approaches have little to say about causal roles (White, *in press*.) The present results therefore support, but could not be said to confirm, the theory. It would therefore appear to be a priority for future research to propose alternative theories to compete with it. It must also be pointed out that it is difficult to make precise or quantitative predictions under the causal powers theory because of the major role played by particular beliefs. If one could obtain measures of causal beliefs and how strongly they are held, then more exact predictions based on information gained about those beliefs might be possible. It may be thought to be an advantage of regularity-based models that quantitative predictions can be derived from them. The force of the critique presented here, though, is that this capacity is bought at the cost of having unacceptably tight methodological constraints. When these constraints are loosened just a little—for example, by allowing multiple causal role interpretations—the ability of regularity-based models to make quantitative predictions is lost,

even when idealized covariation matrices are presented for judgment.

Wasserman (1990) drew a parallel between his results and those of studies of animal learning, as already stated. The present results do not rule out the possibility that such a parallel may be valid, though whether as homology or analogy remains uncertain. In that study, however, the appearance of support for the analogy with animal learning is forced upon us by the fact that subjects were not permitted to assign causal roles to the elements in the stimuli. When subjects are allowed to assign causal roles, the similarity of their judgments to animal learning phenomena depends on the beliefs they hold about the elements in the scenario. On the other hand, studies have shown other interesting parallels between human causal judgment and phenomena in animal learning (Shanks, 1993), and there is still much to learn about the interpretation of such parallels. It is possible, though, that similarities can emerge from the operation of different processes. Humans possibly differ from other species in that they apply concepts in the interpretation of events. The present study suggests that people interpret empirical patterns in the light of a sophisticated conceptual system that enables them to assign distinct causal roles to the different components. It seems implausible to suggest that people are using concepts when their behavior differs from that of other species in analogous paradigms (as, for example, in the low-belief covariate condition of the freeze scenario) but that they are not using concepts when their behavior is similar to that of other species (as in the high-belief covariate condition of the freeze scenario and in Wasserman's study). It is surely more likely that humans are using their causal concepts all the time and that, under some circumstances, their use of those concepts, and their particular acquired causal beliefs, leads them into patterns of judgment that happen to resemble the learning behavior of other species. This is, however, for future research to settle.

REFERENCES

- ALLOY, L. B., & ABRAMSON, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, **108**, 441-485.
- BUNGE, M. (1963). *Causality: The place of the causal principle in modern science*. Cambridge, MA: Harvard University Press.
- CHENG, P. W. (1993). Separating causal laws from causal facts: Pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 215-264). San Diego: Academic Press.
- CHENG, P. W., & NOVICK, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality & Social Psychology*, **58**, 545-567.
- CHENG, P. W., & NOVICK, L. R. (1991). Causes versus enabling conditions. *Cognition*, **40**, 83-120.
- CHENG, P. W., & NOVICK, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, **99**, 365-382.
- EINHORN, H. J., & HOGARTH, R. M. (1986). Judging probable cause. *Psychological Bulletin*, **99**, 3-19.
- FORSTERLING, F. (1989). Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality & Social Psychology*, **57**, 615-625.
- HARRÉ, R., & MADDEN, E. H. (1975). *Causal powers: A theory of natural necessity*. Oxford: Blackwell.
- HARRÉ, R., & SECORD, P. F. (1972). *The explanation of social behaviour*. Oxford: Blackwell.
- HEWSTONE, M., & JASPARS, J. (1987). Covariation and causal attribution: A logical model of the intuitive analysis of variance. *Journal of Personality & Social Psychology*, **53**, 663-672.
- HILTON, D. J., & SLUGOSKI, B. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, **93**, 75-88.
- HUME, D. (1978). *A treatise of human nature*. Oxford: Oxford University Press. (Original work published 1739)
- IACOBucci, D., & MCGILL, A. L. (1990). Analysis of attribution data: Theory testing and effect estimation. *Journal of Personality & Social Psychology*, **59**, 426-441.
- KELLEY, H. H. (1967). Attribution in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.
- KELLEY, H. H. (1973). The processes of causal attribution. *American Psychologist*, **28**, 107-128.
- SCHUSTACK, M. W., & STERNBERG, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, **110**, 101-120.
- SHANKS, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, **84**, 319-354.
- SHANKS, D. R., & DICKINSON, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 229-261). New York: Academic Press.
- SHULTZ, T. R. (1982). *Rules of causal attribution*. *Monographs of the Society for Research in Child Development*, **47**, 1-51.
- SHULTZ, T. R., FISHER, G. W., PRATT, C. C., & RULF, S. (1986). Selection of causal rules. *Child Development*, **57**, 143-152.
- WAGNER, A. R., LOGAN, F. A., HABERLANDT, K., & PRICE, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, **76**, 171-180.
- WASSERMAN, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology*, **103**, 284-297.
- WASSERMAN, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science*, **1**, 298-302.
- WHITE, P. A. (1989). A theory of causal processing. *British Journal of Psychology*, **80**, 431-454.
- WHITE, P. A. (1990). Ideas about causation in philosophy and in psychology. *Psychological Bulletin*, **108**, 3-18.
- WHITE, P. A. (1992). Causal powers, causal questions, and the place of regularity information in causal attribution. *British Journal of Psychology*, **83**, 161-188.
- WHITE, P. A. (1993). *Psychological metaphysics*. London: Routledge.
- WHITE, P. A. (in press). *Understanding of causation and the production of action: From infancy to adulthood*. Hove, U.K.: Erlbaum.

NOTES

1. The term *regularity-based* is used in this paper to refer to any theory or model of causal judgment or attribution based on the idea of empirical regularity of association between cause and effect.
2. Generative or productive relations refer to the notion that a cause actually generates or produces its effect. More specifically, the terms refer to the transmission of energy from the cause to that in which the effect occurs. In the example of the hammer smashing the plate, the generative relation refers to the transfer of energy from hammer to plate at the moment of contact, by virtue of which the effect of smashing is produced. Several philosophical theories of causation have incorporated this notion in some form (Harré & Madden, 1975; White, 1990), and several experiments by Shultz (1982) and Shultz et al. (1986) have found that cues to generative relations are preferred over other possible cues for causal judgment, by adults and by children as young as 2 years old.
3. Wasserman found no important differences in results between the two methods of presentation he used.

4. The difference between a releasing condition and an enabling condition is a technical matter (see Harré & Madden, 1975). For purposes of these scenarios, however, the difference is unimportant, since the intent is only to identify a factor that in some way facilitates the production of the effect by the constant factor, and releasing conditions and enabling conditions both have this general function.

5. These analyses were carried out using the pooled error term, MSW.cell, for simple effects analysis of the between-subjects factor.

6. Because the items in each pair are, or tend to be, related by unidirectional contingency, it was important to ensure that the subjects were not making their judgments merely by assuming the presence of one item and judging the occurrence of the other.

APPENDIX

Calculation of Probabilistic Contrasts for Experiment 2

This shows the calculation of probabilistic contrasts for the furniture-shop scenario. Proportions for the other scenarios are the same, yielding the same calculation of probabilistic contrasts, though the actual frequencies presented to the subjects differed somewhat. All references to line numbers refer to Table 4.

Main Effect Contrast for Being in a Furniture Shop

For p_i , look at occurrences and nonoccurrences of the effect in the presence of being in a furniture shop. This is given in line 1 and shows the effect occurring on 10 out of 10 occasions. Therefore, $p_i = 1$.

For p_{-i} , look at occurrences and nonoccurrences of the effect in the absence of being in a furniture shop. This is given in lines 2 and 3 and shows the effect occurring on 0 out of 30 occasions. Therefore, $p_{-i} = 0$.

Therefore, $\Delta p_i = 1 - 0 = +1.0$.

Main Effect Contrast for Needing New Furniture

For p_j , look at occurrences and nonoccurrences of the effect in the presence of needing furniture. This is given in lines 1 and 2 and shows the effect occurring on 10 out of 30 occasions. Therefore, $p_j = 0.33$.

For p_{-j} , look at occurrences and nonoccurrences of the effect in the absence of needing furniture. This is given in line 3 and shows the effect occurring on 0 out of 10 occasions. Therefore, $p_{-j} = 0$.

Therefore, $\Delta p_j = 0.33 - 0 = +0.33$.

Main Effect Contrast for Not Having Money

For p_k , look at occurrences and nonoccurrences of the effect in the presence of not having money. This is given in line 2 and shows the effect occurring on 0 out of 20 occasions. Therefore, $p_k = 0$.

For p_{-k} , look at occurrences and nonoccurrences of the effect in the absence of not having money (i.e., the presence of having money). This is given in lines 1 and 3 and shows the effect occurring on 10 out of 20 occasions. Therefore, $p_{-k} = 0.5$.

Therefore, $\Delta p_k = 0 - 0.5 = -0.5$.

The sign depends simply on whether p_k is defined as the presence of having money or the presence of not having money. In this case, we are interested in not having money as an inhibitory factor, so p_k is defined as the presence of not having money. If we had been interested in the effect of having money, then p_k could be defined as the presence of having money; in this case, Δp_k for having money would be $+0.5$.

Interactive Contrast for the Combination of Needing Furniture (i) and Not Having Money (k)

For p_{ik} , look at occurrences and nonoccurrences of the effect in the presence of both. This is given in line 2 and shows the effect occurring on 0 out of 20 occasions. Therefore, $p_{ik} = 0$.

For p_{i-k} , look at occurrences and nonoccurrences of the effect in the presence of needing furniture and the absence of not having money (i.e., in the presence of having money). This is given in line 1 and shows the effect occurring on 10 out of 10 occasions. Therefore, $p_{i-k} = 1$.

Note that p_{-ik} is the combination of the absence of needing new furniture and the presence of not having money. This combination is not given in Table 4.

For p_{-i-k} , look at occurrences and nonoccurrences of the effect in the absence of needing furniture and the absence of not having money (i.e., in the presence of having money). This is given in line 3 and shows the effect occurring on 0 out of 10 occasions. Therefore, $p_{-i-k} = 0$.

As stated in the text, the absence of a value for p_{ik} means that the interactive contrast cannot actually be calculated for Experiment 2. For the sake of an example, let us assume that $p_{-ik} = 0$. Then, $\Delta p_{ik} = (0 - 0) - (1 - 0) = -1.0$. This identifies needing furniture and not having money as an inhibitory factor. Further examples can be found in Appendix B of Cheng and Novick (1990).