Do Markov Violations and Failures of Explaining Away Persist with Experience?

Benjamin Margolin Rottman (rottman@pitt.edu)
Learning Research and Development Center, University of Pittsburgh 3939 O'Hara St
Pittsburgh, PA 15260 USA

Reid Hastie (reid.hastie@chicagobooth.edu)
The University of Chicago Booth School of Business, 5807 South Woodlawn Ave
Chicago, IL 60637 USA

Abstract

Making judgments by relying on beliefs about causal relations is a fundamental aspect of everyday cognition. Recent research has identified two ways that human reasoning seems to diverge from optimal standards; people appear to violate the Markov Assumption, and do not to "explain away" adequately. However, these habits have rarely been tested in the situation that presumably would promote accurate reasoning – after experiencing the multivariate distribution of the variables through trial-by-trial learning, even though this is a standard paradigm. Two studies test whether these habits persist 1) despite adequate learning experience, 2) despite incentives, and 3) whether they also extend to situations with continuous variables.

Keywords: Causal Reasoning, Markov Assumption, Explaining Away

Introduction

In the last decade there has been a surge of interest in causal reasoning, particularly whether people reason in line with Causal Bayesian Networks. One question is how people learn causal structures (Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). Another questions is, once one has knowledge of a causal network, how well can he or she make inferences. For example, What are my chances of developing Sickle Cell disease given that my mother and grandmother have it; the causal structure is [*Grandmother* → *Mother* → *Daughter*]? Or what are my chances of getting an A on an exam if I study for 2 more hours but get 2 fewer hour of sleep; [*Study Time* → *Exam Grade* ← *Sleep*]?

Though people are often accurate at making probabilistic causal inferences, two habits deviate from rationality (Rottman & Hastie, 2014). First, people often use cues that are statistically irrelevant for a given inference, a violation of the "Markov Condition." Second, when there are two causes of one effect, people often have difficulty correctly inferring the probability of one cause given knowledge of the other two variables, known as "explaining away."

However, these habits have rarely been tested under circumstances in which participants receive all the information required to perform optimally. Making a quantitative probabilistic judgment requires having knowledge of the statistical covariation between the variables. For example, without knowledge of the precise statistical relations, a reasoner could predict that exam grade will increase with study time and decrease with less sleep, but would not be able to predict how exam grade would change with 2 hours more study time and 2 hours less sleep.

Giving participants knowledge of the covariation between variables through trial-by-trial learning is a common paradigm for studying causal reasoning and is known to improve judgment compared to verbal descriptions (Christensen-Szalanski & Beach, 1982). But surprisingly, these two reasoning habits have rarely been tested in situations when participants have knowledge of this covariation. Thus, the goal of the current studies is to test whether these habits persist despite experience. This research also provides the first systematic test of causal reasoning with Gaussian as opposed to binary variables.

The Markov Condition

Consider estimating the probability that a daughter will have sickle cell disease given that her mother has it. Because sickle cell is a simple autosomal recessive disease, whether the grandmother has sickle cell disease has no influence on whether the daughter has it above and beyond that her mother has it, which can be summarized with the following causal structure [*Grandmother* → *Mother* → *Daughter*]. The mother is a perfect mediator.

More generally, for chain [$X \rightarrow Y \rightarrow Z$] and common cause structures [$X \leftarrow Y \rightarrow Z$], the Markov condition asserts that when inferring the state of X, the state of Z is irrelevant if the state of Y is known. This implies, for example, that $P(z=1|y=1,x=1)$ equals $P(z=1|y=1,x=0)$. However, people tend to violate this assumption. In situations in which all variables are presented as binary (1 or 0), people tend to infer that $P(x=1|y=1,z=1) > P(x=1|y=1,z=0)$. The Markov condition can also be applied with continuous rather than binary variables (e.g., daughter's height is independent of grandmother's height given knowledge of mother's height).

A number of theorists have proposed different explanations for this effect (Park & Sloman, 2013; Rehder, 2014a), and we believe it is likely that multiple explanations apply, likely in unison. However, only one study has tested violations of the Markov assumption after participants experienced the covariation between all three variables (Park & Sloman, 2013 Experiment 3). All the other studies told participants about the causal relations verbally.

However, a problem with verbal descriptions is that it is likely not clear to participants whether a description like "X causes Y, which causes Z" implies that X has no indirect link

on *Z*. Giving participants trial-by-trial experience would allow them, to examine whether there is any effect of *X* on *Z* above and beyond *Y*. The current study serves to replicate this single experiment testing whether people still violate the Markov assumption after obtaining multivariate experience, and extend the results from binary variables to learning about Gaussian-distributed variables.

Explaining Away on Common Effect Structures

The second non-rational inference habit is a violation of "explaining away". Explaining away has long been viewed as an underlying principle in social attribution (Kelley, 1972), legal exoneration, and medical diagnosis.[1] For example, consider two diseases, flu and asthma, both of which can cause a cough [*flu*→ *cough* ← *asthma*]. A patient presents with a cough and has a history of asthma; because of the history of asthma, it is unnecessary to infer that the patient also has the flu to explain the cough, so $P(flu=1|cough=1,asthma=1)$ should be fairly low. However, if it is unknown whether the patient has asthma $P(flu=1|cough=1)$, then flu becomes more likely, and if it is known that the patient does not have asthma $P(flu=1|cough=1,asthma=0)$, flu becomes even more likely. More generally, in common effect structures [$X \rightarrow Y \leftarrow Z$], the normative pattern of reasoning is $P(X=1|Y=1,Z=1) < P(X=1|Y=1) < P(X=1|Y=1,Z=0)$.

A similar pattern of reasoning should occur with continuous variables. Consider if *X* represents the IQ of a father, *Z* represents the IQ of a mother, and *Y* represents the IQ of their child. Further, suppose that *Y* is the average of *X* and *Z*. (In reality *Y* is not a perfect average because there is noise involved, but even with noise the same basic pattern demonstrated here occurs.) Knowing the IQ of the mother ($Z=120$) does not help us infer the IQ of the father (*X*). However, suppose that it is known that $Z=120$ and $Y=100$. We can easily calculate that $X=80$. That is, once the value of *Y* is known, *X* and *Z* become negatively dependent; the higher *Z* is holding *Y* constant, the lower *X* must be.

The most thorough exploration of these judgments had participants make forced choice decisions of which would be higher, $P(X=1|Y=1,Z=0)$ vs. $P(X=1|Y=1)$ and $P(X=1|Y=1)$ vs. $P(X=1|Y=1,Z=1)$ (Rehder, 2014a). These studies revealed a tendency in the opposite direction of explaining away, $P(X=1|Y=1,Z=0) < P(X=1|Y=1) < P(X=1|Y=1,Z=1)$, or ambivalent; there was not a robust explaining away pattern. This study implemented a number of novel counterbalancing and control features, and presents the strongest evidence to date on explaining away.

However, like the few others before it, this study did not present participants with experience by which they could actually learn the correlations between the variables; the reasoning was solely based upon a verbal description of the causal structure. The problem with just a verbal description of the structure ("*X* and *Z* both cause *Y*") is that it does not convey the strength of the causal relations. If in fact *X* and *Z* are both weak causes of *Y*, then the normative amount of explaining away is quite small. Additionally, explaining away normatively varies by exactly how the two causes combine to produce the effect. Explaining away should occur if either cause is sufficient to produce the effect (e.g., flu and asthma for cough), but not when both causes are necessary (e.g., spark and oxygen for fire) (Rehder, 2014b), which may not be fully conveyed verbally.

Thus, in the current studies we gave participants experience that instantiates the multivariate distribution among the three variables, from which normatively correct inferences can be calculated. This also allows us to test not just the qualitative predictions but also whether the judgments are quantitatively on target. We also test the explaining away habit with both binary and Gaussian variables. One other study examined explaining away with continuous variables (Sussman, Abigail & Oppenheimer, 2011), and found insufficient explaining away; however, in that task participants again did not have experience or knowledge quantifying the strength of the causal relations.

Study 1: Binary Variables

In Study 1 participants learned about three causal structures, with binary variables. Afterwards they made a judgments predicting the state of one variable given knowledge of one or two of the other variables to test whether their judgments violated or upheld the Markov condition and whether they demonstrated explaining away appropriately.

Methods

Participants Fifty-one undergraduates at the University of Chicago were paid $12 per hour to participate; the study lasted 17 minutes on average. For further motivation they were also paid 8 cents for each correct inference.

Stimuli and Design Participants reasoned about three scenarios involving a chain [$X \rightarrow Y \rightarrow Z$], a common cause [$X \leftarrow Y \rightarrow Z$], and a common effect [$X \rightarrow Y \leftarrow Z$] structure. The variables (*X*, *Y*, and *Z*) were framed as physiological variables in the human body that could be either high (represented as + or 1) or low (represented as – or 0). There were three cover stories, one about neurotransmitters (amounts of Serotonin, Epinephrine, and Dopamine), another about how the digestive tract absorbs chemicals from food (amounts of Water, Protein, and Fructose Absorption), the last one about components of blood (Red Blood Cell, White Blood Cell, and Platelet Concentration). These variables were chosen so that they could plausibly be causally related to one another probabilistically in any

[1] We use the term "explaining away" because an alternate label, "discounting", has many informal meanings and has been used in psychology to refer to other phenomena. In particular, "discounting" has often been used to refer to lowering one's estimate of the strength of one cause when one learns of a second cause that is strong, which is related to both rational and irrational forms of "cue competition," "blocking," and "conditioning." The judgments assessed here are probability estimates, not causal strength judgments.

possible combination, but participants would be unlikely to have prior beliefs about how they were causally related. The order of the three causal structures, the cover-stories, the assignments of the three labels to variables (*X*, *Y* and *Z*), the position of the three variables on the computer screen, and the order of the learning trials were all randomized.

The sets of learning trials (Table 1) were generated in the following way: For the chain and common cause structure, the chosen parameters produced identical sets of learning trials. When a cause was present it produced its effect with a probability of .75. When a cause was absent its effect still occurred with a probability of .25. The base rates of exogenous causes were .50. For the common effect structure, the base rates of the two causes, $P(x=1)$ and $P(z=1)$ were also .50. The two causes combined through a Noisy-OR gate with strengths of .50, and thus $P(y=1|x=0,z=0) = 0$, $P(y=1|x=1,z=0) = P(y=1|x=0,z=1) = .50$, and $P(y=1|x=1, z=1)=.75$.

The normative point estimate inferences can be computed directly from Table 1. For example, $P(X=1|Y=1,Z=1)$ can be computed by dividing the sum of all the rows in which $X=1$, $Y=1$, and $Z=1$ by the sum in all the rows in which $Y=1$ and $Z=1$ (e.g., [6]/[6+4] = .60 for the common effect). Alternatively, if people learn the parameters of the causal model from experience the normative inferences can be derived from the parameters (Rottman & Hastie, 2014).

Table 1: Learning Trials in Experiment 1.

X	Y	Z	Chain and Common Cause	Common Effect
1	1	1	9	6
1	1	0	3	4
1	0	1	1	2
1	0	0	3	4
0	1	1	3	4
0	1	0	1	0
0	0	1	3	4
0	0	0	9	8

Procedures The general procedure followed a standard trial-by-trial, case-by-case causal learning paradigm in which participants were first told a causal cover-story, then learned the probabilistic relations between the variables from experience, and finally made a series of inferences. Participants were asked to pretend that they were physiologists studying biological processes in the human body. They were told that they would perform studies in which they would bring healthy people into a laboratory and would measure three physiological variables. They were told how to interpret pictures like the ones in Figure 1, where arrows represented causal relations between the three physiological variables and pointed from causes to effects. A "+" sign signified a high amount of the variable and "–" a low amount of the variable. Participants were also told not to use any prior knowledge about physiology and to assume that these three variables are the only ones that mattered within this biological system.

Next, participants completed a learning phase for each of the three causal scenarios in a randomized order, involving a chain, common cause, and common effect. Participants were shown a graphical representation of the causal relationships and they observed whether each of the variables was high or low in a sample of 32 cases ("healthy people"). The cases were presented in a sequential trial-by-trial format (Figure 1a) in a randomized order and the positions of the three variables, X, *Y,* and *Z* on the screen were randomized.

After the learning phase participants made a series of inferences; the order of the questions was randomized. Participants made inferences about each variable, given that the states of the other two variables were high, low, or unknown. Figure 1b shows one of the inferences involved in explaining away. The questions were presented to participants using both a visual diagram and corresponding text. When the state of a variable was unknown it was denoted visually with an X and participants were told that the machine used to test for that variable was broken.

Underneath the variable to be inferred was a gray box that participants used to input their estimates. Following the practice of Waldmann and Hagmayer (2005), we used a frequency format (number of people out of 20) for the question rather than a probability format.

At the end of the study, participants were paid for their time and a bonus for the number of questions that they answered correctly; an answer on the 21-point scale was considered correct if it was the closest response to the normative calculation. Participants were not given feedback during the experiment.

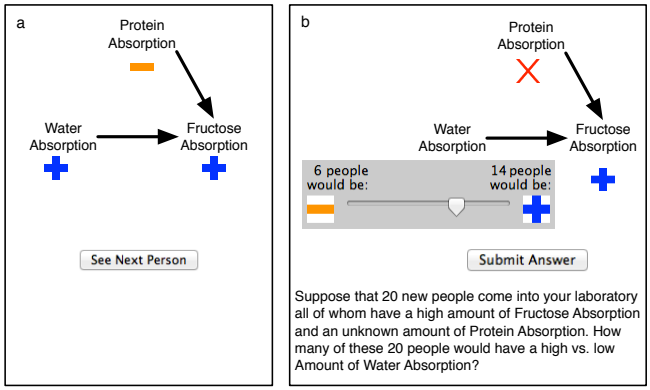

Figure 1: Example Stimuli in Experiment 1 for the Common Effect Structure. Note. Panel A shows an example of one trial in the learning phase. Panel B shows an example of the judgment $P(\text{water}=1 \mid \text{fructose}=1)$.

Analyses

All responses were converted to a probability scale of 0-1. Because a common cause $[X \leftarrow Y \rightarrow Z]$ is symmetric, inferences like $P(z=1|y=1,x=1)$ and $P(x=1|y=1,z=1)$ are essentially duplicates or repeated measures. For the chain $[X \rightarrow Y \rightarrow Z]$, we looked and did not find systematic

differences for the inferences going down versus up the chain, thus we treat inferences like $P(z=1|y=1,x=1)$ and $P(x=1|y=1,z=1)$ as repeated measures. Additionally, on the chain and common cause the normative value for $P(z=1|y=1,x=1)=.75$ and the normative value for $P(z=1|y=0,x=0)=.25$. These inferences showed the same patterns regardless of whether $Y=1$ or $Y=0$, so we converted inferences in the bottom half of the scale to the top half so they could be analyzed together.

Results and Discussion

The Markov Assumption The Markov Assumption implies that pairs of inferences such as $P(x=1|y=1,z=1)$ and $P(x=1|y=1,z=0)$ on the chain and common cause should be equivalent. To test whether the average judgments were reliably different for the $P(x=1|y=1, z=0)$ vs. $P(x=1|y=1, z=1)$ judgments, we used mixed linear regressions. When appropriate, we used a negative square root transformation on the dependent variable to transform the data to rough normality. All confidence intervals reported were back-transformed so that they can be interpreted on the probability scale. [2] For one participant in Experiment 1a the participant's responses were very similar within a scenario, likely reflecting disengagement from the task. Thus, we threw out those observations.

Table 2. Markov Assumption Results in Study 1.

Inference	Norm.	$X\rightarrow Y\rightarrow Z$	$X\leftarrow Y\rightarrow Z$
$P(x=1\|y=1,z=1)$	.75	.78	.77
$P(x=1\|y=1,z=0)$	.75	.59	.65
95% CI of Difference	0	[.13, .22]	[.07, .16]

We ran three mixed effects regressions for the three structures, to test whether the inferences were higher when the screened-off variable (z in Table 2) was 1 instead of 0. By-subject random effects were included for the intercept and for the slope (the difference between the two inferences). See Table 2 for 95% confidence intervals on the size of the Markov violation. There were significant violations of the Markov assumption for both the chain and common cause.

The Markov Assumption also has a role in the common effect [$X\rightarrow Y\leftarrow Z$] structure; X and Z are independent of each other when the state of Y is <u>not</u> known, which means that $P(x=1|z=1) = P(x=1|z=0)$. In this study, both of these two judgments should be .50, and indeed they were very close; the mean for $P(x=1|z=1)$ was .52, and for $P(x=1|z=0)$ was .51, 95% CI of difference = [-.02, .05]

We also assessed whether a small minority of participants were responsible for the Markov violations, or whether violating the Markov Assumption was a common habit. Each participant made 12 inferences relevant to the Markov Assumption (across the three structures, the symmetrical judgments, and when $y=1$ vs. $y=0$ for the chain and common cause). For each participant we conducted a t-test comparing the 6 inferences when the irrelevant variable was 1 against the six inferences when the irrelevant variable was 0. Out of a total of 51 participants, 44 gave higher inferences when the irrelevant variable was 1 than 0, and for 19 participants this effect was significant (despite the fact that each t-test was computed with only 12 judgments). If there really is no overall tendency to violate the Markov Assumption, given a bidirectional α=.05, with 51 subjects only 1 or 2 participants should have a significant positive Markov violation merely due to chance. Only 7 had averages that went in the opposite direction, and none of those were significant. In sum, the habit to violate the Markov assumption is common.

Explaining Away The left side of Table 3 shows the normative calculations and empirical means for the three inferences pertinent to explaining away. The right side shows confidence intervals of the difference between the two inferences such as $P(x=1|y=1,z=0)$ - $P(x=1|y=1,z=1)$ that provide the crucial tests of explaining away. The confidence intervals were calculated using mixed linear regressions with by-subject random effects on the intercept and the slope (the difference between the two judgments) to account for repeated measures. The lower bound of the confidence interval identifies whether the amount of explaining away is significantly higher than zero and the upper bound identifies whether the amount of explaining away is significantly lower than the normative amount.

The inferences for $P(x=1|y=1)$ were on average <u>lower</u> than the inferences for $P(x=1|y=1,z=1)$, <u>not higher</u> as implied by the normative model. The inferences for $P(x=1|y=1,z=0)$ and $P(x=1|y=1,z=1)$ were not significantly different; normatively $P(x=1|y=1,z=0)$ should be higher.

Table 3. Explaining Away Results in Study 1.

	Raw Inferences		Explaining Away Comparisons	
Inference	Norm.	Empi-rical	Norm	95%CI of Difference
$P(x=1\|y=1,z=1)$	.60	.70	-	-
$P(x=1\|y=1)$	.71	.59	.11	[-.16, -.06]
$P(x=1\|y=1,z=0)$	1	.69	.40	[-.08, .06]

In summary, Study 1 found that even when participants experience learning data that instantiates the statistical relations among the variables, people still commit violations

[2] Because it was not always possible to transform the data to normal distributions, all analyses were also run using mixed effect logistic regressions using median splits. For example to compare the inferences $P(x=1|y=1,z=1)$ and $P(x=1|y=1,z=0)$, we took the median judgment across both types of inference, recoded inferences above the median as 1, below as 0, and then used a logistic regression with the same random effects structure as above to test whether $P(x=1|y=1,z=1)$ was more likely than $P(x=1|y=1,z=0)$ to have 1s. These two methods produced very similar results, if anything the median split analysis tended to produce cleaner results. However, we report the normal regressions because they can be back-transformed onto the probability scale which aids interpretability.

of the Markov assumption and still do not show a tendency towards explaining away. We have also run another version of this study with more extreme parameters, which shows that people are sensitive to the parameters (they are paying attention to the learning data), yet they still show violations of the Markov assumption and they show better but still suboptimal explaining away.

Study 2: Continuous Variables

The purpose of Study 2 is to test the same phenomena but with Gaussian-distributed variables. Very little research on causal inference has investigated variables on an ordinal, interval, or ratio scale generally. The normative model we use is linear regression.

It is not possible to directly compare reasoning with binary vs. Gaussian data because the multivariate distributions imply different inferences. However, one study that allowed for close comparisons revealed a shift from exemplar memory (binary) to cue abstraction (continuous) (Juslin, Olsson, & Olsson, 2003). In our studies, both for binary and continuous variables, both exemplar and cue abstraction processes lead normative inferences. Still, evidence of different reasoning processes raises the possibility that the Markov violations and insufficient explaining away may not generalize to Gaussian variables.

Markov Assumption One reason people might violate the Markov Assumption is because they believe that the variables are not perfectly observed when they are presented in a "coarse" binary manner (Rehder & Burnett, 2005, called this the "uncertainty model"). Consider the chain $X \rightarrow Y \rightarrow Z$. Suppose you are told that Y is present but Z is absent and you are asked to infer X. Suppose further that you believe that X, Y, and Z can actually assume any state from 0-100, and "present" refers to a value greater than or equal to 50 and "absent" refers to a value less than 50. Given that the binary states of Y and Z conflict (y=present, but z=absent), one might presume that both Y and Z are close to 50. In that case one might infer that X is also fairly close to 50. However, if you are told that y=present and z=present, you might assume that they are both strongly present (e.g., maybe somewhere near 75), and then infer that X is strongly present. In summary, if people view binary variables as coarse simplifications of variables that are actually magnitudes, one plausible hypothesis is that people will be more likely to respect the Markov Assumption when reasoning about magnitude variables.

Explaining Away It is possible that explaining away might be easier to understand with Gaussian variables. Going back to the example in the introduction in which Y is the average of X and Z, it should be fairly obvious that X and Z must be on opposite sides of Y. This heuristic that X and Z tend to be on the opposite sides of Y captures the basic idea that X and Z are negatively dependent given Y.[3] In fact, Nisbett and Ross suggested that a simple "hydraulic heuristic" was relied on in some circumstances, "as if causal candidates competed with one another in a zero-sum game" (1980, p. 128), though this was not empirically tested. In sum, explaining away might be easier with Gaussian variables.

Methods

Participants Fifty undergraduates were paid $12 per hour; the study lasted 32 minutes on average. They were also paid 10 cents for each judgment accurate within 3 points on either side of the correct response.

Stimuli and Design and Procedure. Study 2 was similar to Study 1 except in the following ways. The learning data comprised 32 trials, but the three variables were integers on the scale [0-100]. The three variables were normally distributed with a mean of 50 and a standard deviation of 20 (constrained such that the minimum and maximum were 0 and 100). For all three causal structures $r_{XY} = r_{YZ} = 2/3$.

During the test phase the values of the known variables were chosen randomly from a multivariate normal distribution with the same parameters as in the learning phase. Participants entered in a response on a scale 0-100. The parameters in the learning data mean that for the chain and common cause, a linear regression $X \sim Y+Z$ on the learning data would produce a regression weight of 2/3 for Y and 0 for Z (due to the Markov assumption). For the common effect structure, a linear regression $X \sim Y+Z$ would produce regression weights of 6/5 for Y, and -4/5 for Z (the negative captures the explaining away). These regression weights are taken as the normative answers.

Results

Markov Assumption For the chain and common cause, the Markov Assumption was tested by running regressions to test whether Z had any effect on the inference of X when the state of Y is known $E(X|Y=y,Z=z)$. The regression was fit with by-subject random effects on the intercept and random effects on the slopes of Y and Z to account for the repeated measures within subjects. Table 3 shows the 95% CIs for Y and Z. Even though the regression weight for Z should have been 0 – it should have had no influence at all above and beyond Y, it had a positive influence for the common cause. The influence of Y was trending in the positive direction but was not significant for the chain. For the common effect structure, a regression was used to test whether Z had any influence on X when the state of Y was not known, and it did have a significantly positive effect.

Table 4. Regression Weights for Markov Assumption in Study 2.

Weight	Norm.	$X \rightarrow Y \rightarrow Z$	$X \leftarrow Y \rightarrow Z$	$X \rightarrow Y \leftarrow Z$
Y	0.66	[.48, .75]	[.32, .61]	-
Z	0.00	[-.06, .20]	[.06, .31]	[.23, .48]

Explaining Away Explaining away was tested with a linear regression with the cues Y and Z as predictors of X for the common effect. The 95% confidence intervals were

[3] This heuristic would not produce perfect responses in our study, but would capture the fundamental idea of explaining away.

[0.87, 1.20] for *Y,* and [-0.42,-0.14] for *Z*. The fact that the confidence interval for *Z* is entirely less than zero implies that there was a significant explaining away effect. However, the lower end of the confidence interval, -0.42, is only about half as low as the normative value of -0.80.

Figure 9 shows a histogram of coefficients for *Z* when separate regressions are run for each participant. There is considerable variance and skew, but the distribution supports a more optimistic view of participants' adherence to normative explaining away principle than the overall regression with random effects on the slope. That is, the mode of the distribution is around -0.50 and the median was -.44. Still, only 5 out of the 50 participants had regression weights less than the normative value of -0.80, though the coefficients should be centered on -0.80 if participants were following the normative principle.

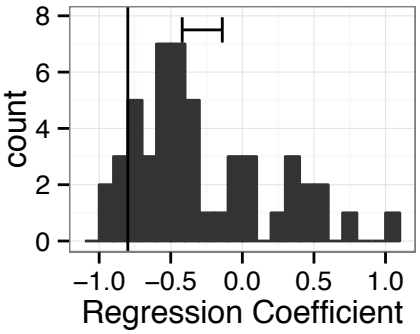

Figure 2. Explaining Away in Study 2. The vertical bar is the normative regression weight and the horizonal bar is the 95% confidence interval.

General Discussion

Two experiments tested whether people could ignore statistically irrelevant variables when making causal predictions (Markov assumption) and whether they could accurately predict the state of one cause given knowledge of its effect and an alternative cause (explaining away). These experiments are unique in that they 1) gave participants learning data that instantiated the statistical relations, 2) incentivized participants for correct responding with monetary rewards, and 3) tested both binary and Gaussian variables. Additionally, even though trial-by-trial learning is one the most common paradigms for studying causal reasoning, it has only been used once previously to examine Markov violations with binary variables (Park & Sloman, 2013 Experiment 3), not for Gaussian variables or explaining away. For the most part, our participants continued to violate the Markov assumption and did not explain away sufficiently, if at all.

Despite the insufficient explaining away observed here, explaining away may occur in other situations. First, it may appear when reasoning about very rare and or very strong causes, parameters not tested here. Second, it might be facilitated by reasoning about concrete mechanisms (Ahn & Bailenson, 1996). Third, it may arise due to domain-specific heuristics. For example, a doctor might conceptualize two rare diseases as essentially mutually exclusive.

These studies imply that even with experience with the multivariate distribution people still have difficulties making accurate judgments. In the future it will be important to identify ways to facilitate normative inference.

References

Ahn, W., & Bailenson, J. (1996). Causal attribution as a search for underlying mechanisms: an explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology*, *31*, 82–123.

Christensen-Szalanski, J. J., & Beach, L. R. (1982). Experience and the base-rate fallacy. *Organizational Behavior and Human Performance*, *29*, 270–8.

Juslin, P., Olsson, H., & Olsson, A.-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology. General*, *132*, 133–156.

Kelley, H. H. (1972). Causal Schemata and the Attribution Process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the Causes of Behavior* (pp. 151–174). Morristown, NJ: General Learning Press.

Nisbett, R. E., & Ross, L. (1980). *Human inference: strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Park, J., & Sloman, S. a. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology*, *67*, 186–216.

Rehder, B. (2014a). Independence and Dependence in Human Causal Reasoning. *Cognitive Psychology*.

Rehder, B. (2014b). The Role of Functional Form in Causal-Based The Role of Functional Form in Causal-Based Categorization. *Journal of Experimental Psychology : Learning , Memory , and Cognition*.

Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, *50*, 264–314.

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*, 109–39.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Sussman, Abigail, B., & Oppenheimer, D. (2011). A Causal Model Theory of Judgment. In C. Hölscher, Carlson & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1703–1708). Austin, TX: Cognitive Science Society.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: two modes of accessing causal knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*, 216–27.