Temporal Causal Strength Learning with Multiple Causes

Cory J. Derringer (cjd78@pitt.edu)
Benjamin M. Rottman (rottman@pitt.edu)
Department of Psychology, University of Pittsburgh
3939 O'Hara Street, Pittsburgh, PA 15260 USA

Abstract

When learning the relation between a cause and effect, how do people control for all the other factors that influence the same effect? Two experiments tested a hypothesis that people focus on events in which the target cause changes and all other factors remain stable. In both four-cause (Experiment 1) and eight-cause (Experiment 2) scenarios, participants learned causal relations more accurately when they viewed datasets in which only one cause changed at a time. However, participants in the comparison condition, in which multiple causes changed simultaneously, performed fairly well; in addition to focusing on events when a single cause changed, they also used events in which multiple causes changed for updating their beliefs about causal strength. These findings help explain how people are able to learn causal relations in situations when there are many alternative factors.

Keywords: causal learning; causality; causal strength; conditionalizing

Introduction

Learning the strengths of causal relationships is essential for successfully manipulating the environment. Consider a student deciding whether to take a GRE prep course. Knowing the extent to which taking the course would improve or harm her score would help the student decide whether the course is worthwhile.

However, causal relationships do not exist in isolation; often multiple causes influence the same effect in different ways (e.g., whether the student has been taking practice tests or drinking at the local pub). Further, the causes could be confounded; one more hour studying may mean one less hour sleeping. In the same way statisticians use regression to estimate the effect of one variable above and beyond another, individuals should attempt to control for alternative causes to estimate the unique strength of a target cause.

In the next section we briefly review the dominant theory for how people control for alternative causes. We then propose a new theory called the Informative Transitions heuristic, and report two experiments that tested this theory.

Controlling for Alternatives Using Focal Sets

Focal set theory (Cheng & Novick, 1990; Cheng & Holyoak, 1995) proposes that when assessing the strength of a target cause (A) on an effect, people examine a subset of the data in which the other factor(s) are constant. For example, given the data in Figure 1, a learner could choose the subset in which B=0 (the gray shading). Within this subset a learner could calculate causal strength. (There are multiple theories for how people calculate causal strength within a focal set; see Hattori and Oaksford (2007) for a summary.)

Time	1	2	3	4
Cause A	1	1	0	0
Cause B	0	1	0	1
Effect	1	1	0	1

Figure 1: A focal set (gray shading) for cause A conditional on cause B's absence

It has been widely demonstrated that people often control for alternative causes when estimating the strength of a target cause (Spellman, 1996; Spellman, Price, & Logan, 2001; Waldmann & Holyoak, 1992; Waldmann, 2000; Waldmann & Hagmayer, 2001). However, the precise mechanism(s) is less clear. We first review two limitations of focal sets, and then discuss other options.

First, when the alternative cause B is a binary variable, two focal sets may be used: the presence or absence of B. Using these different focal sets, the subject could come to different conclusions about the strength of the target cause. In Figure 1, the focal set B=0 implies that A causes the effect (A is correlated with the effect in this focal set), but the focal set B=1 implies that it does not.

The second problem is that it is unclear exactly how people would use focal sets when there are multiple alternative factors. Imagine a situation in which there are four possible causes (A-D) of an effect. When assessing the causal strength of A, a learner could choose a focal set such as (B=0, C=0, D=0); however, using this focal set ignores the other 7 possible combinations of B, C, and D, which means ignoring 7/8ths of the potential data. If there are 8 causes (Experiment 2), choosing any one particular focal set (e.g., when all the other factors are 0), involves ignoring 127/128ths of the potential combinations. If a learner only experiences a small number of observations, there may not be observations in which A=1 and A=0, while all the other factors are 0; the relation between A and the effect cannot be inferred. Thus, selecting one focal set for each cause seems an inefficient strategy. In sum, there is a tension between using just one focal set, which involves discarding data, and using multiple, which requires integrating the conclusions in some yet-unspecified way.

Controlling for Alternatives in Other Ways

Here we list a couple alternative approaches to controlling for alternative causes aside from focal sets. First, in many learning situations, associative theories (e.g., Rescorla &

Wagner, 1972) asymptote to focal sets (see Spellman, 1996). One prediction of the standard version of Rescorla and Wagner's learning rule (RW) is that the associations between a cue and outcome get updated only on trials when the cue is present, which we will assess in the following experiments.

Second, when there are too many alternatives, people may stop controlling for alternative causes due to a working memory overload (Goedert, Harsch, & Spellman, 2005), in which case the causal strength estimates would resemble the bivariate relation between each cause and the effect.

Third, though we are not proposing it as a descriptive theory, it is useful to consider multiple regression as a gold standard computational-level theory, especially when there are many alternatives that need to be controlled for.

Using Informative Transitions

The current experiments test an alternative way that people may calculate conditional causal strength judgments as they experience data over time, which we call the Informative Transitions heuristic. We propose that people leverage the fact that causes in the environment are often autocorrelated—that they often remain in the same state for extended periods of time[1]. For example, if someone is in a bad mood in the morning, they are likely to be in a bad mood at lunchtime. Some degree of stability occurs for many other variables (the economy, one's health etc.). The data in Table 1a are an example in which each of the four causes is fairly stable over time. When causal factors change fairly infrequently, causes often change in isolation (provided that they are independent).

These transitions from one timepoint to the next in which only one potential cause changes and the others remain stable (hereon "informative transitions", or "IT"s) can give the learner unique insight into the relationship between the changing cause and the effect. Any change or lack of change in the effect during an IT informs the learner about the relationship between that cause and the effect holding the other causes constant. If the cause and effect change in the same direction (positive IT), it provides evidence that the cause is positive (generative) such as Table 1a Time 6-7 for Cause A. If they change in the opposite direction (negative IT), then the cause is more likely to be negative (inhibitory) (e.g., Table 1a, Time 4-5 for Cause D). If the cause changes but the effect stays constant (neutral IT), it may not be a cause at all (Time 3-4 for Cause B in Table 1a).

According to the IT heuristic, transitions in which multiple causes change simultaneously (Table 1b) are less informative; a change in the effect cannot be attributed to a single cause. If people use the IT heuristic, they will learn about causal strength more from ITs (Table 1a) than when multiple causes change (Table 1b). By contrast, focal set theory predicts no differences based on trial order.

The IT heuristic retains the most compelling aspect of the focal set theory: it simplifies scenarios with many causes, allowing the subject to examine the bivariate relationship between each cause and the effect. In a sense, the IT heuristic can be viewed as dynamically using different focal sets at different times (e.g., for Time 1-2 in Table 1a, the target cause is B, and the focal set is A=1, C=0, D=1; for Time 2-3 the target cause is C, and the focal set is A=1, B=0, D=1, etc.).

Table 1a: Dataset wherein all transitions are informative

Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Cause A	1	1	1	1	1	0	1	1	1	1	1	1	1	0
Cause B	1	0	0	1	1	1	1	1	1	0	0	0	1	1
Cause C	0	0	1	1	1	1	1	1	0	0	1	1	1	1
Cause D	1	1	1	1	0	0	0	1	1	1	1	0	0	0
Effect	0	0	0	0	1	0	1	0	0	0	0	1	1	0

Table 1b: Data from Table 1a, rearranged to remove ITs

Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Cause A	1	1	0	1	1	1	1	1	1	1	1	1	0	1
Cause B	1	0	1	1	0	1	1	0	1	0	1	0	1	1
Cause C	1	0	1	1	1	0	1	1	1	0	1	1	1	0
Cause D	1	1	0	1	0	1	0	1	0	1	0	1	0	1
Effect	0	0	0	0	1	0	1	0	1	0	1	0	0	0

The current experiments test the IT heuristic by changing the order of the learning data to manipulate the number of ITs. Tables 1a and 1b are an example; they have the same 14 trials, but in a different order. If learners use the IT heuristic, they will be more successful at detecting the underlying causal relations in datasets like Table 1a than 1b.

Experiment 1

Method

Participants For each experiment, 100 participants located in the United States were recruited through Amazon's Mechanical Turk service. Each participant had previously completed at least 100 tasks on MTurk and had an approval rate of at least 95 percent. Participants were paid two dollars, with accuracy bonuses as an incentive. The study lasted approximately 15 minutes.

Design We examined how participants learn about positive causes, negative causes, and neutral factors (non-causes). We created three types of datasets so that participants would not have a strong expectation of the number of each type of cause: one positive, one negative, two neutral; two positive, one negative, one neutral, and one positive, two negative, one neutral. Twenty datasets were created for each type. Each participant viewed six datasets (two of each type), randomly selected and presented in random order.

[1] In reality, the number of causes that change on average from one observation to the next depends on how many causes are being observed and how frequently the causes change on average. If there is a very large number of causes, then typically more than one will change. However, if the causes are very stable, then it is possible for only one (or even less than one) cause to change on

The independent variable (between subjects) was the order of the trials within a dataset: whether each transition was informative (Table 1a) or not (Table 1b).

The datasets were created manually in the following way. On each trial, one cause was selected to change, and the effect was changed in the appropriate direction. For example, if a positive cause was selected to change, the effect would change in the same direction. By contrast, when a negative cause changed, the effect changed in the opposite direction. When a neutral cause changed, the effect did not change. Each cause changed either three or four times overall, and there were 14 trials for each dataset.

With a logistic regression predicting the presence/absence of the effect, this scheme guaranteed that positive and negative causes always had strong logistic regression weights (51.13 and -51.13 respectively), while neutral causes would have weights approaching zero[2] (Table 5).

We created datasets for the No-IT condition by reordering the datasets from the IT condition to remove all of the informative transitions (e.g., Table 1b). For any given transition either no causes change, or two or more change. Still, the regression weights remained the same. For example, in Tables 1a and 1b, Cause A is positive, B and C are neutral, and D is negative.

Procedure Participants were told to imagine that they worked in a nursing home, and were in charge of making sure each patient slept well. Their task was to examine the medications taken by each patient to find out which ones enhanced or interfered with sleep. Participants observed 14 sequential days of data for each patient. For each day, participants were shown which of four medications the patient took that day. Medications were represented using animated pictures of pills, which were either full-color (cause present) or visibly faded (cause absent). These presence/absence cues for each cause remained on-screen for the duration of the trial. Participants predicted whether the patient would sleep well or poorly that night, and received feedback immediately. They were then allowed to adjust sliding scales indicating their beliefs about the causal status of each medication. When they were satisfied with their current judgments, they clicked a button to proceed to the next trial, and the on-screen presence/absence cues changed to reflect the changes during that transition (pictures either faded or came back to full color for causes that changed).

The sliders used the scale -10 (strong sleep inhibitor) to 10 (strong sleep enhancer). The numerical values on the scale were given to the participants during the instruction phase. After viewing data from 14 timepoints, participants submitted final judgments about the influence of each medication using the same scale. This procedure was then repeated for five more patients, using different datasets. After completing the study, participants' bonus amount was calculated and paid. Bonus amounts were calculated based on participants' final judgments for each cause. Five cents were paid for final judgments that were sufficiently accurate, and the maximum possible bonus was $1.20.

Results

Trial-By-Trial Updating of Causal Strength Beliefs Recall that participants had the opportunity to update their current beliefs about each drug's effectiveness after every trial. If transitions are important for causal strength learning, participants would be more likely to update causes that change. For this analysis, we focused on the first five transitions, because most of participants' learning took place toward the beginning of each scenario; the probability of updating a given cause was .27 after Trial 1 and .09 after Trial 14.

We tested whether participants were more likely to update their beliefs about causes that had just changed vs. causes that did not change using a logistic regression, which included a by-subject random intercept and a by-subject random slope for whether the cause changed. For this model, we analyzed a subset of the data that only included positive and negative causes on trials in which the effect changed. Participants in both the IT (B = 1.10, SE = 0.09, p < .001) and No-IT conditions (B = 0.27, SE = 0.09, p < .01) were significantly more likely to update their causal strength beliefs for causes that changed than causes that did not change (Table 2).

Table 2: Trial-by-Trial Probability of Updating Causal Strength Based on Whether the Cause Changed.

	Experiment 1		Experiment 2	
Cause	*IT*	*No-IT*	*IT*	*Random*
Changed	.35	.21	.38	.27
Did Not Change	.17	.18	.09	.13

Participants in the No-IT condition were still more likely to update their beliefs about causes that changed on the most recent trial, even though multiple causes changed simultaneously. This effect highlights one reason we expected people to perform better in the IT condition: participants in the No-IT condition received worse information about the relationships between the causes and the effect from transitions, but were still more likely to update their beliefs to reflect that information.

We used similar regressions to test the RW prediction that participants would be more likely to update present causes than absent ones (Table 3). Participants in the IT (B = 2.66, SE = 0.26, p < .001) and No-IT conditions (B = 2.36, SE = 0.22, p < .001) updated more often for causes that were present.

[2] For a technical and non-obvious reason, our particular data generation process guaranteed that if all the causes were entered into a linear regression to predict the effect, positive, negative, and neutral causes always had regression weights of 1, -1, and 0 respectively; the positive and negative causes were very strong.

Table 3: Trial-by-Trial Probability of Updating Causal Strength Based on Whether the Cause was Present.

	Experiment 1		Experiment 2	
Cause	*IT*	*No-IT*	*IT*	*Random*
Present	.28	.27	.23	.35
Absent	.06	.04	.03	.02

Trial-By-Trial Accuracy of Predictions of the Effect Given that participants were more likely to update their beliefs about the causes that changed in the most recent trial compared to those that did not, it is clear that transitions play a role in temporal causal strength learning. Next we analyzed the differences in accuracy for predicting the effect on each trial. Recall that on each trial, participants were asked to predict whether or not the patient would sleep well. We predicted that participants in the IT condition would be more accurate than those in the No-IT condition.

We tested this using a logistic regression, predicting accuracy on a given trial using trial number, condition (IT vs No-IT), and an interaction between the two. We included a by-subject random intercept and by-subject random slope for trial number, allowing for individual differences in how accuracy changes over the course of the trials. Figure 2 summarizes our findings: participants in the IT condition improved more over time than participants in the No-IT condition (B = 0.05, SE = 0.01, p < .01). This interaction seems to come from the surprising finding that accuracy in the IT condition started out lower than the accuracy in the No-IT condition; this pattern was not found in Experiment 2, so we do not dwell on it. The overall accuracy difference between participants in the IT condition (58.5% correct) and those in the No-IT condition (59.8% correct) was nonsignificant.

Final Judgments of Causal Strength If ITs help participants learn causal strengths, then participants' final judgments of causal strength in the IT condition would be more positive for the positive causes, and more negative for the negative causes, relative to the No-IT condition. Because the data were heavily skewed, with many judgments near the positive extreme for positive causes and the negative extreme for negative causes, we used a Gamma GLM. For positive causes, we transformed the data by multiplying each judgment by -1 and adding 11. For negative causes, we added 11 to each judgment. This ensured that the shape of each distribution would not change, but that they could be mapped onto a gamma distribution, which required all values to be positive. We incorporated a by-subject random intercept to allow for individual differences in subjects' baseline judgments. Participants in the IT condition judged the positive causes to be more positive (B = 0.04, SE = 0.01, p < .01), and the negative causes to be more negative (B = -0.05, SE = 0.02, p < .001), relative to the No-IT condition (Table 4).

Table 4: Mean (SD) of Final Causal Strength Judgments.

	Experiment 1		Experiment 2	
Cause	*IT*	*No-IT*	*IT*	*Random*
Pos.	5.4 (4.5)	3.5 (5.3)	5.4 (5.1)	3.2 (5.8)
Neg.	-5.0 (5.1)	-2.7 (5.5)	-5.2 (5.8)	-3.9 (5.6)

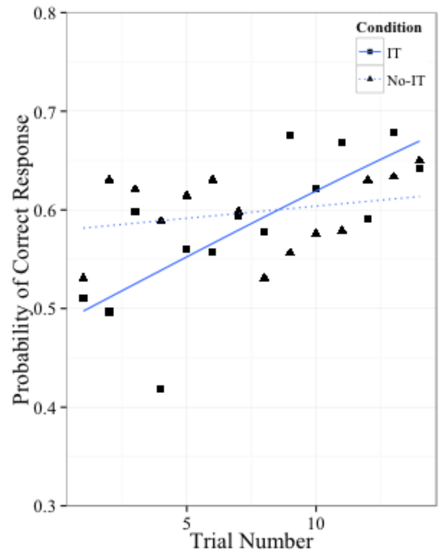

Figure 2: Accuracy of Trial-By-Trial Predictions of Effect in Experiment 1.

Discussion

Participants formed and updated their causal strength beliefs based on which causes changed during transitions. Additionally, participants' predictions of the effect improved more in the IT condition than in the No-IT condition, and participants in the IT condition made more accurate final judgments of causal strength. Together, these findings support the idea that transitions facilitate causal learning.

Interestingly, although participants in the No-IT condition were less accurate in their final judgments, they were clearly learning, and the judgments were significantly different from 0 in the correct directions for both positive and negative causes (p's < .001). This pattern of results raises the question of how participants in the No-IT condition were able to perform as well as they did. As outlined in the introduction, if participants used focal sets, it is not clear exactly what focal sets they were using. One possibility is that the bivariate correlations between each cause and the effect in our datasets were moderately strong (Table 5), so it is possible that participants in the No-IT condition used the bivariate relations and did not control for alternative causes.

Table 5: Regression weights and average bivariate correlations for causes in Experiments 1 and 2

Analysis	*Positive*	*Neutral*	*Negative*
Experiment 1			
Log. Regression Weight	51.13	< 0.001	-51.13
Bivariate Correlation	0.48	0.00	-0.47
Experiment 2			
Log. Regression Weight	51.13	< 0.001	-51.13
Bivariate Correlation	0.23	0.01	-0.22

Experiment 2

In Experiment 2, we attempted to test the limits of causal strength learning by having participants learn causal strengths in a scenario with eight causes. Having eight causes also reduces the bivariate correlations relative to Experiment 1 (Table 5), thereby making the task harder and potentially allowing for a bigger difference between the two conditions. However, the multiple regression weights were still just as strong for the positive and negative causes.

Another advantage of adding more causes in Experiment 2 is that an eight-cause scenario is even more difficult to explain in terms of focal sets. As previously discussed, as the number of causes increases, a smaller fraction of the potential data would be included in any particular focal set.

Importantly, the order of the trials does not matter for the focal set account. By contrast, if participants use the IT heuristic, they will perform better with more ITs.

Method

Participants One hundred workers were recruited through Amazon's Mechanical Turk service using the same qualification criteria from Experiment 1. Participants were paid two dollars, with accuracy bonuses as an incentive. The study lasted approximately 15 minutes.

Design and Procedure There were several differences compared to Experiment 1. Each participant made judgments about eight causes instead of four, over 25 days instead of 14. To compensate for the increased length of each scenario, participants only completed three scenarios instead of six. The IT datasets were created using a script rather than manually, and each cause changed three times. Additionally, rather than creating No-IT datasets, we randomized the order of the trials as a comparison for the IT condition. This means that sometimes there were informative transitions (when only one cause changed) in the Random condition. The modal number of ITs per dataset in the Random condition was four—the maximum number was nine—compared to 24 in the IT condition. On average, 3.36 causes changed for each transition in the Random condition.

We used three types of datasets. The first type had three positive causes, three negative causes, and two neutral causes. The second type had two positive, three negative, and three neutral causes. The third had three positive, two negative, and three neutral causes. Each participant viewed one dataset from each type, with the order of the datasets being randomized. Aside from these changes Experiment 2 was the same as Experiment 1.

Results

Trial-By-Trial Updating of Causal Strength Beliefs Experiment 2 replicated the pattern of results from Experiment 1 (Table 2). Participants updated their beliefs about causal strength more when a given cause changed compared to when it did not change. Furthermore, this occurred in both the IT condition (B = 2.03, SE = 0.14, p < .001) and the Random condition (B = 1.00, SE = 0.10, p < .001). In addition, they also were more likely to update their causal strength judgments when a cause was present in the IT (B = 3.01, SE = 0.30, p < .001) and Random conditions (B = 5.49, SE = 0.79, p < .001) (Table 3).

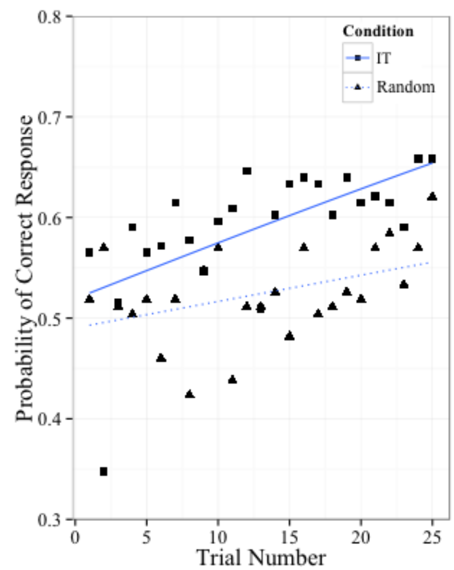

Figure 3: Accuracy of Trial-By-Trial Predictions of Effect in Experiment 2.

Trial-By-Trial Accuracy of Predictions of the Effect We hypothesized that if participants are better able to infer causal strength in the IT condition, they would also be better able to predict the effect. Figure 3 plots the probability of correctly predicting the effect.

A logistic regression with by-subject random intercepts found that participants were more accurate in the IT condition than in the Random condition (B = 0.27, SE = 0.06, p < .001). The interaction from Experiment 1 was marginal in Experiment 2 (B = -0.01, SE = .01, p = 0.06).

Final Judgments of Causal Strength As in Experiment 1, we examined the differences in participants' final judgments of positive and negative causes in the IT and Randomized conditions (Table 4). A mixed effects Gamma regression with random intercepts for each participant found stronger causal strength judgments in the IT condition for both positive ($B = 0.08$, $SE = 0.02$, $p < .001$) and negative causes ($B = -0.07$, $SE = 0.02$, $p < .01$). Participants in the Random condition performed significantly above chance (p's < .001). Overall, the results were similar to Experiment 1.

General Discussion

Our main hypothesis was that participants learn causal strengths better in stable environments (IT condition). Indeed, participants' final causal strength estimates were stronger—closer to the normative regression weights—in these environments, suggesting that participants capitalize on the stability of the environment to learn causal strengths.

More generally, participants were more likely to update their beliefs about a cause immediately after that cause changed, even in learning environments in which more than one cause changed from one observation to the next. This finding builds upon research concerning how people learn cause-effect relations in single-cause environments; that research also found that people are more likely to update beliefs about a cause after it has changed (Soo & Rottman, 2015). The current paper generalizes this finding to situations with multiple causes and demonstrates the effectiveness of this learning habit in stable environments.

Our finding that participants update their beliefs more often after a cause changes parallels another finding: participants updated their beliefs more for present causes than absent ones. RW predicts this as well (cf. Van Hamme & Wasserman, 1994).

Several questions are as yet unanswered. Participants in the No-IT and Random conditions could infer causal direction, despite low bivariate relations in Experiment 2. Understanding this may yield other causal learning insights.

Additionally, there are some differences in the results of Experiment 1 vs. 2, particularly in the accuracy of the trial-by-trial predictions of the effect. In Experiment 2 the accuracy in the IT condition was generally better than the Random condition (Figure 3). By contrast, in Experiment 1, the accuracy in the IT condition started lower than in the No-IT condition, and rose more quickly (Figure 2). This raises the possibility that informative transitions may, in certain situations, temporarily impair learning. Aside from the main difference of 8 vs. 4 causes, there are other differences between the two experiments that could explain these different patterns: weaker bivariate relations and more learning trials in Experiment 2, and the fact that the Random condition in Experiment 2 did not exclude informative transitions. Investigating these factors can elucidate whether situations with many informative transitions may temporarily result in worse causal learning.

Individuals often face situations in which multiple causes can influence an effect. One way people may cope with this complexity is by focusing on times when a cause changes, especially in situations with fairly stable causes over time.

Acknowledgments

This research was supported by NSF 1430439.

References

Cheng, P. W., & Holyoak, K. J. (1995). Complex Adaptive Systems as Intuitive Statisticians: Causality, Contingency, and Prediction. In H. L. Roitblat & J. -A. Meyer (Eds.), *Comparative Approaches to Cognitive Science*. Cambridge, MA, US: The MIT Press.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*(4), 545-567.

Goedert, K. M., Harsch, J., & Spellman, B. A. (2005). Discounting and conditionalization: Dissociable cognitive processes in human causal inference. *Psychological Science*, *16,* 590-595.

Hattori, M., & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science, 31*(5), 765-814.

Rescorla, R., & Wagner, A. R. (1972). A Theory on Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Theory and Research* (pp. 64-99). New York, NY: Appleton-Century-Crofts.

Soo, K., & Rottman, B. M**.** (2015). Elemental causal learning from transitions. In R. Dale, et al. (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7*(6), 337-342.

Spellman, B. A., Price, C. M., & Logan, J. M. (2001). How two causes are different from one: The use of (un)conditional information in Simpson's paradox. *Memory & Cognition, 29*(2), 193-208.

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

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology, 26*(1), 53-76.

Waldmann, M. R., & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition, 82*(1), 27-58.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*(2), 222-236.