Causal Learning With Continuous Variables Over Time

Kevin W. Soo (kevin.soo@pitt.edu)
Benjamin M. Rottman (rottman@pitt.edu)
Department of Psychology, University of Pittsburgh
3939 O'Hara Street, Pittsburgh, PA 15260 USA

Abstract

When estimating the strength of the relation between a cause (X) and effect (Y), there are two main statistical approaches that can be used. The first is using a simple correlation. The second approach, appropriate for situations in which the variables are observed unfolding over time, is to take a correlation of the change scores – whether the variables reliably change in the same or opposite direction. The main question of this manuscript is whether lay people use change scores for assessing causal strength in time series contexts. We found that subjects' causal strength judgments were better predicted by change scores than the simple correlation, and that use of change scores was facilitated by naturalistic stimuli. Further, people use a heuristic of simplifying the magnitudes of change scores into a binary code (increase vs. decrease). These findings help explain how people uncover true causal relations in complex time series contexts.

Keywords: causal learning; causal reasoning; time

Introduction

Knowing the strength underlying cause-effect relationships (e.g. how strongly a drug suppresses a symptom) allows people to decide which causes to use to achieve desired outcomes (Hagmayer & Meder, 2013). Past research has uncovered various ways people infer the strength of a causal relation after observing the covariation between the cause and effect (e.g. Cheng, 1997; Griffiths & Tenenbaum, 2005; Hattori & Oaksford, 2007). Most research has focused on how people assess causal relations among *binary* causes and effects – usually 'absent' or 'present' (for exceptions, see Pacer & Griffiths, 2011; Rottman, in press; Saito, 2015). Additionally, the temporal order of the observations is often random and typically viewed to be an irrelevant factor by the researchers. The present study investigates causal learning with continuous variables that are observed in time series exhibiting increasing or decreasing trends. To anticipate the findings, we found that reasoning about continuous variables and reasoning about time series trends are interrelated processes; sequentially-presented continuous variables are treated as binary, which simplifies the learning process.

Learning with continuous variables

In the real world, variables are often *continuous* or at least ordinal in scale – e.g. a drug is not either 'present' or 'absent' but administered with a particular dosage. How do people infer causal strength when a cause and effect can each assume multiple levels?

One theory is that people simplify continuous variables by mentally dichotomizing them into binary variables, which would make it easier to summarize the values of the cause and effect for computing causal strength. One study found that people assimilate intermediate values and treat them as either the high or low value on the scale (Marsh & Ahn, 2009). However, the variables in that study were not really continuous; they primarily had a high or low value, and occasionally an intermediate value. It is unknown whether people mentally dichotomize variables when the variables take on values along a fuller range of possible levels, which would require arbitrarily choosing cutoff values.

Another theory is that when judging the causal strength of a continuous cause on a continuous effect, people perform a mental computation similar to Pearson's correlation coefficient r. In fact, in correlation-learning paradigms, peoples' correlation estimates are sensitive to many of the components that are used for calculating r like the slope, error variance and the variance of each variable (Lane, Anderson, & Kellam, 1985). A weakness of r as a model of human learning is that it is a computational-level theory that fails to explain how the learner actually processes the information in a tractable way. Computing r would be computationally intensive; all the X and Y values would need to be remembered and integrated into one score.

In the next section we introduce causal learning in time series contexts, and later discuss how causal learning over time and with continuous variables are interrelated.

Observations over time

There are two standard paradigms for causal learning: situations in which the order of the data is meaningful (e.g., perhaps X or Y undergoes a trend), and situations in which the data have no inherent temporal ordering and are presented randomly. We call these situations "longitudinal" and "cross-sectional", respectively; an analogy to the terms used to describe experimental research designs.

The correlation coefficient r is appropriate for causal learning in cross-sectional contexts, but can be misleading when used in longitudinal contexts. For example, consider the data in the "negative transitions" condition in Figure 1. The order of observations is denoted with the numbers 1-20 in the plot. Both X and Y gradually increase over time, and overall there is a positive correlation between X and Y. However, from one observation to the next there is a negative correlation between X and Y; when X increases, Y decreases, and vice versa. Which is the "right" way to interpret the causal relation between X and Y? Is the strength of the relationship positive or negative?

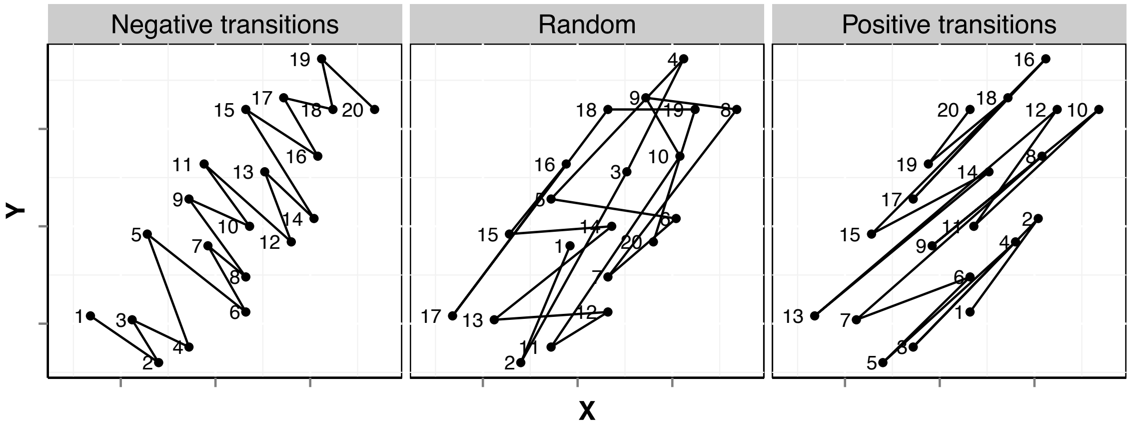

Figure 1: Positively correlated data presented in different orderings (denoted by numbers).

We assert that the latter interpretation based on how the variables change together is appropriate in a longitudinal context. For example, in Figure 1 (Negative transitions), X and Y both increase over time. It would be inappropriate to conclude that X and Y are positively related just because they both increase over time, as many variables exhibit temporal trends. For example, the US economy and the price of oil have generally increased over time (positive correlation), even though increases in the price of oil cause the economy to contract at a smaller time scale. Data sets like this arise naturally whenever 1) a cause and effect pair, X and Y, both increase (decrease) over time due to the influence of a common cause, Z; and 2) X has a negative (positive) causal influence on Y. Figure 2 shows a causal graph depicting this scenario. To estimate the true causal relationship between X and Y, we need to control for Z.

Table 1: Information from longitudinal observations of a cause (X) and effect (Y) relationship. The data are the same as in Figure 1 (negative transitions condition).

Time	States		Transitions			
			$\Delta_{Continuous}$		Δ_{Binary}	
	X	Y	X	Y	X	Y
1	17	27				
2	35	15	18	–12	+1	–1
3	25	26	–7	11	–1	+1
4	43	19	15	–7	+1	–1
5	32	48	–11	29	–1	+1
	•		•		•	
	•		•		•	
	•		•		•	
18	81	80	13	–3	+1	–1
19	78	93	–3	13	–1	+1
20	92	80	14	–13	+1	–1
	$r_{States} = .70$		$r_{\Delta Cont} = -.97$		$r_{\Delta Binary} = -1$	

Examining the *changes* in Y in response to changes in X accounts for secular trends in both variables (i.e. controls for the effect of the common cause influencing them over time). A way to account for secular trends in time series analysis is to take difference scores (Δ) of X and Y, and then compute a correlation on the difference scores; in Table 1 the Δ scores are denoted with $\Delta_{Continuous}$, from which $r_{\Delta Cont}$ is computed. Using Δ scores controls for first-order non-stationarity (linear trends) in a time series (Shumway & Stoffer, 2011).

The main question of this article is whether humans intuitively compute causal strength from the absolute or difference scores of the cause and effect (i.e. based on r_{States} or $r_{\Delta Cont}$). This is tested using datasets like those in Figure 1. The three plots have the same 20 data points so $r_{States} = .70$ for all three, but $r_{\Delta Cont}$ changes from very negative to very positive.

In addition to the fact that $r_{\Delta Cont}$ is better at uncovering the true causal relation in contexts with temporal trends, there is a second reason why $r_{\Delta Cont}$ may be a better model than r_{States} of how humans infer causal strength. Across many sensory modalities (e.g., sound, light, pain), humans encode relative changes rather than the absolute magnitudes of stimuli because our senses adapt to the current level of stimulation (Stewart, Brown, & Chater, 2005). This means that, even though Table 1 suggests that $r_{\Delta Cont}$ requires an additional step of computing Δ scores from raw scores, in naturalistic

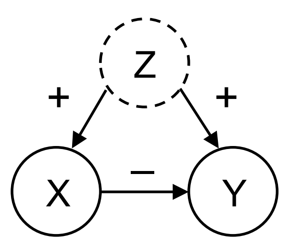

Figure 2: The relationship between X and Y when there is a common cause (Z) influencing both over time. X negatively influences Y. Z causes X and Y to increase over time.

settings Δ scores may be the primary form of data that we have access to through our sensory systems. For this reason, humans may naturally attend more to to Δ scores; making $r_{\Delta Cont}$ a better model of human learning than r_{States}.

Continuous vs. Binary Representation

A question raised above is how a learner could compute a correlation between two variables when it would require remembering all the values of X and Y as well as computationally integrating all those values; $r_{\Delta \mathrm{Cont}}$ would be just as challenging to compute as r_{States}. Earlier, we proposed that in certain contexts learners may discretize continuous variables. In longitudinal situations, there is a natural cutoff that may ease discretization; whether the variable increased (coded as +1), decreased (–1), or stayed the same (0) from the previous time point (Δ_{Binary} in Table 1). Computing the correlation between Δ_{Binary} scores ($r_{\Delta Binary}$) only requires keeping track of the number of times that X and Y increase and decrease together and separately, raising the possibility of discretization as a plausible heuristic.

Experiment 1 tested whether learners infer causal strength based on the absolute values of the cause and effect (similar to r_{States}) in addition to how the cause and effect changed over time (similar to $r_{\Delta Cont}$ and $r_{\Delta Binary}$). Experiment 2 investigated the extent to which presentation format influences the reliance on transitions. Experiment 3 tested whether people use continuous or binary representations of change scores for inferring causal strength.

Experiment 1

Experiment 1 tested whether learners use information about transitions (Δ scores) in addition to information about states (raw scores) when inferring the causal strength between a cause (X) and effect (Y) in a time series context. We predicted a stronger effect of transitions relative to states on causal strength judgments.

Method

Subjects 50 subjects were recruited using Amazon Mechanical Turk (MTurk) and paid $0.60. The experiment lasted approximately 5 minutes.

Design and stimuli Subjects inferred causal strength from data sets consisting of 20 observations of X and Y, where X and Y could take on values ranging from 0 to 100. We manipulated the correlation between the states of X and Y as well as the transitions (see Figure 1) in a 2 (positive vs. negative r_{States}) × 3 (negative vs. random vs. positive transitions) within-subjects design.

Table 2: Means and SDs of $r_{\Delta Cont}$ in different conditions

	$r_{\Delta Cont}$ Mean (SD)		
	Negative transitions	Random transitions	Positive transitions
r_{States} = .7	–.96 (.02)	.73 (.07)	.97 (.01)
r_{States} = –.7	–.97 (.01)	–.73 (.07)	.96 (.02)

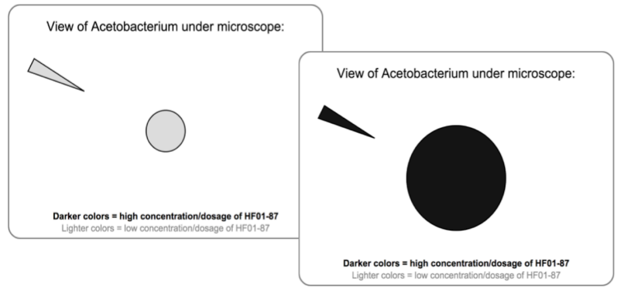

Figure 3: Stimuli viewed by subjects. Two consecutive observations are shown.

The manipulation of r_{States} involved data sets with either positive or negative r_{States}. Data sets with r_{States} = .7 were generated, and copies were made with the values of X flipped around the midpoint of the scale (X = 50), creating data sets with r_{States} = –.7.

The manipulation of transitions was achieved by reordering the observations to produce three conditions (as in Figure 1). In the negative transitions condition, the relationship between the Δ scores of X and Y at every time point is negative – increases in X are always accompanied by decreases in Y, and vice versa. In the positive transitions condition, increases in X are always accompanied by increases in Y. In the random transitions condition, the order of the 20 observations was randomized, resulting in a mix of positive and negative transitions. In the positive (negative) r_{States} condition, most of the transitions are positive (negative) (see Table 2).

20 data sets were created per condition. Each subject viewed one randomly chosen data set in each of the six conditions, and the order of the conditions was randomized.

Procedure Subjects were told they would evaluate how the dosage of a drug (X) affected the size of a microorganism (Y) over 20 observations (‘days’). On each day, a new dosage of the drug was administered to the same microorganism and the size was observed under a microscope. The microorganism was represented using a circle (see Figure 3). Each scenario was presented as data of a different drug-microorganism pair.

The value of X was mapped onto the opacity of the circle, with darker shades representing higher doses. The value of Y was mapped to the diameter of the circle.

After 20 observations, subjects judged the causal strength, on a scale of 8 (“high levels of the drug strongly cause the microorganism to increase in size”) to –8 (“high levels of the drug strongly cause the microorganism to decrease in size”). A rating of 0 indicated no causal relationship.

Results

Causal strength judgments were analyzed with a within-subjects factorial ANOVA (see Figure 4a for means). As expected, subjects rated causal strength higher in the positive r_{States} condition than in the negative condition; $F(1, 288) = 11.21$, $p < .001$, $\eta_p^2 = .04$. The primary question was

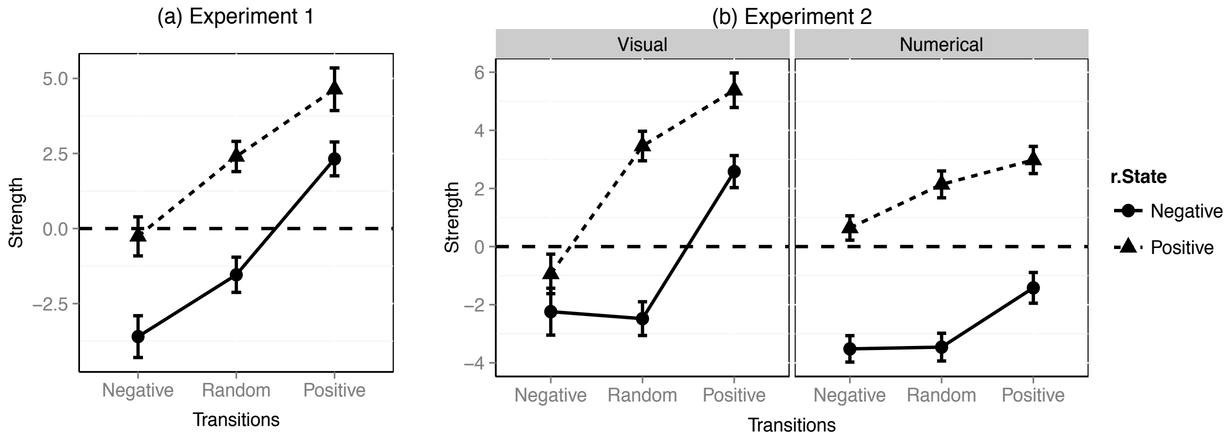

Figure 4: Means by Condition. Error bars represent standard errors.

whether the transitions had an influence on causal strength ratings above and beyond r_{States}. Indeed, subjects rated causal strength highest in scenarios with positive transitions, followed by random transitions, and negative transitions; $F(2, 288) = 20.38$, $p < .001$, $\eta_p^2 = .12$. The effect size of transitions was about three times larger than the effect of states (r_{States}). No interaction effect was observed. In the condition with negative r_{States} but positive transitions, subjects actually judge the causal strength to be positive; the effect of transitions 'overrides' the effect of states in this case (see Figure 4a). With positive r_{States} but negative transitions, subjects' judgments are not significantly different from zero; the negative transitions 'neutralized' the effect of states.

Experiment 2

In the introduction we argued that people may use transitions for estimating causal strength when the cause and effect are presented as naturalistic perceptual stimuli (such as in Experiment 1); adaptation of sensory systems to the level of stimuli results in an emphasis on changes in the environment. This feature of our sensory systems could be beneficial in helping us learn causal relations in time series contexts, which requires analyzing Δ scores to detect true causal relations. In Experiment 2 we test if the effect of transitions on causal strength judgments is larger for perceptual stimuli than for stimuli presented numerically. Such a finding would suggest that humans are better at uncovering causal relations in more naturalistic settings.

Method

Subjects 100 subjects were recruited using Amazon Mechanical Turk (MTurk) and paid \$0.60. The experiment lasted about 5 minutes.

Design, stimuli and procedure The same method as Experiment 1 was used, except a between-subjects factor was added so that half the subjects viewed the data presented in a visual format (identical to Experiment 1 and shown in Figure 3), while half the subjects viewed the data with X and Y presented numerically.

Results

Replicating Experiment 1, there was a main effect of r_{States}, $F(1, 582) = 44.33$, $p < .001$, as well as transitions $F(2, 582) = 19.82$, $p < .001$. The main question was whether the influence of transitions was larger in the visual than numerical condition. Though the effect of transitions was obtained in both visual ($p < .001$) and numerical ($p < .001$) presentation formats separately, the effect of transitions is more dramatic (steeper slopes of the lines in Figure 4b) in the visual condition; $F(2, 582) = 8.10$, $p < .001$, $\eta_p^2 = .03$. Presenting stimuli in a numerical format attenuated the effect of transitions. When stimuli are presented in a naturalistic format (as in most causal learning contexts involving real-world stimuli), people more naturally attend to transitions, which helps uncover the true causal strength.

Experiment 3

Experiments 1 and 2 showed that people use transitions for inferring causal strength above and beyond states (r_{States}). However, computing a correlation on Δ scores is still a computational challenge. In the introduction we proposed that perhaps instead of using the continuous Δ scores ($\Delta_{Continuous}$), that people instead code changes as increasing or decreasing (Δ_{Binary}) (Table 1), which could simplify the memory and inference process. The stimuli in Experiments 1 and 2 conflated the two; e.g. in the negative transitions condition, $r_{\Delta Cont} \approx -1$, and $r_{\Delta Binary} = -1$ (because every transition was negative). In Experiment 3, we created data sets in which $r_{\Delta Cont}$ and $r_{\Delta Binary}$ diverged to test whether people discretize Δ scores for assessing causal strength.

Method

Subjects 50 subjects were recruited using Amazon Mechanical Turk (MTurk) and paid \$1.50. The experiment lasted about 10-12 minutes.

Stimuli and Design We created datasets that hold r_{States} constant, and vary $r_{\Delta Binary}$ and $r_{\Delta Continuous}$. This is challenging because $r_{\Delta Binary}$ and $r_{\Delta Continuous}$ in a given data set are typically highly similar. This was accomplished in the following way:

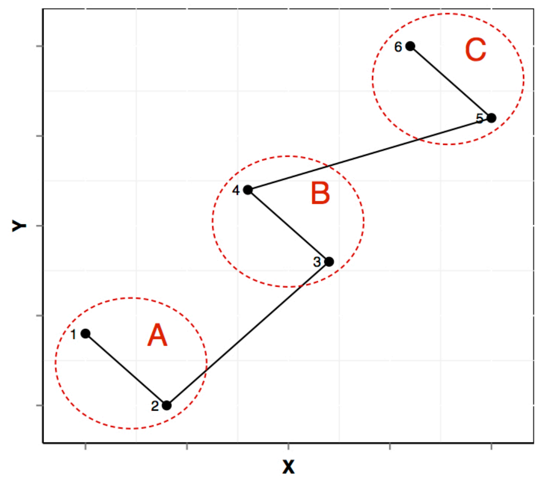

Figure 5: Sample observations from Experiment 3. Numbers reflect the order of the observations.

Each dataset had 12 observations – two observations of each of the following six joint states of (X, Y): (0, 20), (20, 0), (40, 60), (60, 40), (80, 100) and (100, 80). The r_{States} correlation was .83. The observations formed three clusters (A-C, see Figure 5) with low, medium, or high values. Transitions between two observations within a cluster necessarily involved a negative transition (X increasing and Y decreasing, or vice versa). Transitions across clusters necessarily involved a positive transition. Varying the ratio of within to between-cluster transitions changes both $r_{\Delta Cont}$ and $r_{\Delta Binary}$.

However, for a given ratio (and fixed $r_{\Delta Binary}$), $r_{\Delta Cont}$ is also influenced by the between-cluster transition path. Transition paths with large jumps between clusters (e.g. A, C, A…) result in higher $r_{\Delta Cont}$ values, whereas transition paths with small jumps between clusters (e.g. A, B, C…) result in lower $r_{\Delta Cont}$ values.

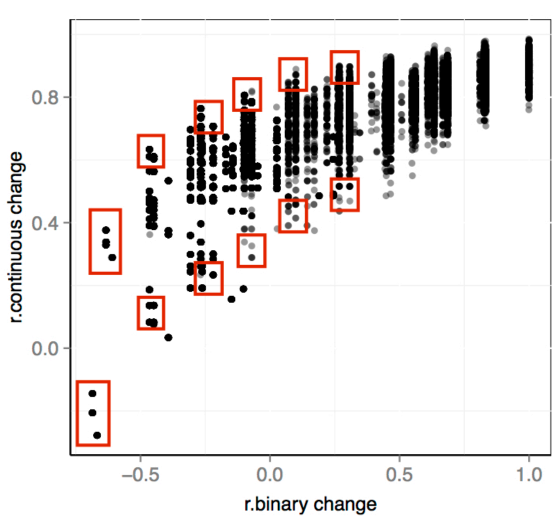

Figure 6: $r_{\Delta\mathrm{Cont}}$ and $r_{\Delta\mathrm{binary}}$ for generated data sets. Sampled regions of data sets have been marked with squares.

Ten thousand data sets were generated by randomly ordering the 12 observations, which produced data with varying degrees of $r_{\Delta Cont}$ and $r_{\Delta Binary}$ (Figure 6). In order to maximally discriminate $r_{\Delta Cont}$ and $r_{\Delta Binary}$, we selected data sets from 12 regions on the periphery of the two-dimensional space (Figure 6). Each subject worked with 12 data sets – one from each region. $r_{\Delta binary}$ values ranged from –.69 to .31, and $r_{\Delta Cont}$ values ranged from –.21 to .90.

The 12 data sets were presented in a random order. Randomly, half the time, the values of X were flipped around the midpoint so that r_{States} was either .83 or –.83; judgments from the flipped conditions were reverse-coded for analysis.

Procedure The procedure was the same as Experiment 1 (the data was only presented in the visual format), except that each data set had only 12 observations.

Results

The exact $r_{\Delta Cont}$ and $r_{\Delta Binary}$ values of a particular data set were used as predictors in a regression of subjects' final judgments of causal strength. The regressions had a by-subject random intercept for repeated measures, and by-subject random slopes for $r_{\Delta Cont}$ and $r_{\Delta Binary}$ to capture the possibility that some subjects might use $r_{\Delta Cont}$ or $r_{\Delta Binary}$ more strongly or weakly than other subjects.

A bivariate analysis found that $r_{\Delta Binary}$ was a significant predictor of the causal strength judgments (B = 2.42 and SE = .48, p < .001, R^2 = .037). However, somewhat surprisingly, $r_{\Delta Cont}$ was not a significant predictor even in a bivariate analysis (B = .76 and SE = .46, p = .10, R^2 = .006). The reason for the smaller effect size here compared to Experiment 1 is that the $r_{\Delta\mathrm{Cont}}$ values here were less extreme. A multivariate analysis again found that $r_{\Delta Cont}$ was not significant (B = –.76 and SE = .57, p = .19, ΔR^2 = .002), and $r_{\Delta Binary}$ remained significant (B = 2.87 and SE = .59, p < .001, ΔR^2 = .032). $r_{\Delta Binary}$ is a significant predictor over and above $r_{\Delta Cont}$.

Overall, Experiment 3 suggests that people do discretize transitions as "positive" or "negative" for the purpose of estimating causal strength.

General Discussion

Past research on causal strength learning has focused primarily on binary variables, and variables that do not exhibit temporal trends (e.g., increasing or decreasing). The present research was concerned with how people judge causal strength from a continuous cause and effect when they are observed over time. We find that in longitudinal situations how the cause and effect change over time is crucial when judging the strength of the causal relationship.

In Experiment 1, we presented subjects with data sets with constant *state-based* information (r_{States}). By manipulating the order of observations to create all positive or all negative transitions (varying $r_{\Delta Cont}$), we created stimuli with state and transition-based information leading to diverging conclusions about causal strength. In conditions where they were in conflict, transition-based information overcame (or at least neutralized) the effect of the state-based information.

Experiment 2 demonstrated that people attend more to transition-based information in causal learning contexts involving naturalistic visual stimuli rather than numerical stimuli. This finding suggests that our sensory systems, which are attuned to changes in the environment through the process of adaptation, may naturally code information in a way that helps us uncover causal relations in time series contexts. In this case, the visual system is involved.

Finally, Experiment 3 showed that people use the direction of change in a variable's state (increase vs. decrease) more than the magnitude of change for estimating causal strength. This result suggests that people discretize continuous changes into binary changes to determine if transitions are positive or negative, which may be a heuristic for calculating causal strength.

Soo & Rottman (2014) showed that people use transitions for inferring causal strength from binary variables. The present study generalizes this finding to the more complex (and one might argue, more real-world) problem of judging causal strength from continuous variables. As demonstrated in Figure 1, using transitions for estimating causal strength in a time series context is especially important when the cause and effect are continuous variables; only using the overall correlation between the states can result in concluding that there is a positive causal relation when it is in fact negative, or vice versa.

One important question to be answered is whether people use transitions for inferring causal strength even in situations when the order is statistically irrelevant – e.g. when each observation comes from different entities rather than consecutive observations of the same entity, or when there are no temporal trends. For example, in the random conditions in Experiments 1 and 2, if r_{states} was positive (negative), the transitions were mainly positive (negative). The way we tested for transitions involved introducing temporal trends, but this methodology cannot provide insight into situations when there are no temporal trends. Other important questions include understanding individual differences in the role of memory and attentional processes for keeping track of transitions, especially if observations are spaced out.

Overall, these experiments provide a positive view of human causal learning in longitudinal contexts; people are able to uncover causal relations despite complex temporal trends. Furthermore, this process appears to be fairly easy, and is facilitated by naturalistic presentations. It is possible that people also use changes over time for other sorts of learning processes such as learning correlations (as opposed to causal relations), or learning to discriminate between categories.

Acknowledgements

This research was supported by NSF 1430439.

References

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–84.

Hagmayer, Y., & Meder, B. (2013). Repeated causal decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 33–50.

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.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 451–60.

Lane, D. M., Anderson, C. A., & Kellam, K. L. (1985). Judging the relatedness of variables: The psychophysics of covariation detection. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 640–649.

Marsh, J. K., & Ahn, W.-K. (2009). Spontaneous assimilation of continuous values and temporal information in causal induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 334–52.

Mooij, J., Janzing, D., Heskes, T., & Schölkopf, B. (2011). On Causal Discovery with Cyclic Additive Noise Models. *Advances in Neural Information Processing Systems*, *24*, 1–9.

Pacer, M. D., & Griffiths, T. L. (2011). A rational model of causal induction with continuous causes. *Advances in Neural Information Processing Systems*, *24*.

Rottman, B. M. (in press). Searching for the Best Cause: Roles of Mechanism Beliefs, Autocorrelation, and Exploitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Saito, M. (2015). How People Estimate Effect Sizes : The Role of Means and Standard Deviations. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 2075–2079).

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal Contiguity and the Judgement of Causality by Human Subjects. *The Quarterly Journal of Experimental Psychology*, *41B*(2), 139–159.

Soo, K., & Rottman, B. M. (2015) Elemental Causal Learning from Transitions. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (2254-2259).

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute Identification by Relative Judgment. *Psychological Review*, *112*(4), 881–911.