

Causal learning about tolerance and sensitization

BENJAMIN MARGOLIN ROTTMAN AND WOO-KYOUNG AHN
Yale University, New Haven, Connecticut

We introduce two abstract, causal schemata used during causal learning. (1) *Tolerance* is when an effect diminishes over time, as an entity is repeatedly exposed to the cause (e.g., a person becoming tolerant to caffeine). (2) *Sensitization* is when an effect intensifies over time, as an entity is repeatedly exposed to the cause (e.g., an antidepressant becoming more effective through repeated use). In Experiment 1, participants observed either of these cause-effect data patterns unfolding over time and exhibiting the tolerance or sensitization schemata. Participants inferred stronger causal efficacy and made more confident and more extreme predictions about novel cases than in a condition with the same data appearing in a random order over time. In Experiment 2, the same tolerance/sensitization scenarios occurred either within one entity or across many entities. In the many-entity conditions, when the schemata were violated, participants made much weaker inferences. Implications for causal learning are discussed.

Whereas most causal learning research has focused on bottom-up inferences of causal relations from data (e.g., Cheng, 1997; Dickinson, Shanks, & Evenden, 1984; Jenkins & Ward, 1965), some has also examined how abstract, domain-general knowledge can shape causal inference. A classic example is Kelley's (1972) "causal schemata," which people can use to make inferences from limited data. For example, if a person believes that his or her stomach-ache could be caused by eating spoiled shrimp or catching a stomach virus (a *multiple sufficient causes* schema) and subsequently learns that the shrimp were, indeed, spoiled, he or she might *discount* the possibility of also having a stomach virus. More recently, Waldmann and colleagues (for a review, see Waldmann, 1996) have shown that causal learning is influenced by knowledge about causal structures. For example, when estimating causal strength, people account for other potential causes of one effect, but not for other effects of one cause (Waldmann & Hagmayer, 2001).

Following Kelley's (1972) suggestion to uncover more causal schemata, the present article introduces two new schemata. The first schema is *tolerance* (see Figure 1). For instance, the first time a person drinks a cup of coffee, he or she may feel very awake; but after repeatedly drinking one cup of coffee, he or she becomes tolerant. The person may then drink two cups of coffee and initially feel very alert; but after repeatedly drinking two cups of coffee, he or she again becomes tolerant.

The second schema is *sensitization* (see Figure 1). For example, two antidepressant pills may initially have no effect; but after repeated exposure, two pills may be sufficient to make a person very happy. If the person cuts down to one pill, the decrease may initially result in lowered happiness;

but as the person becomes sensitized to the reduced amount of antidepressant, one pill may become sufficient.

Note that, for the tolerance/sensitization scenarios in Figure 1, the strength of the simple correlations between causes and effects is 0. Yet, one may still conclude that coffee causes wakefulness and antidepressants cause happiness. Indeed, in real life, people have experienced and can easily understand tolerance/sensitization situations: A drug addict may become so habituated that he consumes doses sufficient to kill an unhabituated person (Carpenter, 1855, as cited in Goudie & Emmett-Oglesby, 1989), and people may not find such an event to be impossible. Many diverse phenomena exhibit tolerance/sensitization patterns (Peeke & Petrinovich, 1984), but no prior studies have examined whether people use tolerance/sensitization schemata when learning novel causal relations.

There are two key attributes of the tolerance/sensitization schemata that inform our experiments. First, the tolerance/sensitization schemata have signature cause-effect, temporal patterns. When the cause is held fixed at the same level across repeated exposures, tolerance situations have a decreasing effect, and sensitization situations have an increasing effect. The scenarios in Figure 1 show a linear increase/decrease, but other instantiations could include nonlinear patterns.

A simple version of tolerance/sensitization can be described by a linear regression equation. For instance, with high = 4 and low = 1, the equation for the tolerance scenario in Figure 1 is $Wakefulness = 1 + 4 \cdot CupsOfCoffee - Time$, and the regression equation for the sensitization scenario in Figure 1 is $Happiness = -8 + 4 \cdot PillsOfAntidepressant + Time$.

B. M. Rottman, benjamin.rottman@yale.edu; W. Ahn, woo-kyoung.ahn@yale.edu

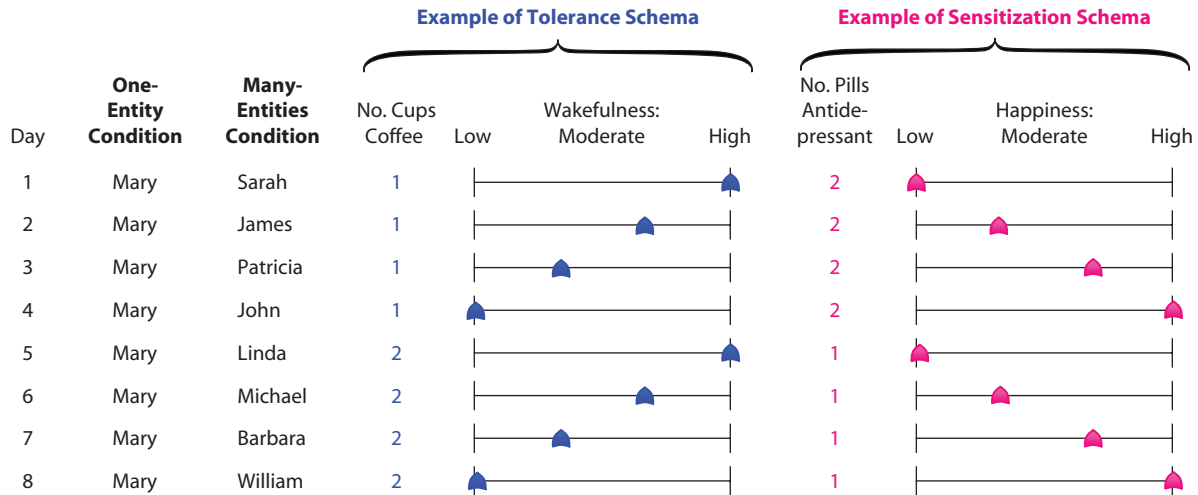


Figure 1. Examples of tolerance and sensitization schemata.

The positive/negative weights on time for sensitization/tolerance represent the increasing/decreasing effects, respectively. Experiment 1 examines the importance of the signature tolerance/sensitization temporal patterns in causal learning.

Second, the tolerance/sensitization schemata apply to repeated treatments to a single entity, not multiple entities: Something within the entity builds up with repeated exposure to the same cause, making that specific entity sensitive or tolerant to the cause over time. Thus, it would be strange if Sarah’s drinking coffee on Day 1 makes James tolerant to coffee on Day 2 (see the relation between the “Many-Entities Condition” and “Examples of Tolerance Schema” columns in Figure 1), because what happens to Sarah is unlikely to influence James’s reaction to coffee. Restated, the trials are not independent: The effect at one time influences the effect at a future time. Although it makes sense for consecutive trials to be dependent with one entity, it is less plausible that trials would be dependent across many entities.¹

Traditionally, contingency-based models (e.g., Cheng, 1997; Jenkins & Ward, 1965) were designed to describe situations with many independent entities (e.g., 50 patients receive a medicine and 50 do not), whereas animal learning models (e.g., Rescorla & Wagner, 1972) were designed to describe situations with one entity over time (e.g., a single rat repeatedly receiving or not receiving a shock). These two sets of models have been routinely compared without considering this important difference. Experiment 2 examines whether people distinguish between the one-entity and many-entities conditions when presented with sensitization and tolerance scenarios (for a demonstration of similar distinctions in categorization, see Barsalou, Huttenlocher, & Lamberts, 1998).

EXPERIMENT 1

Participants observed the signature increasing/decreasing sensitization/tolerance patterns in the same entity

over 14 trials (see the “Tolerance Ordered” column in Figure 2). Then two measures were taken to demonstrate the use of the schemata. First, participants considered another case in which the cause was repeatedly present, and participants predicted the outcome. If they use the schemata, people would predict increasing effects over time for the sensitization scenario, whereas they would predict decreasing effects over time for the tolerance scenario. Second, participants rated their confidence in causal efficacy of the observed cause. There was no correlation between the cause and the effect in our stimuli if the temporal dimension is not taken into consideration. Yet, if people use the schemata and take time into consideration, they would be fairly confident that the cause is efficacious.²

Method

Participants. Sixteen undergraduates from Yale University completed the study for payment at \$10 per hour.

Stimuli and Procedure. In order to demonstrate that tolerance/sensitization schemata are abstract—see, for example, Kelley’s (1972) causal schemas—we created novel scenarios. Participants first read a cover story about a company testing the emissions of machines (noise, light, heat, and smell) and saw one of four scenarios (see below for details).

Each scenario began with the following instructions: “Below you will see the position of the lever (1 = *weak* to 3 = *strong*) and the amount of [X] emitted by machine [Y] over 14 consecutive trials,” where X is the type of emission (e.g., light) and Y is the name of a particular machine (e.g., WDU913). Then, participants saw 14 rows of data (see, e.g., the “Tolerance Ordered” column in Figure 2). Each row displayed the trial number, the position of the lever, and the amount of emission on a scale unique to the type of emission. The 14 rows appeared on the screen in a sequential order, at a slower pace at the beginning (4-sec pause between trials) and at a faster pace as participants got used to the new scenario. After each row appeared, it remained visible until participants answered all of the questions about that scenario.

After all the data for a given scenario were presented, participants answered two prediction questions involving a new machine. (1) Trial 4 prediction: “Suppose a new machine [X], is tried on **position 2 for 11 consecutive trials**. Please rate how much [Y] it would emit on the **4th trial at position 2,**” where X is the name of a machine (e.g., SPE928) and Y is the type of emission (e.g., light). (2) Trial 11

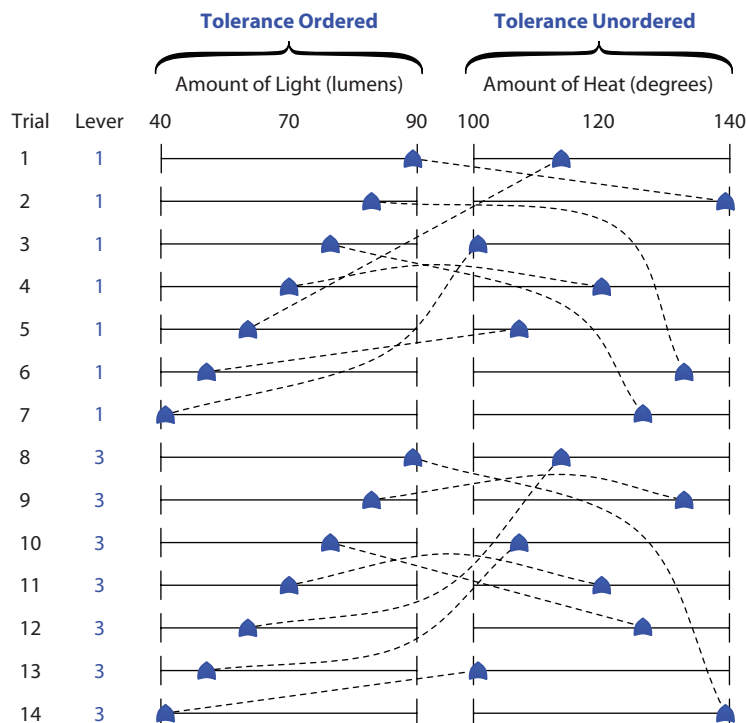


Figure 2. Summary of tolerance-ordered and tolerance-unordered conditions in Experiment 1. Dashed lines (not shown to participants) demonstrate how the same data were reordered from the ordered to the unordered conditions.

prediction: “Please rate how much [Y] machine [X] would emit on the **11th trial** at **position 2.**” The scales of these prediction ratings were the same scales used to display the emission scores and were later converted to a 0–100 scale for data analysis. Immediately after each prediction question, participants rated their confidence about each of the prediction judgments. Finally, participants rated “how confident you are that the lever has the capacity to affect the amount of [emission; e.g., noise].”³ For all of these confidence judgments, participants used a sliding scale from *not confident* to *very confident*, which was later converted to a 0–100 scale for data analysis.

The experiment was a 2 (tolerance vs. sensitization) \times 2 (ordered vs. unordered) within-subjects design. Each participant saw the four conditions using different emissions, which were presented to each participant using a Latin square design blocked by ordered/unordered.

The tolerance-ordered condition showed the signature tolerance pattern, as illustrated in Figure 2. For the sensitization-ordered condition, the lever was at Position 3 for Trials 1–7 and at Position 1 for Trials 8–14, and the pattern of emissions was the mirror image of tolerance: The emission increased from the lowest to the highest points on the scale during Trials 1–7 and again during Trials 8–14. We predicted that, for the tolerance-ordered condition, the average Trial 4 prediction would be much higher than the average Trial 11 prediction, whereas the opposite would occur for the sensitization-ordered condition.

The two unordered conditions were created by randomizing the emission scores from the ordered conditions within the same lever positions (see the “Tolerance Unordered” column in Figure 2). For the sensitization-unordered condition, the pattern of emission was the mirror image of the tolerance-unordered condition, and the lever positions were the same as in the sensitization-ordered condition. The unordered conditions did not display the signature temporal patterns; thus, no abstract causal schema could be applied. We ensured that the emissions for Trials 4 and 11 were set in the middle of the scale, as in the ordered conditions, so that all four conditions had identical emission scores on these critical trials. We predicted

that participants in the ordered conditions would make increasing/decreasing predictions for Trials 4 and 11 despite their having observed identical emission scores on these trials. However, since participants in the unordered conditions had no schema to inform their predictions, we figured they would rely on the observed emission scores on Trials 4 and 11, and would thus predict the same outcome on these two trials.

Most important, the unordered conditions served as controls for the causal efficacy judgments. Had the unordered conditions not triggered any abstract schemas, participants would have based their causal efficacy judgments only on the observed correlations between causes and effects. However, had the ordered conditions triggered the tolerance/sensitization schemata and participants accounted for time in the correlations between causes/effects, they would have given much higher confidence ratings in the causal efficacy judgments for the ordered than for the unordered conditions.

Results and Discussion⁴

Prediction judgments. The prediction judgments showed a decreasing pattern for the tolerance-ordered condition and an increasing pattern for the sensitization-ordered condition. Two separate 2 (order: ordered vs. unordered) \times 2 (trial: 4 vs. 11) repeated measures ANOVAs showed significant main effects of order [for tolerance, $F(1,15) = 22.29, p < .01, \eta_p^2 = .60$; for sensitization, $F(1,15) = 26.24, p < .01, \eta_p^2 = .64$] and main effects of trial [for tolerance, $F(1,15) = 5.84, p = .03, \eta_p^2 = .28$; for sensitization, $F(1,15) = 14.78, p < .01, \eta_p^2 = .50$], but they should be interpreted in light of the significant interactions [for tolerance, $F(1,15) = 15.59, p < .01, \eta_p^2 = .51$; for sensitization, $F(1,15) = 13.97, p < .01, \eta_p^2 = .48$]. For the tolerance-ordered condition (see Figure 3A), the

Trial 4 predictions were higher than the Trial 11 predictions [$t(15) = 4.77, p < .01$]. However, for the tolerance-unordered condition, there was no significant difference [$t(15) = 1.83, p = .09$]. For the sensitization-ordered condition (see Figure 3B), the Trial 11 predictions were higher than the Trial 4 predictions [$t(15) = 5.12, p < .01$], but not for the sensitization-unordered condition [$t(15) < 1$].

Confidence of prediction judgments. If a learner can use a schema to explain the data pattern, he or she should be more confident in making predictions. Separate 2 (order: ordered vs. unordered) \times 2 (trial: 4 vs. 11) repeated measures ANOVAs were run for tolerance and sensitization. Participants were more confident making predictions for the ordered ($M = 60, SD = 25$) than for the unordered ($M = 23, SD = 23$) conditions [for tolerance, $F(1,15) = 20.79, p < .01, \eta_p^2 = .58$; for sensitization, $F(1,15) = 18.76, p < .01, \eta_p^2 = .56$]. Participants were also more confident for Trial 4 ($M = 53, SD = 24$) than for Trial 11 ($M = 30, SD = 19$) predictions [for tolerance, $F(1,15) = 11.71, p < .01, \eta_p^2 = .44$; for sensitization, $F(1,15) = 29.76, p < .01, \eta_p^2 = .67$]. The Trial 11 predictions were more novel because participants never saw one machine tested 11 times on the same lever position.⁵

Confidence of causal efficacy judgments. For the ordered scenarios, participants could attribute the increasing/decreasing patterns to sensitization/tolerance and judge the lever to be causally efficacious. However, for the unordered scenarios, it is unlikely that there was a plausible schema to account for the pattern. Although both groups of participants saw identical cause and effect data not accounting for time, participants gave higher causal efficacy judgments in the ordered than in the unordered conditions, both for tolerance [$z(N = 16) = 2.27, p = .02$] (Figure 3C) and sensitization [$z(N = 16) = 2.48, p = .01$] (Figure 3D).⁶

EXPERIMENT 2

Experiment 1 demonstrated that people apply the sensitization/tolerance schemata when data exhibit the signature increasing/decreasing temporal patterns. The purpose of Experiment 2 was to demonstrate that people apply the tolerance/sensitization schemata more for single entities than for multiple entities. If people believe that a causal mechanism does not transfer across entities, they would likely ignore or discount the temporal information. Consequently, we predicted that participants would give stronger increasing/decreasing predictions, would be more confident in these predictions, and would believe the cause to be more efficacious for one entity than for many entities.

Method

Twenty participants were tested. The experiment was a 2 (tolerance vs. sensitization) \times 2 (one entity vs. many entities) within-subjects design. The two one-entity conditions were identical to the ordered conditions in Experiment 1, except that the scenarios had 10 trials instead of 14, which resulted in somewhat different emission scores, as shown in Figure 4. The many-entity conditions were identical to the one-entity conditions, except that they involved different types of emissions, and each machine was tested only once.

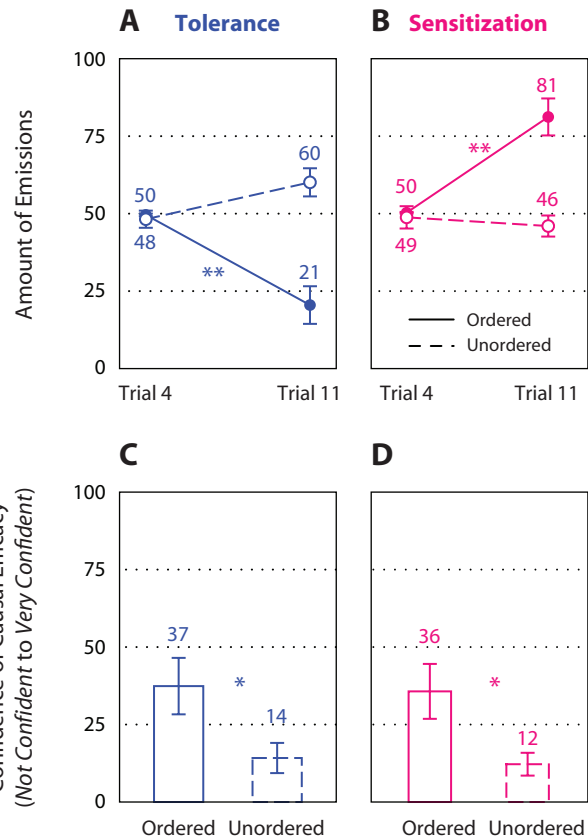


Figure 3. Prediction and causal efficacy judgments for Experiment 1. Bars represent standard errors. * $p < .05$. ** $p < .01$.

Figure 4 summarizes the four conditions. For the one-machine conditions, the name of one machine appeared on all 10 rows. For the many-machine conditions, 10 different machine names appeared, one on each row. The sensitization/tolerance conditions showed the lever and emission patterns, as presented under the two respective brackets in Figure 4.

Participants made predictions for new machines on Trials 3 and 8. For the one-entity conditions, the prediction judgments were identical to those in Experiment 1, except the predictions were for Trials 3 and 8 instead of 4 and 11. For the many-entities conditions, the predictions were framed as occurring to different machines. For the Trial 3 prediction, participants were given the following instructions: "Suppose that **8 new machines** are tried on **position 2 for the first time each**. Please rate how much [X] the **3rd** machine, [Y1], would emit at **position 2**," where X is the type of emission (e.g., heat) and Y1 is the name of a machine (e.g., MWY266). For the Trial 8 prediction, participants were instructed as follows: "Please rate how much [X] the **8th** machine, [Y2], would emit at **position 2**," where Y2 is the name of another machine (e.g., PAQ715).

As in Experiment 1, participants also rated their confidence about each of the prediction judgments and rated the causal efficacy of the lever in each condition. Since the tolerance/sensitization schemata explained lever/emission patterns only in the one-entity conditions, participants should have been more confident in making predictions and should have given higher causal efficacy ratings for the one-entity than for the many-entities conditions.

Results and Discussion

Prediction judgments. The prediction judgments suggest that participants used the tolerance/sensitization sche-

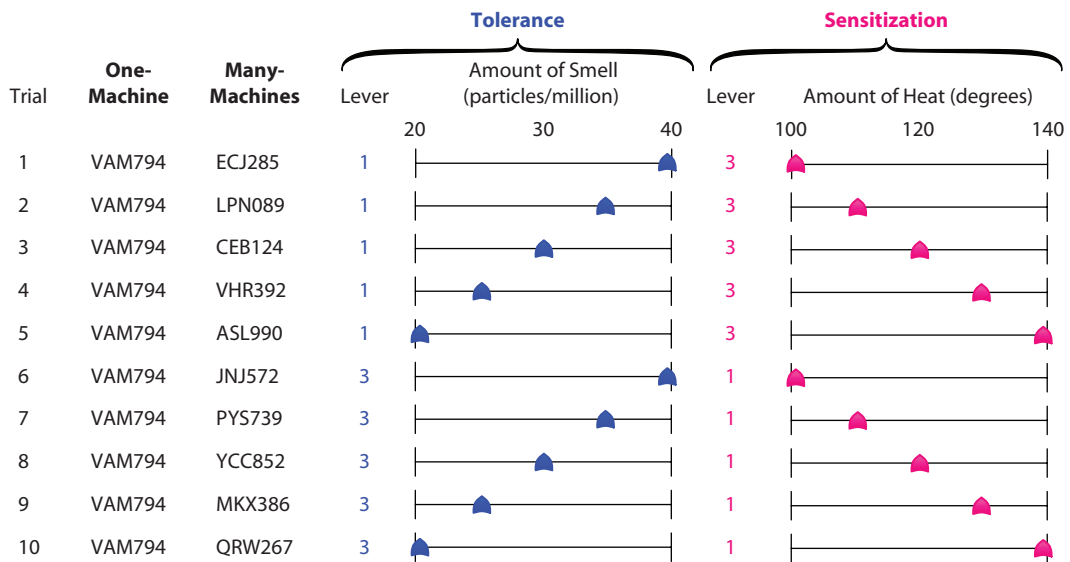


Figure 4. Summary of four conditions in Experiment 2.

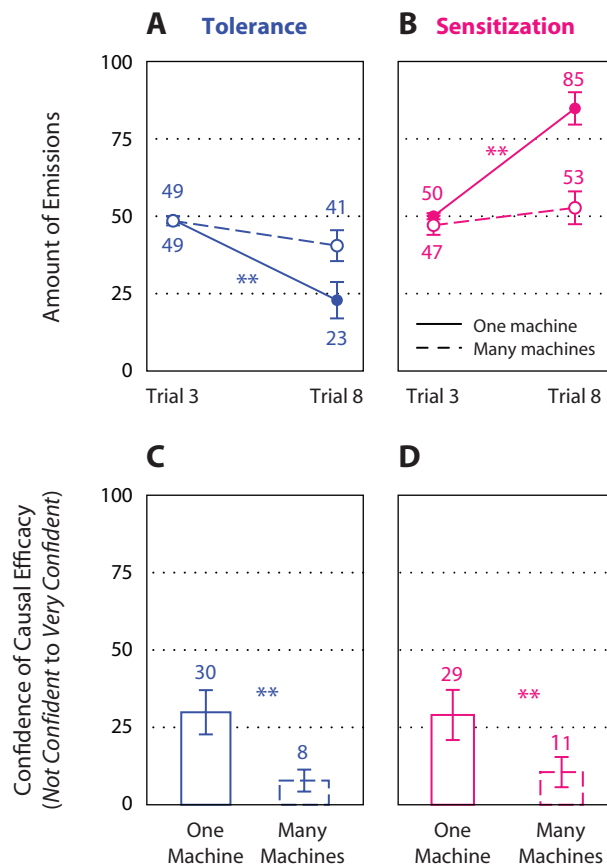


Figure 5. Prediction and causal efficacy judgments for Experiment 2. Bars represent standard errors. ** $p < .01$.

mata more frequently for one-entity than for many-entities conditions. Two separate 2 (entity: one vs. many) \times 2 (trial: 3 vs. 8) repeated measures ANOVAs showed significant main effects of entity [for tolerance, $F(1,19) = 8.01, p = .01, \eta_p^2 = .30$; for sensitization, $F(1,19) = 33.53, p \leq .01, \eta_p^2 = .64$] and significant main effects of trial [for tolerance, $F(1,19) = 14.80, p < .01, \eta_p^2 = .44$; for sensitization, $F(1,19) = 21.52, p < .01, \eta_p^2 = .53$], but they should be interpreted in light of the significant interactions [for tolerance, $F(1,19) = 5.89, p = .03, \eta_p^2 = .24$; for sensitization, $F(1,19) = 14.29, p < .01, \eta_p^2 = .43$]. In the tolerance–one-entity condition (Figure 5A), participants predicted a considerable decrease in emission from Trial 3 to Trial 8 [$t(19) = 4.02, p < .01$]. However, in the tolerance–many-entities condition, there was no significant difference [$t(19) = 1.61, p = .12$]. For sensitization (see Figure 5B), the predictions increased considerably from Trial 3 to Trial 8 for one entity [$t(19) = 6.66, p < .01$], but not for many entities [$t(19) < 1$].

Confidence of prediction judgments. Separate 2 (entity: one vs. many) \times 2 (trial: 3 vs. 8) repeated measures ANOVAs were run for tolerance and sensitization. In both, participants were more confident making predictions for one-entity ($M = 57, SD = 29$) than for many-entities ($M = 28, SD = 25$) conditions [for tolerance, $F(1,19) = 7.75, p = .01, \eta_p^2 = .29$; for sensitization, $F(1,19) = 23.57, p < .01, \eta_p^2 = .55$], suggesting that they had applied the tolerance/sensitization schemata more in the one-entity conditions. As in Experiment 1, since Trial 8 predictions were more novel, participants tended to be more confident for Trial 3 ($M = 47, SD = 22$) than for Trial 8 ($M = 38, SD = 23$) [for tolerance, $F(1,19) = 3.69, p = .07, \eta_p^2 = .16$; for sensitization, $F(1,19) = 4.78, p = .04, \eta_p^2 = .20$].

Confidence of causal efficacy judgments. Participants were more confident in the causal efficacy of the lever in the one-entity than in the many-entities conditions for both tolerance [$z(N = 20) = 2.73, p < .01$] and sensitization [$z(N = 20) = 2.92, p < .01$] (Figure 5C). These results suggest that participants more frequently use the tolerance/sensitization schemata to make sense of the data patterns in the one-entity conditions than in the many-entities conditions.

GENERAL DISCUSSION

In two experiments, use of the tolerance/sensitization schemata influenced causal learning. Participants more frequently applied the schemata when a scenario depicted the signature tolerance/sensitization patterns between causes and effects, rather than when the same data were randomly ordered, as well as when a scenario depicted one entity over time, rather than many entities. When they applied the schemata, participants made decreasing/increasing predictions, were more confident in these predictions, and were more confident in the efficacy of causes.

In many causal learning scenarios, it is not known whether data are independent over time, and it is important to make this determination in order to draw appropriate conclusions. The present results suggest that, when the

scenario involves many entities, people are more likely to assume that the data are independent; consequently, those people ignore time. However, when the scenario occurs within one entity, people are more likely to infer that the trials are dependent and are, therefore, more likely to account for time when estimating the causal efficacy of the primary cause.

Models of Causal Learning

Although the temporal order of trials is crucial in the tolerance/sensitization schemata—and our participants were indeed sensitive to trial order—many models of causal learning (e.g., Cheng, 1997; Jenkins & Ward, 1965) simply aggregate over all observations, irrespective of trial order (for similar discussions, see Dennis & Ahn, 2001). The Rescorla–Wagner (1972) model is sensitive to trial order in that it continually minimizes the error between current predictions and outcomes. However, it cannot capture our participants' prediction judgments, since its prediction for all future trials is based on the current associative strength (e.g., after Trial 3, Rescorla–Wagner would make the same prediction for Trial 4 and Trial 11). Furthermore, since the effect in sensitization/tolerance scenarios is constantly increasing/decreasing, the associative strength would not converge on one value that could be used as a causal efficacy judgment.

Furthermore, it is unclear how many models could be modified to account for time. First, many influential models handle only binary-valued causes and effects, but the temporal variable is necessarily multivalued (but see Griffiths & Tenenbaum, 2005). Additionally, although methods (e.g., Cheng, 1997; Cheng & Novick, 1992; Spellman, 1996) have been proposed for how people condition on alternative binary causes, conditioning on a multivalued variable, such as time, is more complicated.⁷ Nonetheless, our participants did account for time when estimating the causal efficacy of the primary cause.

Future Directions

Much of the previous work on causal learning has focused on how people learn causal strength from a binary correlation. The present experiments expand the types of scenarios studied by demonstrating that people are sensitive to the temporal order of trials. Future models of causal learning may look toward animal habituation learning models (e.g., Wagner, 1981) or normative statistical models that do not assume independence of observations over time, such as time series analysis. In addition to tolerance/sensitization, people likely use other causal schemata, such as sinusoidal or alternating patterns (see Rottman, Ahn, & Luhmann, in press); and as Kelley (1972) originally suggested, it would be useful to catalog other schemata.

Another important area for future research is to examine how people acquire domain-general schemata, such as tolerance/sensitization. One viable possibility is that they form a generalized schema by analogizing between specific instances (e.g., tolerance to caffeine and tolerance to distractors while working). Analogizing from an instance to a schema may allow people to make complex predictions on the basis of little data, much in the same

way that complex predictions can be made if a specific mechanism is known. In general, given the importance of causal reasoning, diverse causal schemata deserve more attention from researchers.

AUTHOR NOTE

This research was supported by a National Science Foundation Graduate Research Fellowship to B.M.R. and National Institute of Mental Health Grant MH57737 to W.A. The authors thank Batel Isenstein, Chris Cabrall, and Rachel Litwin for data collection, and Marc Buehner, Josh Tenenbaum, and two anonymous reviewers for useful comments. Correspondence concerning this article should be addressed to B. M. Rottman or W. Ahn, Department of Psychology, Yale University, P.O. Box 208205, New Haven, CT 06520-8205 (e-mail: benjamin.rottman@yale.edu or woo-kyoung.ahn@yale.edu).

REFERENCES

- BARSALOU, L. W., HUTTENLOCHER, J., & LAMBERTS, K. (1998). Basing categorization on individuals and events. *Cognitive Psychology*, **36**, 203-272. doi:10.1006/cogp.1998.0687
- CHENG, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367-405. doi:10.1037/0033-295X.104.2.367
- CHENG, P. W., & NOVICK, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, **99**, 365-382. doi:10.1037/0033-295X.99.2.365
- DENNIS, M. J., & AHN, W. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition*, **29**, 152-164.
- DICKINSON, A., SHANKS, D., & EVENDEN, J. (1984). Judgment of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, **36A**, 29-50.
- GOUDIE, A. J., & EMMETT-OGLESBY, M. W. (EDS.) (1989). *Psychoactive drugs: Tolerance and sensitization*. Clifton, NJ: Humana Press.
- GRIFFITHS, T. L., & TENENBAUM, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, **51**, 334-384. doi:10.1016/j.cogpsych.2005.05.004
- JENKINS, H. M., & WARD, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General & Applied*, **79**(Whole No. 594).
- 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.
- LUHMANN, C. C., & AHN, W. (2005). The meaning and computation of causal power: Comment on Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, **112**, 685-693. doi:10.1037/0033-295X.112.3.685
- PEEKE, H. V. S., & PETRINOVICH, L. (EDS.) (1984). *Habituation, sensitization, and behavior*. Orlando, FL: Academic Press.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- ROTTMAN, B. M., AHN, W., & LUHMANN, C. C. (in press). When and how do people reason about unobserved causes? In P. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences*. New York: Oxford University Press.
- SPELLMAN, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, **7**, 337-342. doi:10.1111/j.1467-9280.1996.tb00385.x
- WAGNER, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.
- WALDMANN, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 47-88). San Diego: Academic Press.
- WALDMANN, M. R., & HAGMAYER, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition*, **82**, 27-58. doi:10.1016/S0010-0277(01)00141-X

NOTES

1. Tolerance/sensitization may occur across many entities in certain domains that have interentity mechanisms (e.g., social situations).
2. For instance, the regression weights for the causes in the linear regression models presented in our introductory remarks are positive when one accounts for time.
3. We believe that having causal efficacy is an all-or-none condition (Luhmann & Ahn, 2005): A drug either does or does not have the capacity to increase memory. To make this judgment more sensitive, we asked participants to rate confidence in causal efficacy.
4. To ensure that the findings would not be due entirely to order effects in the four conditions, we also examined results from only the first half of the conditions presented to each participant. The patterns of results hold for both Experiments 1 and 2.
5. The only other significant finding was an interaction between order and trial for tolerance [$F(1,15)=6.97, p=.02, \eta_p^2=.32$]. There was a larger difference in confidence between Trial 4 and Trial 11 for the ordered than for the unordered conditions.
6. Nonparametric Wilcoxon signed-rank tests were used in both experiments because the causal efficacy judgments were not normally distributed; the distributions were positively skewed because many of the responses were near *not confident*.
7. A learner can condition on a binary variable by looking only within trials when the conditioned variable is held fixed at one state (present/absent). For a multivalued variable, however, there are potentially infinite states on which to condition. Furthermore, in the present scenarios, only one trial occurs at a particular time. Thus, it is not possible to determine whether the cause has an effect within multiple trials when time is held constant; one would have to compare trials close in time.

(Manuscript received January 7, 2009;
revision accepted for publication July 17, 2009.)