European Journal of Social Psychology
Eur. J. Soc. Psychol. **31**, 155–175 (2001)

When more observations are better than less: a connectionist account of the acquisition of causal strength

FRANK VAN OVERWALLE* AND
DIRK VAN ROOY
Vrije Universiteit Brussel, Belgium

Abstract

The statistical law of large numbers prescribes that estimates are more reliable and accurate when based on a larger sample of observations. This effect of sample size was investigated on causal attributions. Subjects received fixed levels of consensus and distinctiveness covariation, and attributions were measured after a varying number of trials. Whereas prominent statistical models of causality (e.g. Cheng & Novick, 1990; Försterling, 1992) predict no effect of sample size, adaptive connectionist models (McClelland & Rumelhart, 1988) predict that subjects will incrementally adjust causal ratings in the direction of the true covariation the more observations are made. In three experiments, sample size effects were found consistent with the connectionist prediction. Possible extensions of statistical models were considered and simulated, but none of them accommodated the data as well as connectionist models. Copyright © 2001 John Wiley & Sons, Ltd.

It is a common observation that people's judgments are often made with more confidence and accuracy when they are drawn from many cases rather than only a few. This aspect of everyday reasoning is paralleled by the statistical law of large numbers, which prescribes that our evaluations 'should be more confident when they are based on a larger number of instances' (Nisbett, Krantz, Jepson, & Kunda, 1993, p. 339). For instance, when making causal inferences, people make more extreme judgments after receiving more information (Baker, Berbier, & Vallée-Tourangeau, 1989; Försterling, 1992; Shanks, 1985, 1987, 1995; Shanks, Lopez, Darby, & Dickinson, 1996). Similarly, when receiving more supportive information, people tend to hold more extreme impressions about other persons (Anderson, 1967, 1981), make more polarized group decisions (Fiedler, 1996; Ebbesen & Bowers, 1974), endorse more firmly an hypothesis (Fiedler, Walther, & Nickel, 1999), make more extreme predictions (Manis, Dovalina, Avis, & Cardoze, 1980) and agree more with persuasive messages (Eagly & Chaiken, 1993). The converging evidence in these different areas demonstrates that people have an intuitive appreciation of the law of large numbers and seem to recognize that when making inferential judgments, more evidence is better than less.

*Correspondence to: Frank Van Overwalle, Department of Psychology, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium. E-mail: Frank.VanOverwalle@vub.ac.be

Received 2 May 1999
Accepted 26 June 2000

Do people also obey the law of large numbers when they make causal judgments of social events? Most studies that documented an effect of sample size on causal explanations were conducted in the domain of experimental psychology and typically used free-operant tasks that contained little social material (Baker *et al.*, 1989; Shanks, 1985, 1987, 1995; Shanks *et al.*, 1996). On the other hand, in the social attribution literature, the topic of sample size has received remarkably little attention. We are aware of only one study by Försterling (1992) who, unfortunately, did not manipulate sample size independently of covariation. Perhaps more surprising is that many major attribution models developed during the last 30 years in social psychology ignore the question of sample size altogether (e.g. Cheng & Novick, 1990; Försterling, 1989; Hewstone & Jaspars, 1987; Hilton & Slugoski, 1986; Kelley, 1967; Orvis, Cunningham, & Kelley, 1975; Read & Marcus-Newhall, 1993; but see Försterling, 1992). Even more recent theories which explicitly claim to describe the 'attribution process' (Försterling, 1989, p. 624) or 'causal induction' (Cheng & Novick, 1990, p. 549) are silent with respect to this issue and are only capable of accounting for asymptotic performance, that is, after learning has consolidated given a sufficient number of trials. Contrary to intuition, they imply that the number of observations should *not* affect people's causal estimates.

Recently, however, some models have been proposed that do take into account the size and order of causal information. These models either introduced an updating rule to previous statistical approaches (Hogarth & Einhorn, 1992; Busemeyer, 1991; see also Wasserman, Kao, Van Hamme, Katagiri, & Young, 1996) or are based on novel connectionist principles (Van Overwalle, 1998; Read & Montoya, 1999). Given these novel theoretical developments, the aim of this article is, first, to examine whether sample size is a reliable phenomenon in causal attribution in the social domain, and second, to evaluate how well existing models can account for this effect. Before introducing these novel approaches in some more detail, we will first briefly explain why mainstream attribution models are insensitive to sample size.

STATISTICAL ATTRIBUTION MODELS

The dominant, rational view on the attribution process in social psychology is that people are intuitive statisticians who extract covariation information by tallying frequencies and applying some kind of a rule on them (e.g. Cheng & Novick, 1990; Försterling, 1989, 1992; Hewstone & Jaspars, 1987; Hilton & Slugoski, 1986; Kelley, 1967). It is assumed that perceivers tally four major types of evidence: when the cause is present and the effect is (*a*) present or (*b*) absent, and when the cause is absent and the effect is (*c*) present or (*d*) absent. A rule is then applied on all or some of these four types of frequency (*a–d*) to estimate the degree of covariation and causality.

One of the most popular rules was formalized in the probabilistic contrast model of Cheng and Novick (1990), and 'requires that people ... estimate and compare proportions' (p. 549). Because this probabilistic contrast rule closely parallels statistical measures such as χ^2, it has become widely accepted as the normative standard for computing the covariation between a cause and an effect. Other valuable statistical accounts are the Bayesian model proposed by Fales and Wasserman (1992, see also Anderson & Sheu, 1995), and Försterling's (1992) formulation of Kelley's ANOVA model that is based on the statistical parameter of effect size (i.e. η^2).

A key element of statistical models is that the proposed rule is not directly based on raw frequencies but rather on internal proportions between these frequencies. For this reason, most of these models have difficulties explaining the effect of sample size. This can be illustrated with the following example. Take whatever values for the four types of frequencies *a–d* as discussed earlier. Now keep these frequencies fixed and repeat them a number of times. Although the overall frequencies increase,

obviously, their internal proportions remain the same. Many statistical models therefore predict that the level of covariation and causality should remain identical, no matter how often these frequencies are repeated. This limitation exists for statistical models such as the probabilistic contrast model (Cheng & Novick, 1990), the Bayesian model[1] (Anderson & Sheu, 1995; Fales & Wasserman, 1992), and the ANOVA formulation (Försterling, 1992).

JUDGMENT-UPDATING MODELS

Recently, a number of models have been proposed that avoid these limitations of earlier statistical approaches by incorporating some form of updating rule that make them sensitive to sample size. These models are the step-by-step belief-adjustment model of Hogarth and Einhorn (1996) and the serial averaging strategy of Busemeyer (1991) that extends Anderson's (1981) impression-formation theory into the domain of attribution. However, a major restriction is that the proposed rule involves only a single cause and does not take into account the influence of alternative causes (e.g. discounting and augmentation). In addition, it is easy to show that these models are mathematically identical to a simplified version of the delta algorithm applied in connectionist models, one that deals with only one cause at the time (Wasserman *et al.*, 1996, Appendix D). Because these models can be considered special instances of the more general delta algorithm, we will ignore them for the most part and immediately turn to the connectionist approach.

A CONNECTIONIST PERSPECTIVE

A radically different approach to causal reasoning is based on connectionist principles. This perspective assumes that causal judgments are represented in memory as connections or links between the representations of causes and effects. The strength or weight of this connection reflects the perceived influence of the cause on the effect. In adaptive connectionist network models, these cause–effect connection weights vary on the basis of the evidence (Allan, 1993; McClelland & Rumelhart, 1988; Read & Montoya, 1999; Shanks, 1995; Smith, 1996; Van Overwalle, 1998).

One of the most widely accepted learning algorithms that governs this weight adjustment is the *delta algorithm* (McClelland & Rumelhart, 1988). This algorithm has been applied in many investigations on human categorization (see Gluck & Bower, 1988a,b; Shanks, 1991) and causality (for reviews, see Allan, 1993; Van Overwalle & Van Rooy, 1998; Shanks, 1987, 1995), and is formally identical to the Rescorla–Wagner (1972) model of animal conditioning.

To simplify the exposition of the delta algorithm, we focus here on its working in a minimalist but powerful adaptive connectionist model, known as the *feedforward* network (Van Overwalle, 1998), because this network displays the sample size effect much like more complex connectionist networks do (e.g. Read & Montoya, 1999). A feedforward network consists of two layers of nodes. Each potential cause is represented by an *input node* that is connected by an adjustable weight to an *output*

[1]The argument that Bayesian reasoning is not influenced by sample size is based on the assumption – shared by most models – that perceivers evaluate to what extent the causal hypothesis can explain all the available data (i.e. with a probability of 1). As noted by Fales and Wasserman (1992), this assumption can be relaxed by assuming that perceivers simultaneously entertain a number of causal hypotheses with mutually exclusive degrees of causal probability, for instance eleven hypotheses reflecting a probability of 0, 0.1, 0.2, 0.3, and so forth to 1. Because this assumption is psychologically very implausible as it puts a great burden on human information processing, and because it is not required by any of the other models discussed, this extension is not considered in this article.

node. Every time a cause is present, its corresponding input node is activated, and then this activation spreads automatically to the output node in proportion to the weight of the connection. The term feedforward reflects the assumption of this particular model that activation is spread only from input to output nodes. All activations received at the output node are linearly summed to determine the output activation, and this activation reflects the effect predicted by the network on the basis of the causal input given. Note that this summation of all causal inputs makes the delta algorithm much more general and powerful than judgment-updating models (Hogarth & Einhorn, 1996; Busemeyer, 1991) that have taken only a single cause into account.

The key assumption of the delta algorithm is that it attempts to reduce the error between the effect predicted by the output activation and the actual effect that occurred, and it does so by adjusting the weights of the cause–effect connections. For instance, when the occurrence of the effect is underestimated, the weights are adjusted upward; and when the occurrence of the effect is overestimated, adjustments are made downward. The adjustments are made in proportion to the magnitude of the error, so that larger errors result in larger adjustments and faster causal learning (for more details on this delta-processing mechanism, see Van Overwalle, 1998, pp. 314–317).

The delta algorithm reflects several types of processes that are well known. First, the slow incremental adjustments of cause–effect connections seem to capture gradual causal learning from direct experiences, in contrast to fast routine explanations in everyday life which are typically based on prior knowledge, that is, on the reactivation of learned connections in memory. Second, most learning in the network occurs when the discrepancy between what is expected and what actually happens is large. This is consistent with research showing that an unexplained effect (e.g. a disconfirmed expectancy or goal; Weiner, 1985) instigates considerable attributional activity that often leads to changes of causal beliefs. Third, although the algorithm does not tally frequencies nor computes statistical probabilities, it has been shown mathematically that it forces the weights to converge to Cheng and Novick's (1990) probabilistic contrast norm after a sufficient number of observations (Chapman & Robbins, 1990; Van Overwalle, 1996). Thus, the algorithm reproduces the statistical properties of covariation. The delta algorithm is also able to account for other familiar phenomena like discounting and augmentation (Van Overwalle & Van Rooy, 1998).

Most importantly for the present purpose, sensitivity to sample size is incorporated into the delta algorithm. This can be clearly seen if we set the initial connection weights to a moderate starting value (e.g. the midpoint of the response scale). If the actual cause–effect covariation is very high, the moderate starting weight produces an underestimation of the actual effect. This error is gradually eliminated with increasing observations as each cycle through the network adjusts the connection upward toward the true (high) covariation. The same logic holds for low cause–effect covariation, which results in an overestimation and downward adjustments. Thus, the statistical law of large numbers is a natural emergent property of the delta learning algorithm.

DESIGN AND HYPOTHESES

We have argued that people's judgments become more accurate given a growing sample size, an effect which is not anticipated by many statistical attribution formulations, but which is predicted by an adaptive connectionist approach. Three experiments with a common design were devised to explore this sample size prediction. The design exploits the idea from our earlier example that simply repeating identical frequencies of information will increase sample size, but will leave unaffected statistical measures of covariation between cause and effect. This unconfounds the effect of covariation (which is kept fixed) from that of sample size (which varies).

We focus on consensus and distinctiveness covariation because they are among the prime social sources from which people infer causal attributions (cf. Kelley, 1967). Consensus refers to the extent to which behaviors or outcomes of an actor generalize to other, similar, actors, whereas distinctiveness refers to the extent to which outcomes given a stimulus do *not* generalize to other, similar, stimuli. High covariation of an actor is implied given low consensus (i.e. only this actor behaved in this manner) and high covariation of a stimulus is implied given high distinctiveness (i.e. the behavior occurred only with this stimulus). Conversely, low covariation is implied given the reversed patterns of high consensus or low distinctiveness. After receiving this covariation information, subjects rated the causal influence of the actor or stimulus. To examine the effect of sample size, this rating was repeated after a varying number of trials.

As noted above, existing attribution models including the probabilistic (Cheng & Novick, 1990), Bayesian (Anderson & Sheu, 1995; Fales & Wasserman, 1992) and ANOVA models (Försterling, 1992) predict that for each fixed level of covariation, the causal ratings should be identical no matter how many trials are given. In contrast, our connectionist hypothesis is that subjects will incrementally adjust their ratings in the direction of the true covariation as more evidence is given. Thus, the earlier statistical formulations essentially predict flat acquisition curves, whereas we predict increasing curves given high covariation and decreasing curves given low covariation.

EXPERIMENT 1

In the first experiment, subjects received two different levels of covariation (0% or 100%) across six blocks of trials. After each block, they rated the causal contribution of the target cause. Thus, the design of the experiment consisted of two within-subjects factors Covariation (0% or 100%) and Block (1 to 6).

Method

Subjects

Subjects were 97 male and female freshmen from the Dutch-speaking Vrije Universiteit Brussel, who participated for a partial requirement of an introductory psychology course. They were tested in groups of one to four.

Materials

All instructions, materials and questions appeared on an IBM-compatible PC screen, and the whole experiment was monitored by MEL software. Each subject read four experimental event descriptions and four filler descriptions. Two experimental events depicted high- and low-consensus information involving a target actor and other comparison actors (with the same stimulus). Another two experimental events depicted high- and low-distinctiveness information involving a target stimulus and other comparison stimuli (with the same actor). Each event described whether the target's outcome was present or not. (To ensure correct encoding of this information, the negative wording of an outcome was always given in capitals.)

Sample Size Manipulation. An event description was broken down in six blocks of two trials, or 12 trials in total. The trials were presented one after the other and described as consecutive points in time. After each block of two trials, subjects had to give ratings on the target cause. Throughout all blocks of

an event, the name of the target actor or stimulus remained the same, while the name of the comparison actors or stimuli differed at each block (to induce comparison with a range of comparison actors or stimuli as is customary in social attribution research).

To illustrate, an event involving consensus manipulation of a target actor, Helen, with a comparison actor, Eva, is shown here for the first two trials. The outcome of the comparison actor was either identical or opposite (as indicated by square brackets) depending on the covariation condition:

- Helen liked the perfume.
- Eva liked [did NOT like] the perfume.

Event descriptions involving distinctiveness manipulation of a target stimulus, Corinne, with a comparison stimulus, Karen, were constructed in a similar manner as illustrated below for the first two trials:

- Jasmine deceived her friend, Corinne
- Jasmine deceived [did NOT deceive] her friend, Karen

Level of Covariation. By varying whether the outcome of the target actor (or stimulus) was also obtained by the comparison actors (or stimuli), covariation between target and outcome was manipulated. An illustrative summary is shown in Table 1. Conforming to probabilistic theory (Cheng & Novick, 1990), an outcome that occurs only with the target and not with comparison cases (i.e. low consensus or high distinctiveness) is denoted by 100% covariation, while an outcome that is the same for both the target and comparison cases (i.e. high consensus or low distinctiveness) is denoted by 0% covariation. The level of covariation remained fixed across all blocks of an event.

The experimental events were evenly distributed over all consensus and distinctiveness conditions for each subject. The order in which they appeared was randomized for each subject, as well as the order of the trials within each block. The number of action versus state verbs was counterbalanced between the events to control for potential effects of implicit causality (Rudolph & Försterling, 1997).

Procedure

Instructions appeared on the screen and the use of the rating scale was practiced. Each subject then read an event description. After each block, subjects had to rate how much influence *something special about*

Table 1. Design of Experiment 1 illustrated for the event 'Els detested the singer Karel'

		Covariation (consensus)	
Trial type	Actor + stimulus	100% (Low Cs)	0% (High Cs)
Target	*Els* + Karel	Detested	Detested
Comparison	Ilse + Karel	No	Detested
		Covariation (distinctiveness)	
Trial type	Actor + stimulus	100% (High Di)	0% (Low Di)
Target	Els + *Karel*	Detested	Detested
Comparison	Els + Peter	No	Detested

Note: The target is in italics; Cs = consensus; Di = distinctiveness; the same information in each covariation condition was repeated over six consecutive blocks with different comparison names.

[target actor/stimulus] had on the outcome, using an 11-point rating scale ranging from 0 (*absolutely no influence*) to 100 (*very strong influence*), with midpoint 50 (*partial influence*). Subjects indicated their answer by moving through the scale points in steps of 10, using the left and right arrow keys. For example, in order to explain why 'Helen liked the perfume', subjects had to judge the influence of *something special about Helen*. Similarly, in order to explain why 'Jasmine deceived her friend, Corinne', subjects had to judge the influence of *something special about Corinne*.

Results

Because we made the same predictions for consensus and distinctiveness covariation, the ratings for the actor and stimulus target were collapsed and analyzed together. This seemed justified as preliminary multivariate analyses of variance (MANOVAs) with Measure (actor or stimulus), Covariation (0% or 100%) and Block (1 to 6) revealed that Measure did not interact with any factor. We also collapsed the target ratings in all subsequent experiments, because similar preliminary MANOVAs revealed almost no interactions with Measure.[2]

The mean attribution ratings in function of covariation condition and trial number are depicted in Figure 1. As expected, the ratings showed an increase over trials in the 100% covariation condition, and a decrease over trials in the 0% condition.

A repeated-measures analysis of variance (ANOVA) with Covariation (0% and 100%) and Block (first versus last) as within-subjects factors revealed a significant main effect of Covariation, $F(1,78) = 48.01$, $p < 0.0001$, indicating that subjects were very sensitive to the different levels of covariation. The main effect of Block was not significant, $F(1,78) = 1.98$, $p = 0.163$. More importantly, as expected, the interaction between Covariation and Block was significant, $F(1,78) = 35.60$, $p < 0.0001$.

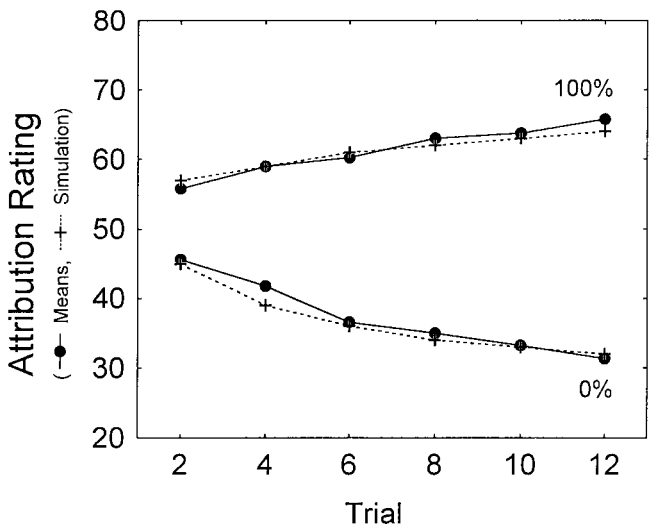

Figure 1. Experiment 1: attribution ratings and feedforward simulation in function of covariation level and number of trials (100% = low Cs/high Di; 0% = high Cs/low Di)

[2] There was only one significant interaction between Measure and Block in Experiment 2, $F(3, 66) = 3.27$, $p < 0.05$, but the more relevant triple interaction including Covariation did not reach significance, $F(3, 66) = 1$, $p < 0.904$.

We predicted that sample size adjustments should occur for all levels of covariation. To verify this, we explored the interaction further by conducting trend analyses in each covariation condition. The delta algorithm would predict a substantial linear trend indicating an increase or decrease across blocks, together with a smaller quadratic trend indicating that the linear change becomes smaller toward the final rating (i.e. close to asymptote). In line with this prediction, there was a significant linear increase in the 100% condition, $F(1,78) = 10.17$, $p < 0.002$, and a linear decrease in the 0% condition, $F(1,78) = 44.72$, $p < 0.0001$. There was also a smaller quadratic trend in the 0% condition, $F(1,78) = 4.77$, $p < 0.05$, but not in the 100% condition ($F < 1$). These results confirm our prediction that the attribution ratings were adjusted over trials as described by connectionist models.

Discussion

Our findings demonstrate that there was a reliable tendency for subjects to progressively adjust their causal ratings towards the true level of covariation with the target as more evidence was provided that confirmed their initial judgments. Given high covariation, judgments were adjusted upward; and given low covariation, judgments were adjusted downward, showing an increasing sensitivity to covariation given a larger sample size. This supports the connectionist model and contradicts current statistical models.

EXPERIMENT 2

Although the previous results suggest that subjects adjust their causal judgments in view of a growing sample size, a possible limitation of our procedure is that repeated ratings were requested intermittently while reading new information about the same event. It is unclear what effect this might have had on our subjects, as it may have induced a demand to change responses from trial to trial, something subjects would perhaps never do when they make spontaneous causal inferences. Moreover, this procedure is not very typical of attribution research.

To address this issue, in the next experiment we employed a design in which subjects were asked to make judgments only once at the end of the trials. The design consisted of the same within-subjects factors Covariation (0% or 100%) and Block (1 to 4) as in the previous experiment, although now the number of trials was manipulated between (rather than within) events. Our hypothesis is that subjects are sensitive to sample size even with this less contrived and demand-inducing procedure.

Method

Subjects

Subjects were 69 male and female freshmen from the Dutch-speaking Vrije Universiteit Brussel, who participated for a partial requirement of an introductory psychology course.

Materials and Procedure

All instructions, materials, and questions were similar as the previous experiment, with the following modifications. Subjects read eight stories without fillers (selected randomly from a larger pool of sixteen stories). Instead of interrupting the presentation of covariation information by repeated ratings,

the information was presented in one whole block with varying number of trials, and attribution ratings were requested at the end of the whole block. These blocks contained either two, four, six, or eight trials. This varying number of trials was not announced.

One other difference between the experiments deserves mentioning. Specifically, the target case in the present procedure was not repeated in each block, unlike the previous experiment where it was repeated together with the comparison case. This was done to make the task more natural and because information was not interrupted by causal ratings. As a consequence, there were more comparison cases given the same total amount of trials, that is, Blocks 1 to 4 contained respectively one, three, five and seven comparison cases, whereas they contained only one, two, three and four comparison cases in Experiment 1. This may have facilitated the speed by which subjects adjusted their ratings.

Results

The mean attribution ratings in function of covariation condition and trial number are depicted in Figure 2. As expected, the target ratings showed an increase over trials in the 100% covariation condition, and a decrease over trials in the 0% condition, although the ratings appear to reach asymptote more rapidly than in the previous experiment.

A repeated-measures ANOVA with Covariation (0% and 100%) and Block (smallest versus largest) as within-subjects factors revealed a significant main effect of Covariation, $F(1, 68) = 135.46$, $p < 0.0001$, indicating that subjects were very sensitive to the different levels of covariation. The main effect of Block was not significant, $F(1, 68) < 1$, *ns*. More importantly, as expected, the interaction between Covariation and Block was significant, $F(1, 68) = 26.93$, $p < 0.0001$.

This interaction was further explored by conducting trend analyses in each covariation condition. The results revealed a significant linear increase in the 100% condition, $F(1, 68) = 12.88$, $p < 0.001$, as well as a linear decrease in the 0% condition, $F(1, 68) = 14.95$, $p < 0.0001$, together with smaller quadratic trends in the two conditions, Fs$(1, 68) = 6.60$–7.84, $p < 0.02$. These results confirm that causal ratings were sensitive to sample size as predicted by the delta algorithm.

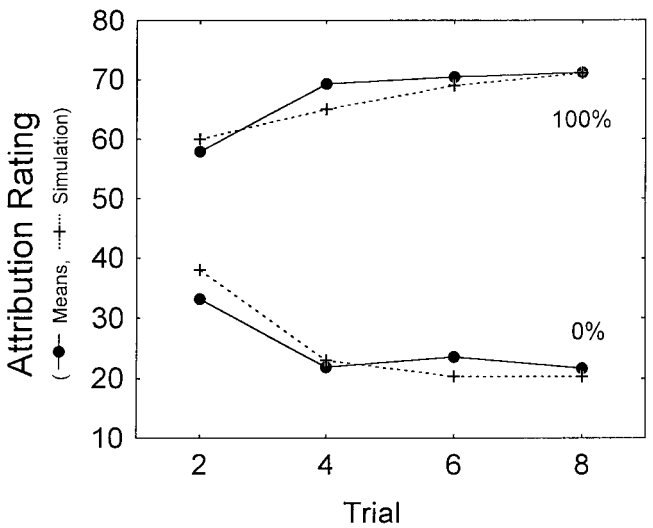

Figure 2. Experiment 2: attribution ratings and feedforward simulation in function of covariation level and number of trials (100% = low Cs/high Di; 0% = high Cs/low Di)

Discussion

The results are consistent with the previous experiment, and confirm that causal ratings are sensitive to sample size, even when these ratings were given only once at the end of a whole block of trials rather than repeatedly while new trial information is presented. Although we found that the causal ratings were adjusted more rapidly toward the scale extremes, this effect was presumably more apparent than real, as it might have been mainly due to a greater number of comparison cases per block in the present procedure.

EXPERIMENT 3

So far we have seen that subjects progressively adjusted their causal ratings when subsequent covariation information confirmed their initial judgments. A different question is whether they would also adjust their judgments when subsequent information conflicts with their initial judgments, and how far this correction would go. This is an important question, because connectionist models have been heavily criticized on the ground that they adapt to temporal changes so rapidly and drastically that initial knowledge is fully overridden by the novel input, a feature that is known as catastrophic interference (Ratcliff, 1990).

This question was also addressed in impression-formation research (Anderson, 1981). It was found that information either at the beginning or at the end of a series of items may dominate impression judgments. When judgments were made at the end of a series, initial information tended to dominate (reflecting primacy). Conversely, when judgments were continuously given while receiving novel information, final information tended to have a greater impact (reflecting recency). However, in exploring causal judgments, Wasserman *et al.* (1996, Experiment 3) found no primacy or recency. From a normative perspective, this latter finding is exactly what should be expected if the same amount of novel covariation information contradicts previous causal information. People should revoke previous causal judgments if there is equally strong novel information indicating that they were wrong, leading to catastrophic interference without primacy or recency.

The issue of catastrophic interference is explored in the next experiment, where subjects were given a first block of four trials in which a given target always covaried with the outcome (to build up sufficient causal strength), and then received information in a second block of four trials where the same amount of 100% covariation was given (confirming initial judgments) or was lowered to 50% (partly disconfirming initial judgments) or to 0% (entirely disconfirming initial judgments). To test whether catastrophic learning takes place in causal judgments, the latter 0% condition was compared with a control condition in which subjects received 0% covariation all the time. Attribution ratings were measured after each trial from the last trial of the first block onwards, so that the disconfirmatory process could be monitored very closely. Hence, the design consisted of two within-subjects factors Covariation (0%, 50%, 100% or control) and Trial (4 to 8).

As before, our prediction is that the judgments will be gradually adjusted towards the novel degree of covariation, that is, a further increase of target ratings in the 100% condition, a mild decrease in the 50% condition, and a strong decrease in the 0% condition (as well as in the control condition). In addition, we predict that the disconfirming 0% information will have a catastrophic effect so that at the end of all trials, this condition will yield the same judgments as the control condition.

In contrast, all current statistical models predict, as noted above, that consistent information over trials (100% and control condition) will not lead to changes in the attribution ratings. For inconsistent information (50% and 0% conditions) these models anticipate either no or little changes. The probabilistic contrast model predicts a sudden downward shift in the 0% condition after the first

conflicting information without further changes, and a gradual decrease in the 50% condition. Likewise, the Bayesian model predicts a sudden downward shift without further changes in the 0% and 50% conditions. In contrast, the ANOVA model predicts a moderate decrease in the 50% condition, while η^2 is undetermined in the 0% condition.

Method

Subjects

Subjects were 101 male and female freshmen from the Dutch-speaking Vrije Universiteit Brussel, who participated for a partial requirement of an introductory psychology course.

Materials and Procedure

All instructions, materials, and questions were similar to those in the first experiment, with the following modifications. Subjects read eight stories without fillers (selected randomly from a pool of sixteen stories). Each story consisted of two Blocks of four trials. In the first Block, regardless of experimental condition, subjects received four trials in which the target actor or stimulus was always followed by the outcome. This allowed the target factor to acquire sufficient causal strength. In the second Block, four trials were given. These trials had a fixed covariation level. The 0% and 100% covariation conditions were similar to the first experiment, while the novel 50% covariation condition involved comparison cases which were followed by the outcome in half of the trials (i.e. midrange consensus or distinctiveness). The design is illustrated in Table 2. Trial order was fixed for all subjects, as shown in the table.

In addition, there was also a control condition in which the degree of covariation was kept to 0% from the first trial on. This condition involved seven comparison trials and one target trial (presented at the last trial of Block 1), which were always followed by the outcome. For all covariation conditions, ratings of the target factor were measured after each trial, from the fourth trial (or last trial of Block 1) onwards.

Results

The mean target ratings in function of covariation conditions and trial number are depicted in Figure 3. As expected, the ratings showed an increase over trials in the 100% condition, a very marginal

Table 2. Design of Experiment 3 illustrated for the event 'Els detested the singer Karel'

			Convariation (consensus)		
Trial no.	Trial type	Actor + stimulus	100% (Low Cs)	50% (Mid Cs)	0% (High Cs)
1-4	Target	*Els* + Karel	Detested	Detested	Detested
5	Comparison	Ilse + Karel	No	No	Detested
6	Comparison	Jana + Karel	No	Detested	Detested
7	Comparison	Olga + Karel	No	No	Detested
8	Comparison	Linda + Karel	No	Detested	Detested

Note: The target is in italics; Cs = consensus, Mid = Midrange; a similar manipulation was used for distinctiveness by varying the stimulus names instead of actor names (see also Table 1).

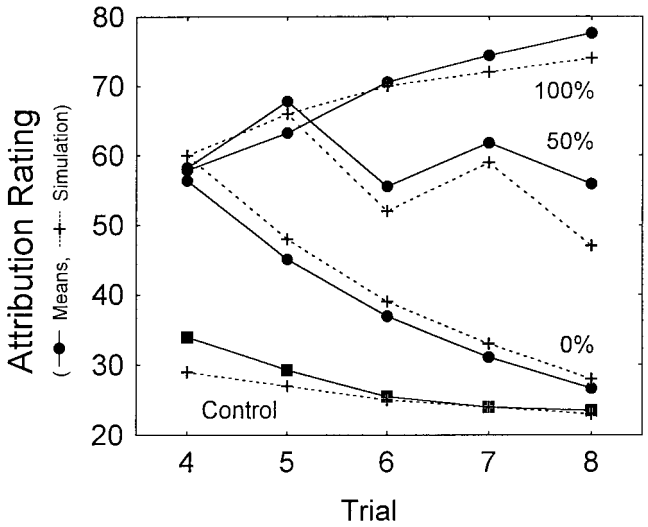

Figure 3. Experiment 3: attribution ratings and feedforward simulation in function of covariation level and number of trials (100% = low Cs/high Di; 50% = midrange Cs/Di; 0% = high Cs/low Di)

decrease in the 50% condition, and a substantial decrease in the disconfirming 0% condition which closely approached the control condition at the last trial. (The jagged pattern of the 50% condition is due to the fact that 50% covariation was reached on even trials only.)

A repeated-measures ANOVA with Covariation (100%, 50%, 0% and control) and Trial (fourth versus last) as within-subjects factors revealed a significant main effect of Covariation, $F(3, 98) = 94.26$, $p < 0.0001$, indicating again that subjects were very sensitive to the different degrees of covariation. The main effect of Trial was also significant, $F(1, 100) = 20.89$, $p < 0.0001$, which is due to the fact that the covariation levels were not evenly distributed in the first block (they were all 100% except for the control condition). Of most importance, the predicted interaction between Covariation and Trial was significant, $F(3, 98) = 55.12$, $p < 0.0001$.

To explore this interaction further and to examine whether adjustments were made after confirming trials as well after disconfirming trials, we conducted trend analyses in each covariation condition. In the confirming 100% condition, the predicted linear increase of the target ratings was significant, $F(1, 100) = 62.38$, $p < 0.0001$, while the quadratic trend was marginally significant, $F(1, 100) = 3.86$, $p < 0.06$. Like the previous experiments, this result is at variance with the statistical models.

In the disconfirming 50% condition, the expected linear decrease was significant, $F(1, 100) = 4.71$, $p < 0.05$, although the quadratic trend was much stronger $F(1, 100) = 11.51$, $p < 0.0001$, because of the jagged (non-randomized) pattern of ratings. The disconfirming 0% condition showed the predicted linear decrease, $F(1, 100) = 145.91$, $p < 0.0001$, just like the control condition, $F(1, 100) = 37.06$, $p < 0.0001$, while these two conditions also revealed a smaller quadratic trend, $Fs(1, 100) = 12.57$, $ps < 0.001$. More importantly, conforming to our catastrophic interference prediction, at the last trial, the disconfirming 0% condition did not differ reliably from the control condition, $F(1, 100) = 2.43$, $p = 0.122$.

Discussion

The present results extend the previous experiments where covariation information confirmed initial judgments, to instances where new information conflicts with earlier information. When subjects

learned that a target covaried with an outcome, they first developed positive causal estimates. But when later information disconfirmed these earlier covariation patterns, their estimates progressively decreased, the more trials were presented. When this novel information was completely opposite (as in the 0% condition), their ratings completely override the prior estimates, showing a typical catastrophic interference pattern, and no primacy or recency. These results replicate the findings of Wasserman *et al.* (1996, Experiment 3) and provide further support for our connectionist hypothesis.

MODEL SIMULATIONS

The present experiments revealed adjustment effects with an increasing number of observations that can be readily explained by connectionist principles. In order to evaluate how closely this formulation can predict our data, we computed simulations of a feedforward model, identical to that of Van Overwalle (1998), and correlated these to the observed attribution ratings.

As a way of comparison, we also ran simulations of two major statistical models: the probabilistic contrast (Cheng & Novick, 1990) and the ANOVA model (Försterling, 1992). However, it would be of little interest to reiterate the fact that these models fail to be sensitive to the number of observations. Therefore, we extended these existing models with additional parameters that take into account sample size, in order to evaluate whether these extensions would be sufficient to account for the observed size effect.

Specifically, we followed the suggestion by Cheng and Holyoak (1995, p. 273) that 'confidence in the assessment of a contrast is presumed to increase monotonically with the number of cases observed'. That is, we weighted the major theoretical variables of the models (conditional probabilities: Cheng & Novick, 1990; sum of squares: Försterling, 1989) in proportion to the number of observations available. We allowed two different confidence weights for frequencies that involved the presence and absence of the target cause (ω_t and ω_x respectively). This procedure is identical to the one used recently by Lober and Shanks (2000, p. 207), and parallels that of the connectionist models, where we also allowed two different learning rate parameters for target and comparison factors (ε_t and ε_x respectively). More technical details on the model specifications are given in the Appendix, Sections A–C.

Method

The models were run using exactly the same order of trials and blocks as in the experiments. The connectionist model was updated after each trial; and when trial order was randomized within each block (i.e. Experiments 1 and 2), we ran 100 simulations with a random trial order within each block, and averaged the results. The fit of the models was measured by taking the mean ratings for each combination of covariation level and block/trial, and then correlating that with the mean simulated values for those conditions. We sought the best-fitting parameters of each model by searching the maximum correlation given all admissible parameter values (Gluck & Bower, 1988a; Nosofsky, Kruschke, & McKinley, 1992). It should be noted that small changes of 0.05 in the obtained parameter values did not alter the results meaningfully.

In order to evaluate the performance of the models specifically with respect to the effect of sample size, we also computed the same correlation fit separately for each 0% and 100% covariation condition, using the same best-fitting model parameters obtained for the whole dataset.

Table 3. Fits of the models to the data

	Experiment			
Model	1	2	3	Mean *r*
Weighted ANOVA				
Overall	0.959	0.972	0.805	0.912
100%	0.940	0.000	0.748	0.563
0%	0.000	0.000	0.000	0.000
	$\omega_t = 0.22$	$\omega_t = 0.10$	$\omega_t = 0.20$	
	$\omega_x = 0.40$	$\omega_x = 0.00$	$\omega_x = 1.00$	
Weighted probabilistic				
Overall	0.988	0.984	0.951	0.974
100%	0.740	0.000	0.000	0.247
0%	0.948	0.988	0.993	0.976
	$\omega_t = 0.72$	$\omega_t \geq 0.00$	$\omega_t = 0.00$	
	$\omega_x = 0.32$	$\omega_x = 0.74$	$\omega_x = 1.30$	
Feedforward connectionist				
Overall	0.997	0.992	0.985	0.991
100%	0.985	0.946	0.993	0.975
0%	0.983	0.947	1.000	0.977
	$\varepsilon_t = 0.70$	$\varepsilon_t = 0.91$	$\varepsilon_t = 0.70$	
	$\varepsilon_x = 0.35$	$\varepsilon_x = 0.38$	$\varepsilon_x = 0.30$	

Note: Cell entries are correlations between mean observed ratings and mean simulated values (over all random runs); ω_t=best–fitting weight when target is present; ω_x=when target is absent; ε_t=best–fitting learning rate for target node; ε_x=for contextual node.

Results and Discussion

Table 3 displays the performance of all simulated models. The last column depicts the mean correlation across all three datasets. As can be seen in this column, all models reach substantial overall fit correlations, indicating that they accurately simulated the actual covariation to which our subjects were also very sensitive. However, the fit of the weighted ANOVA model was lowest, followed by the fit of the weighted probabilistic and connectionist models which was about equally high. Difference tests revealed that the correlations of the weighted ANOVA model were significantly lower than those of the connectionist model for all three experiments ($p < 0.05$, one-sided), whereas the correlations of the weighted probabilistic model were marginally lower than the connectionist model for the first two experiments ($ps < 0.08$-0.11, one-sided) and reached significance for the last experiment ($p < 0.05$, one-sided).

Of most importance are the fit correlations within the 0% and 100% conditions because they reflect to what extent the models are able to replicate the sample size effect. As expected, the fit was generally poor for the two statistical models, as some datasets revealed zero correlations. In contrast, consistent with predictions, the connectionist model showed in all 0% and 100% conditions very high positive correlations above 0.90 in all datasets.[3]

We predicted that Busemeyer's (1991) averaging strategy and Hogarth and Einhorn's (1992) belief-adjustment model would give similar predictions as the more general delta algorithm (although we used different parameters, see the Appendix, Section D). In addition, Read and Montoya (1999)

[3]Statistics of differential fit make little sense here because of the reduced number of observations (no higher than 10) as we computed correlations between mean observed ratings and one simulation run (for statistical models because they are independent of order) or mean simulation values (for the connectionist model).

argued that a recurrent network might be a better model to simulate causal judgments because it has a more extensive architecture than the feedforward network (see the Appendix, Section E). Therefore, we also ran simulations with these alternative models and found that the results were very similar to those of the feedforward network. Averaged across all conditions (i.e. Overall, 100% and 0% conditions) and all datasets, the mean correlation of the belief-adjustment model was 0.971 and that of the recurrent network 0.974, which is very close to the 0.981 mean correlation of the feedforward network. This suggests that the error-correcting delta algorithm common to all these models, rather than anything particular about the feedforward architecture was essential in reproducing the present data.

To give a better appreciation of the high fit of the feedforward model, its simulation results are graphically depicted in Figures 1 to 3. Note that to visually match the simulation results with the observed data, the simulated values were regressed on the observed ratings.

GENERAL DISCUSSION

The present experiments accomplished two things. First, they demonstrated unambiguously the existence of a sample size effect in causal attribution, and second, they suggested that adaptive models using the error-correcting delta algorithm can account for this effect, unlike earlier statistical models (e.g. Cheng & Novick, 1990; Försterling, 1989). We focus on each point in turn.

Sample Size Effect

Our findings confirm that people are quite sensitive to sample size while making causal attributions. They begin with a relatively moderate causal estimate and, as more observations are made, adjust their estimate in the direction of the true covariation between cause and effect. These sample size adjustments were made when the information confirmed or disconfirmed initial judgments, and when judgments were made repeatedly during acquisition of new information or after all information had been acquired.

These data suggest that people respect not only the statistical norm of covariation but also the statistical law of large numbers, that is, they are more accurate when judgments are based on a larger number of instances. This is the first study on social attributions that demonstrates people's adherence to this principle of large numbers independently of other attributional principles. A number of similar studies conducted in the past failed to unconfound the effect of sample size from covariation (Försterling, 1992) or contained material that was less rich in social content (Baker *et al.*, 1989; Shanks, 1985, 1987; Lopez and Shanks: in Shanks, 1995).

However, there is a potential limitation in our research in that we did not measure subjects' confidence in their attribution ratings. Perhaps subjects would not have changed their causal ratings if given the opportunity to indicate their degree of confidence in them. If true, this would imply for Experiment 1, for instance, that after a single first piece of evidence, subjects believed that the novel cause completely influenced the outcome, but at the same time expressed uncertainty about this belief. Or, for Experiment 3, it would imply that after multiple confirmatory pieces of evidence and a single disconfirmatory piece, they believed that the cause was completely irrelevant, but again expressed uncertainty about this. Such reasoning seems very implausible and anomalous. Although the present

data do not directly speak to this issue, research in social attribution by Försterling (1992) did not reveal any changes in confidence ratings after manipulating sample size and covariation.[4]

Theoretical Implications

Our findings may have important implications for current theories in social psychology on how people make causal attributions. Contrary to statistical models (Cheng & Novick, 1990; Försterling, 1992), our data and simulations showed that the delta algorithm used in adaptive connectionist models (McClelland & Rumelhart, 1988) can readily describe how people incrementally adjust their estimates when more cases are observed. This algorithm is based on the assumption that people attempt to reduce the error between their mental representation of the environment and what actually happens (Van Overwalle, 1998; Read & Montoya, 1999; Smith & DeCoster, 1998).

Given this capacity, perhaps adaptive connectionist networks can simulate other judgmental biases and sample size effects such as illusory correlation, group homogeneity, group polarization (Fiedler, 1996; Ebbesen & Bowers, 1974), and the impact of increasing information on impression formation (Anderson, 1981), attitude change (Eagly & Chaiken, 1993) and hypothesis testing (Fiedler *et al.*, 1999). This would place social process models in line with current conceptions of the connectionist workings of the brain (McClelland & Rumelhart, 1988; Smith, 1996).

Of course, an interesting question that arises is whether it is possible to extend mainstream statistical models with a mechanism that makes them sensitive to sample size. The simulations demonstrated that simply adding confidence weights in proportion to the number of observations as suggested by Cheng and Holyoak (1995) does not work. Perhaps another possibility is to assume that positive (i.e. 100%) information increases confidence over trials whereas negative (i.e. 0%) information decreases confidence. However, this reasoning tends to be circular, because it confounds confidence with covariation assessment (i.e. it requires sorting out positive from negative information independently of the model's theoretical variables that define what information is positive or negative). Moreover, it shift the burden of proof from attribution ratings to confidence ratings, about which the statistical theories have little to say.

Another possibility is to incorporate a sort of error noise term in the statistical formulations of the probabilistic and ANOVA models and to make the assumption that increasing sample size will reduce the noise component (for a similar approach, see Fiedler, 1996). However, this extension seems implausible because the variance between subjects was typically higher at the last block than at the beginning block, contrary to what one would expect if noise were reduced. Taken together, we see no way to solve the sample size limitation within the boundaries of earlier statistical theories.

Perhaps a more interesting question is whether recent statistical models that incorporate an updating mechanism are more sensitive to sample size. The belief-adjustment mechanism proposed by Hogarth and Einhorn (1992) as well as the serial averaging strategy suggested by Busemeyer (1991) are able to accomplish that, as was documented by the simulations. However, as noted earlier, these models are limited to a single cause at the time and are therefore less powerful than the more general delta algorithm.

In sum, many current attribution models have difficulties in integrating and formalizing the concept of sample size, and recent proposals that have made advancements on this issue are in fact identical to a restricted form of the delta algorithm. Whatever direction future developments of current statistical models will take, our data and simulations surely highlight the need of an error-correcting learning component.

[4]This failure to find any effects of certainty measures, as well as similar failures in the experimental literature (Shanks, 1985, 1987) made us decide not to include certainly measures in the present studies.

Conclusion

We started with the question of what type of process people are using when making and adjusting causal judgments. Earlier attribution theorists often warned that their statistical equations did not address directly the cognitive operations involved in making causal attributions, but that people probably performed a more simple, analogous reasoning (Cheng, 1997; Försterling, 1992). Although it might be argued that statistical models should be considered at a computational level (What is computed?) rather than at an algorithmic level that keeps track of the development of causal judgment (How is it computed?), the fact that the connectionist approach accommodates both levels makes that argument questionable.

We believe that connectionist models may complement or even replace earlier statistical models by providing a low-level description of the attribution process. This is consistent with the observation that causal attribution in the hustle of daily life is often effortless, preconscious and spontaneous. The statistical models may, of course, describe causal judgments at a more explicit symbolic level, for instance, when subjects are processing verbal summary information (Baker *et al.*, 1996; Lober & Shanks, 2000). However, recent evidence suggests that even within a verbal format, covariation information can automatically facilitate or inhibit spontaneous trait inferences (Van Overwalle, Drenth, & Marsman, 1999). Together with the present data, this seems to indicate at the very least that causal induction often proceeds without any symbolic rule.

ACKNOWLEDGEMENTS

This research was in part supported by the Belgian National Foundation of Scientific Research (NFWO) under grant 8.0192.95. We are grateful to B. Timmermans for his comments on an earlier version of the manuscript and to I. Ponjaert for allowing us to use her subject pool.

REFERENCES

Allan LG. 1993. Human contingency Judgments: Rule based or associative? *Psychological Bulletin* **114**: 435–448.

Anderson JR, Sheu C-F. 1995. Causal inferences as perceptual judgments. *Memory and Cognition* **23**: 510–524.

Anderson NH. 1967. Averaging model analysis of set-size effects in impression formation. *Journal of Experimental Psychology* **75**: 158–165.

Anderson NH. 1981. *Foundations of Information Integration Theory.* Academic Press: New York.

Baker AG, Berbier MW, Vallée-Tourangeau F. 1989. Judgments of a 2×2 contingency table: sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology* **41B**: 65–97.

Baker AG, Murphy RA, Vallée-Tourangeau F. 1996. Associative and normative models of causal induction: Reacting to versus understanding cause. *The Psychology of Learning and Motivation* **34**: 1–46.

Busemeyer JR. 1991. Intuitive statistical estimation. In *Contributions to Information Integration Theory, Vol. 1: Cognition*, Anderson N (ed.). Erlbaum: Hillsdale, NJ.

Chapman GB, Robbins SJ. 1990. Cue interaction in human contingency judgment. *Memory and Cognition* **18**: 537–545.

Cheng PW. 1997. From covariation to causation: A causal power theory. *Psychological Review* **104**: 367–405.

Cheng PW, Holyoak KJ. 1995. Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In *Comparative Approaches to Cognitive Science*, Roiblatt HL, Meyer J.-A. (eds). MIT Press: Cambridge, MA; 271–302.

Cheng PW, Novick LR. 1990. A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology* **58**: 545–567.
Eagly AH, Chaiken S. 1993. *The Psychology of Attitudes*. Harcourt Brace: San Diego, CA.
Ebbesen EB, Bowers RJ. 1974. Proportion of risky to conservative arguments in a group discussion and choice shifts. *Journal of Personality and Social Psychology* **29**: 316–327.
Fales E, Wasserman EA. 1992. Causal knowledge: What can psychology teach philosophers? *Journal of Mind and Behavior* **13**: 1–28.
Fiedler K. 1996. Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environment. *Journal of Personality and Social Psychology* **103**: 193–214.
Fiedler K, Walther E, Nickel S. 1999. The auto-verification of social hypotheses: Stereotyping and the power of sample size. *Journal of Personality and Social Psychology* **77**: 5–18.
Försterling F. 1989. Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality and Social Psychology* **57**: 615–625.
Försterling F. 1992. The Kelley model as an analysis of variance analogy: How far can it be taken? *Journal of Experimental Social Psychology* **28**: 475–490.
Gluck MA, Bower GH. 1988a. From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General* **117**: 227–247.
Gluck MA, Bower GH. 1988b. Evaluating an adaptive network model of human learning. *Journal of Memory and Language* **27**: 166–195.
Hewstone M, Jaspars J. 1987. Covariation and causal attribution: A logical model of the intuitive analysis of variance. *Journal of Personality and Social Psychology* **53**: 663–673.
Hilton DJ, Slugoski BR. 1986. Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review* **93**: 75–88.
Hogarth RM, Einhorn HJ. 1992. Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology* **24**: 1–55.
Johnson-Laird PN. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press: Cambridge.
Kelley HH. 1967. Attribution in social psychology. *Nebraska Symposium on Motivation* **15**: 192–238.
Lober K, Shanks DR. 2000. Causal induction based on causal power? Critique of Cheng 1997. *Psychological Review* **107**: 195–212.
Manis M, Dovalina I, Avis NE, Cardoze S. 1980. Base rates can affect individual predictions. *Journal of Personality and Social Psychology* **38**: 231–248.
Markman B. 1989. LMS rules and the inverse base-rate effect: Comment on Gluck and Bower (1988). *Journal of Experimental Psychology: General* **118**: 417–421.
McClelland JM, Rumelhart DE. 1988. *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs and Exercises*. Bradford: Cambridge, MA.
Nisbett RE, Krantz DH, Jepson C, Kunda Z. 1983. The use of statistical heuristics in everyday inductive reasoning. *Psychological Review* **90**: 339–363.
Nosofsky RM, Kruschke JK, McKinley SC. 1992. Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory and Cognition* **18**: 211–233.
Orvis BR, Cunningham JD, Kelley HH. 1975. A closer examination of causal inference: The role of consensus, distinction and consistency information. *Journal of Personality and Social Psychology* **32**: 605–616.
Ratcliff R. 1990. Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review* **97**: 285–308.
Read SJ, Marcus-Newhall A. 1993. Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology* **65**: 429–447.
Read SJ, Montoya JA. 1999. An autoassociative model of causal reasoning and causal learning: Reply to Van Overwalle's critique of Read and Marcus-Newhall (1993). *Journal of Personality and Social Psychology* **76**: 728–742.
Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, Black AW, Prokasy WF (eds). Appleton-Century-Crofts: New York; 64–98.
Rudolph U, Försterling F. 1997. The psychological causality implicit in verbs: A review. *Psychological Bulletin* **121**: 192–218.

Shanks DR. 1985. Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology* **37**b: 1–21.
Shanks DR. 1987. Acquisition functions in contingency judgment. *Learning and Motivation* **18**: 147–166.
Shanks DR. 1991. Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory and Cognition* **17**: 433–443.
Shanks DR. 1995. Is human learning rational? *Quarterly Journal of Experimental Psychology* **48**a: 257–279.
Shanks DR, Lopez FJ, Darby RJ, Dickinson A. 1996. Distinguishing associative and probabilistic contrast theories of human contingency judgment. In *The Psychology of Learning and Motivation* (Vol. 34), Shanks DR, Holyoak KJ, Medin DL (eds). Academic Press: New York; 265–311.
Smith ER. 1996. What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology* **70**: 893–912.
Smith ER, DeCoster J. 1998. Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology* **74**: 21–35.
Tassoni CJ. 1995. The least mean square network with information coding: A model of cue learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* **21**: 193–204.
Van Hamme L, Wasserman EA. 1994. Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation* **25**: 127–151.
Van Overwalle F. 1996. The relationship between the Rescorla–Wagner associative model and the probabilistic joint model of causality. *Psychologica Belgica* **36**: 171–192.
Van Overwalle F. 1997. A test of the Joint model of causal attribution. *European Journal of Social Psychology* **27**: 221–236.
Van Overwalle F. 1998. Causal explanation as constraint satisfaction: a critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology* **74**: 312–328.
Van Overwalle F, Van Rooy D. 1998. A connectionist approach to causal attribution. In *Connectionist Models of Social Reasoning and Social Behavior*, Read SJ, Miller LC (eds). Erlbaum: New York.
Van Overwalle F, Drenth T, Marsman G. 1999. Spontaneous trait inferences: Are they linked to the actor or to the action? *Personality and Social Psychology Bulletin* **25**: 450–462.
Wasserman EA, Kao S-F, Van Hamme L, Katagiri M, Young ME. 1996. Causation and association. *The Psychology of Learning and Motivation* **34**: 207–264.
Weiner B. 1985. 'Spontaneous' causal thinking. *Psychological Bulletin* **97**: 74–84.

APPENDIX

This appendix discusses how the models were specified for the simulations, and how the statistical models were extended to account for sample size.

A. Weighted Probabilistic Model

According to the probabilistic contrast formulation of causality (Allan, 1993; Cheng & Novick, 1990), causal strength is defined as:

$$\Delta P = P(O \mid T) - P(O \mid \tilde{T}) \tag{A1}$$

where P is a conditional probability, O is the outcome and T is the target cause. To make the probabilistic model sensitive to the number of observations, we extended this ΔP formulation by weighting each of the conditional probabilities P with a freely estimated proportion (between 0 and 1) of the frequencies involved. That is, if ω_t denotes a proportion of the frequencies when cause T is present and if ω_x denotes a proportion of the frequencies when cause T is absent, then the strength of a target cause may be formalized as:

$$\Delta P = \omega_t P(O \mid T) - \omega_x P(O \mid \tilde{T}) \tag{A2}$$

B. Weighted ANOVA Model

The ANOVA formulation defines causal strength as an analog to the effect size η^2 of a standard ANOVA (Försterling, 1992), which is given by:

$$\eta^2 = SS_{\text{between}}/SS_{\text{total}} \tag{A3}$$

where the presence of the outcome is indicated by 1 and its absence by 0 in the standard sum of squares (*SS*) formula. To make Försterling's (1992) model sensitive to the number of observations, we weighted η^2 with a freely estimated proportion of the frequencies involved, that is, with the same ω_t and ω_x as defined above. Thus, this becomes:

$$\eta^2 = [\omega_t SS_{\text{between}}]/[(\omega_t + \omega_x)SS_{\text{total}}] \tag{A4}$$

C. Feedforward Model

The feedforward network involves two layers of nodes (McClelland & Rumelhart, 1988). The first layer consists of two input nodes, a target node *T* that represents the target and a contextual node *X* that represents all comparison factors. The second layer consists of an output node representing the outcome. The input nodes are connected to the output nodes via weighted, unidirectional links. When a cause is present at a trial, its input node is activated to the default level 1; when a cause is missing at a trial, its input node is activated to a negative value α_n, which lies between 0 and -1 (Van Hamme & Wasserman, 1994). However, to keep the number of free parameters in all models equal, this parameter was arbitrary set at an intermediate value $\alpha_n = -0.50$. The activation of all input nodes is spread automatically to the output node in proportion to the weights of the links, and linearly summarized to represent the output activation.[5]

After each trial, the weights of the links are incrementally adjusted by reducing the error between the output activation (the outcome predicted by the network) and a teaching activation representing the actual outcome, which is 1 when the outcome is present and -1 when absent. This adjustment is mathematically expressed by the delta algorithm (McClelland & Rumelhart, 1988, p. 87):

$$\Delta_W = \varepsilon(a_{\text{t}} - a_{\text{o}})a_{\text{i}} \tag{A5}$$

where ε is the learning rate (typically between 0 and 1), and where a_{t}, a_{o}, a_{i} denote respectively the teaching, output and input activations. Before running the simulations, the weights were set at zero starting values. We assumed that there were separate learning rates for the target node and the contextual node, denoted respectively by ε_t and ε_x. In addition, we also assumed that the contextual node was always activated during learning (see Van Overwalle, 1996, 1997, 1998; Van Overwalle & Van Rooy, 1998).

[5]The negative activation value for absent but relevant factors proposed by Van Hamme and Wasserman (1994) solves a number of problems of the original Rescorla–Wagner (1972) model where absent factors always received zero activation. It can be demonstrated following the logic of Chapman and Robbins (1990) and Van Overwalle (1996) that this revision converges to the statistical norm, as was the case for the original Rescorla–Wagner coding. Note that Van Hamme and Wasserman's (1994) formulation implies that negative activation plays a role only on adjustments, not in the computation of the outcome activation. Other revisions, which assumed that the negative activation also spreads to determine the output activation (Markman, 1989; Tassoni, 1995), do not converge to the statistical norm.

The strength of a causal explanation is reflected in the weight of the connection between the input node and the output node, which can also be measured by activation the input node and reading off the resulting output activation.

D. Belief-Adjustment Model

We applied Equations 7a and 7b of Einhorn and Hogarth (1992) for negative and positive evidence respectively, with learning rates as formalized in Equations 6a and 6b and hypothesis $R = 0$ for causal explanations. Given a subjective evaluation of 1 for positive evidence on a trial (contingency table frequencies a or $d > 0$) and an evaluation of -1 for negative evidence (frequencies b or $c > 0$), the belief-adjustment model of Einhorn and Hogarth (1992) can be restated as:

$$S_k = S_{k-1} + w_k(1 - S_{k-1}), \text{for positive evidence} \tag{A6}$$

$$S_k = S_{k-1} + w_k(0 - S_{k-1}), \text{for negative evidence,} \tag{A7}$$

with as free parameters the initial subjective estimate, S_o, and weight adjustment rate, w_k. The serial averaging strategy of Busemeyer (1991) reduces to exactly the same set of equations. It can be easily verified that equations (A6) and (A7) are a simplified version of the delta algorithm (equation (A5)) with only a single causal estimate S_{k-1} in the right-most part of the equation, whereas the delta algorithm incorporates all previous causal estimates in the output activation a_o.

E. Recurrent Model

For a detailed description of the recurrent model, we refer to Read and Montoya (1999). The model simulations were run using the same specifications as the feedforward model, with the following additional recurrent parameters: $\text{istr} = \text{estr} = \text{decay} = 1$, using the linear activation rule with one internal processing cycle (McClelland & Rumelhart, 1988).

