

Decision-Making and Biases in Causal-Explanatory Reasoning

Samuel G. B. Johnson¹, Marianna Zhang², & Frank C. Keil¹

(samuel.johnson@yale.edu, mariannaz@uchicago.edu, frank.keil@yale.edu)

¹Department of Psychology, Yale University, 2 Hillhouse Ave., New Haven, CT 06520 USA

²Department of Psychology, University of Chicago, Chicago, IL 60637 USA

Abstract

Decisions often rely on judgments about the probabilities of various explanations. Recent research has uncovered a host of biases that afflict explanatory inference: Would these biases also translate into decision-making? We find that although people show biased inferences when making explanatory judgments in decision-relevant contexts (Exp. 1A), these biases are attenuated or eliminated when the choice context is highlighted by introducing an economic framing (price information; Exp. 1B–1D). However, biased inferences can be “locked in” to subsequent decisions when the judgment and decision are separated in time (Exp. 2). Together, these results suggest that decisions can be *more* rational than the corresponding judgments—leading to choices that are rational in the output of the decision process, yet irrational in their incoherence with judgments.

Keywords: Decision-making; causal reasoning; inductive reasoning; explanation; behavioral economics.

Introduction

Our decisions often depend on prior inferences. For instance, a patient decides on a treatment that matches the disease likeliest to ail her, based on diagnostic tests; an investment banker chooses an asset allocation expected to maximize profits, based on past returns; a consumer chooses the toothpaste likeliest to keep his teeth white, based on persuasive advertising.

As others have argued, such inference-based decisions are often *causal* (e.g., Sloman, 2005). The patient’s treatment will *cause* her to recover; the investment banker’s choice will *cause* profits to be maximized; the toothpaste will *cause* the consumer’s teeth to be white. Much is known about how people make causal predictions and evaluate causal explanations (e.g., Rottman & Hastie, 2014; Sloman, 2005; Waldmann & Holyoak, 1992). In particular, recent work has triangulated a set of heuristics used in making diagnostic inferences, including causal explanations (Johnson, Rajeev-Kumar, & Keil, 2014, 2015a; Khemlani, Sussman, & Oppenheimer, 2011; Lombrozo, 2007). How do these mechanisms translate into choice behavior?

One possibility is that we use these mechanisms to arrive at judgments, and then translate those judgments into decisions. (This is roughly the view of classical decision theory; e.g., Jeffrey, 1965.) Although this pathway from judgment to decision is itself normative in preserving coherence, it can lead to errors in decisions to the extent that the judgments are themselves biased. Since judgments arrived at through diagnostic reasoning are subject to systematic biases (e.g., Khemlani et al., 2011; Lombrozo, 2007), one would expect decisions that

depend on those judgments to also be biased.

Alternatively, decision-making may recruit additional mechanisms beyond judgment. In dual process terms, we could think of the intuitive judgments as relying on System 1, and the biases result because they are not corrected by System 2 (Kahneman, 2003). If decision-making recruits additional System 2 resources that are not available in judgment, then the decisions may be less biased, or even unbiased.

Some previous results are consistent with this more nuanced picture of judgment and decision. For example, in addition to judgments leading to our decisions, our decisions also seem to affect our judgments (Johnson, Rajeev-Kumar, & Keil, 2015b). When people choose a course of action that is more consistent with one diagnostic judgment rather than another, people tend to think that the corresponding judgment is more likely to be true—even if the reason for choosing the corresponding action is independent of the judgment (i.e., the stakes are higher for being wrong given the other choice). Likewise, people are more likely to search for information that confirms a decision (Fischer & Greitemeyer, 2010) and judge disconfirmatory evidence more harshly (Chaxel, Russo, & Kerimi, 2013). All of these findings point to a bidirectional relationship between judgment and decision.

Although it is well-known in behavioral economics circles that monetary incentives improve performance in decision-making contexts (e.g., Levitt & List, 2007), it is less clear whether the ‘pseudoincentive’ of a decision-making *task* (with the same monetary compensation as a judgment task) would be sufficient to induce System 2 monitoring.

We examine these issues in two sets of studies. First, we test whether a bias against explanations making unverified predictions propagates from judgment to decision (Exp. 1A and 1B), and test boundary conditions of these effects (Exp. 1C and 1D). Second, we look at individual differences in a context where *both* judgments and decisions are elicited from the same participants and are separated in time (Exp. 2).

Experiments 1A and 1B

When interpreting evidence to distinguish between hypotheses, people are unwilling to settle for ignorance (Khemlani et al., 2011; Sussman et al., 2014).

For example, suppose that you are hunting. There are two types of deer in the forest, one with white spots on its tail (species *W*) and another without spots (species *N*), which roam the forest in equal numbers. Species *W* has a

wide explanatory scope, because it can explain more potential features (i.e. white spots on tail) than species *N*, which has a *narrow* scope.

Suppose that, due to the policies of your local government, the deer have overlapping hunting seasons, but species *W* must be shot with a bow-and-arrow, whereas species *N* must be shot with a gun. Now, suppose you see a deer in the distance, but its tail is occluded by a tree. Do you shoot with a gun or with a bow-and-arrow?

In reasoning through problems like this, people attempt to *infer* whether this particular deer *would* have spots, if the tree were not in the way. Unfortunately, since the forest has equal numbers of *W* and *N* deer, this strategy is not helpful—it has exactly an equal chance of having spots (if it is *W*) or not (if it is *N*). Nonetheless, people do not settle for ignorance, and use the base rate of the diagnostic feature—in this case, the proportion of deer *in general* that have white spots on their tails—to guess whether this *particular* deer will have white spots (Johnson, Rajeev-Kumar, & Keil, 2015a). Since most deer do not have white spots on their tails, people erroneously infer that this deer is also unlikely to have white spots, and will conclude it is more likely to belong to species *N*. Because most effects and features are relatively uncommon in general, people generally are averse to explanations with a wide scope of unverified predictions (see also Johnson, Johnston, Koven, & Keil, 2015, for evidence of this bias in 4-year-old children).

We make inferences in large part so that we can make choices in the world. In this case, the inference ought to influence whether you use a gun or a bow-and-arrow to shoot the deer. More generally, we often make decisions in economic contexts which depend on explanatory inference where evidence is unavailable. Would a bias against wide scope explanations, making unverified predictions, also arise in decision-relevant contexts?

To test this, participants in Exp. 1 read about situations where two different explanations (one wide and one narrow) had different choice implications. For example, suppose you've been having problems with your robotic lawnmower—it has been running into trees and making strange noise. There are two possible problems that could lead to this behavior—it could be a faulty hesolite axle (which makes no other predictions) or a faulty transduction spindle (which also makes the prediction that the spindle should remain cool during use). However, because safety precautions make it impossible to lift the lawnmower's lid, you cannot check whether or not the spindle is cool.

Thus, based on previous research, we would predict that participants should favor the narrow explanation that did *not* make the unverified prediction, even if the two explanations actually have equal posterior probabilities, given the information in the problem. Exp. 1 tested this prediction in two different ways. Some participants were asked to make an explicit causal/explanatory judgment, identifying which part was the most likely cause of the

problem (Exp. 1A). Previous research suggests that participants should favor the narrow explanation here. Other participants were asked to choose which replacement part they would buy (Exp. 1B). If their causal inferences translate directly into decisions, they should also favor the narrow explanation here. Conversely, if the decision-making context leads them to recruit System 2 resources that correct for bias, then they should be more likely to provide normative responses.

Method

We recruited 383 participants from Amazon Mechanical Turk ($N = 186$ for Exp. 1A, $N = 197$ for Exp. 1B); 48 were excluded from analysis due to poor performance on check questions (see below).

Participants in both experiments completed 5 items (concerning robotic lawnmowers, pest control, junk mail, television repair, and household detergents) in a random order. For example, in Exp. 1A, the lawnmower item read:

Imagine your autonomous robotic lawnmower hasn't been working. It's definitely a problem with either the transduction spindle or the hesolite axle. These two problems occur equally often.

A faulty hesolite axle causes disorientation and makes noise.

A faulty transduction spindle causes disorientation, makes noise, and stays cool during use.

Your lawnmower has been running into trees and making strange noise, but you can't tell whether the transduction spindle stays cool during use because the lawnmower's lid cannot be opened during use as a safety precaution.

That is, the *narrow* explanation (*faulty hesolite axle*) makes two confirmed predictions (*disorientation* and *noise*). The other *wide* explanation (*faulty transduction spindle*) makes the same two confirmed predictions, plus one latent or unverified prediction (*stays cool*). The order of the wide and narrow explanations was counterbalanced for each participant.

In Exp. 1A, participants answered a *cause* question, reporting which causal explanation they favored (e.g., "Which part do you think caused the problem?"), on a scale from 0 ("Definitely transduction spindle") to 10 ("Definitely hesolite axle").

The items in Exp. 1B were the same, except they were also given some additional information about the decision-making context, focusing on what interventions they could make to solve the problem. For the lawnmower example, participants read:

To fix it, you must replace one of the parts and check if the lawnmower is fixed. You can buy a new transduction spindle for **\$40** or a new hesolite axle for **\$40**.

They then answered a *choice* question, reporting which choice they would make (e.g., "Which part would you buy?") on a scale from 0 ("Definitely buy transduction

spindle”) to 10 (“Definitely buy hesolite axle”). For both Exp. 1A and 1B, the left/right orientation of the scales was adjusted to match the order in which the explanations were listed, and the default setting on all scales was the midpoint.

At the end of each study, participants were asked to check off of a list of items that had appeared throughout the study, as an attention check. Participants incorrectly answering more than 30% of these questions were excluded from analysis.

Results

All measures were scaled so that negative scores correspond to *narrow* scope inferences or decisions, and so that positive scores correspond to *wide* scope inferences or decisions.

When participants were asked to evaluate explanation in Exp. 1A, they had a significant bias toward the negative latent scope explanation [$M = -0.25$, $SD = 0.84$; $t(167) = -3.87$, $p < .001$, $d = -0.30$]. This bias is consistent with previous work on causal explanation (e.g., Khemlani et al., 2011), where people tended to favor explanations that did not posit unverified predictions.

However, when participants were asked to choose between potential interventions based on explanations, in Exp. 1B, they no longer had any bias [$M = 0.00$, $SD = 1.41$; $t(166) = 0.02$, $p = .98$, $d = 0.00$]. This led to a significant difference between Exp. 1A and 1B [$t(333) = 2.00$, $p = .047$, $d = 0.22$]. See Table 1 for means and confidence intervals across Experiments 1A–D.

Exp.	DV	Prices	Mean	CI
1A	Cause	Yes	-0.25	(-0.38, -0.12)
1B	Choice	Yes	0.00	(-0.21, 0.22)
1C	Cause	No	-0.03	(-0.16, 0.11)
1D	Cause	Cheap	-0.14	(-0.31, 0.03)

Table 1: Results of Experiments 1A–D

Discussion

These results suggest a nuanced role for explanatory inference in decision-making. Exp. 1A demonstrated that a signature bias of explanatory reasoning—found previously in causal diagnosis (Khemlani et al., 2011), categorization (Sussman et al., 2014), stereotyping (Johnson, Kim, & Keil, 2016), and causal strength judgment (Johnson, Johnston, Toig, & Keil, 2014)—also appears in the kinds of causal reasoning problems that feed directly into decision-making.

However, somewhat surprisingly, this bias did not translate into biased decisions in Exp. 1B. Taken together, participants in these experiments indicated that *N* was a more likely cause than *W*, yet they were equally likely to intervene on *N* and *W*. These decisions at once violate and affirm the tenets of rationality: They *violate* rationality in the sense that individuals’ decisions were inconsistent with their beliefs; yet, they *affirm* rationality in the sense

that their decisions were unbiased. This unbiased decision, while inconsistent with their beliefs, is rational taken in isolation.

Something about making an inference-based decision, rather than a mere inference, appears to be pushing people toward more rational behavior. In dual process terms (Kahneman, 2003), one possibility is that explanatory heuristics produce System 1 responses which can be overridden by System 2 monitoring. Perhaps the stakes of decision-making invoke more monitoring of intuitive judgments, leading to more normative responses. Exps. 1C and 1D explore implications of this account.

Experiment 1C

Exp. 1C aimed to pinpoint which difference between Exp. 1A and 1B drove the difference in outcomes. These studies differed in two ways: (1) They used different dependent measures and tasks (a causal diagnosis versus a choice); and (2) They invoked different judgment contexts (a reasoning context versus a choice context) in that Exp. 1B provided information about interventions to fix the problem, such as the prices of the options. Which of these factors led to the biased inferences in Exp. 1A but unbiased decisions in Exp. 1B?

On the one hand, it may be the *task* itself (causal diagnosis versus choice) that is crucial. On the assumption that decision-making invokes more System 2 monitoring than mere inference, it seems plausible that the nature of the question itself is driving the results: Forcing participants to appreciate the stakes of the problem by using a decision process may lead them to more normative responses.

Alternatively, the mere *context* of making an economic decision could suffice to raise the stakes. The contextual information supplied in Exp. 1B indicated that the judgment *implied* a course of action, and perhaps that implication is sufficient even in the absence of an overt decision.

In Exp. 1C, we distinguished between these factors by using the same dependent measure as Exp. 1A (a causal diagnosis) but including the contextual information from Exp. 1B, to establish the decision-making context. If the task itself led to more rational judgment, then we would expect biased judgment in Exp. 1C (as in Exp. 1A); but if the choice context is sufficient to invoke rational judgment, then we would expect unbiased judgment (as in Exp. 1B).

Method

We recruited 206 participants from Amazon Mechanical Turk; 21 were excluded from analysis due to poor performance on check questions.

The procedure was identical to Exp. 1B, including the same paragraph of contextual information (“To fix it...”; see Exp. 1B methods). However, the dependent measure was the same causal question used in Exp. 1A (“Which part do you think caused the problem?”).

Results and Discussion

Participants' judgments were normative, even though these judgments were causal inferences rather than choices. That is, participants' judgments were unbiased [$M = -0.03$, $SD = 0.96$; $t(184) = -0.37$, $p = .71$, $d = -0.03$]. Correspondingly, the causal judgments in Exp. 1C (with the choice context) differed significantly from the causal judgments in Exp. 1A (without the choice context) [$t(351) = 2.34$, $p = .020$, $d = 0.25$] but not from the choices in Exp. 1B [$t(350) = -0.22$, $p = .82$, $d = -0.02$]. See Table 1.

These results suggest that a judgment that *implies* a decision is sufficient to induce System 2 monitoring, just as much as a decision itself. Exp. 1D further probes the boundary conditions of this normative choice behavior.

Experiment 1D

What is it about a choice context that induces System 2 monitoring? It could be that having to make a decision, regardless of the stakes, is sufficient to induce monitoring. Alternatively, it could be that the importance of the choice could be the key factor, in which case the economic stakes of the choice should be critical.

Exp. 1D sought to tease apart these mechanisms by introducing "dirt cheap" prices. If *any* choice is sufficient to induce monitoring, regardless of the stakes, then we should expect unbiased inferences. Conversely, if it is the stakes themselves that are critical, then we should expect the bias to return when they are minimized.

Method

We recruited 198 participants from Amazon Mechanical Turk; 18 were excluded from analysis due to poor performance on check questions.

The procedure was identical to Exp. 1C, except the prices were lowered to "dirt cheap" levels. For example:

To fix it, you must replace one of the parts and check if the lawnmower is fixed. From the local junkyard, you can buy a replacement transduction spindle for **\$0.75** or a replacement hesolite axle for **\$0.75**.

The methods were otherwise identical to Exp. 1B.

Results and Discussion

The results were mixed. On the one hand, participants' causal judgments were somewhat non-normative, leading to a marginally significant bias [$M = -0.14$, $SD = 1.15$; $t(179) = -1.66$, $p = .100$, $d = -0.12$]. However, despite the significant difference between Exp. 1A and 1C, the current bias did not significantly differ from either experiment [$t(346) = 1.00$, $p = .32$, $d = 0.11$ and $t(363) = -1.05$, $p = .29$, $d = -0.11$, respectively]. See Table 1 for means and comparisons across experiments.

These results are not conclusive, but they are suggestive. The marginally significant bias seems to suggest that extremely low stakes allow for some degree of System 1 bias that is uncorrected by System 2 monitoring. However, the results falling midway between Exp. 1A and 1C (albeit not significantly differing from

either) suggests that both mechanisms may be at play in bias reduction: The stakes appear to play a role, but the mere act of implying a choice also appears to play a role.

Experiment 2

We have been describing the theoretical picture supported by these results in dual process terms—that people make intuitive judgments which are then corrected by more explicit reasoning when making decisions. A more radical view of these results is that causal-explanatory reasoning is simply not a force in decision-making, or that decision-making relies on separate reasoning processes, as opposed to the heuristics known to be used in explanatory reasoning (e.g., Johnson, Rajevee-Kumar, & Keil, 2014, 2015a; Lombrozo, 2007). Could this view be right?

Exp. 2 capitalized on the fact that reasoners do not make uniform judgments in the face of explanations varying in scope—indeed, Exp. 1A revealed considerable variability in judgments ($SD = 0.84$) despite the mean favoring the narrow explanation. That is, participants varied greatly in the magnitude and even direction of their bias (see Johnson, Rajevee-Kumar, & Keil, 2014, 2015a for discussion of the mechanisms underlying this bias, which can lead to biases in either direction, depending systematically on individuals' prior beliefs).

In Exp. 2, participants were asked to make a judgment (as in Exp. 1A) followed by a choice (as in Exp. 1B). If the unbiased choices in Exp. 1B occurred because people are relying on a different computational system for choice that circumvents diagnostic judgment heuristics, then individuals who make biased diagnostic inferences in judgment would be unlikely to make the same biased inferences in choice, or should at least be far less biased. Conversely, if the reasoning mechanisms are the same, then once locked into a judgment, a participant would likely make a choice that matches that judgment.

Method

We recruited 299 participants from Amazon Mechanical Turk; 1 was excluded from analysis due to poor performance on check questions.

The procedure combined the dependent measures of Exp. 1A and 1B, in a within-subjects design. Participants were randomly assigned to one of the five vignettes, and completed both the *cause* question (from Exp. 1A) and the *choice* question (from Exp. 1B), in that order.¹ The procedure was otherwise identical to the other experiments.

Results and Discussion

Among the 113 participants who favored the narrow

¹ The reverse order was not used because this order does not test our hypothesis—if participants made the choice first, then the congruence between choice and inference would be explained by our earlier results showing that economic contexts lead to debiasing.

explanation in responding to the cause question [$M = -2.36$, $SD = 1.51$], these participants also tended to *choose* the option corresponding to that diagnosis [$M = -2.09$, $SD = 2.11$]. Likewise, among the 101 participants who favored the wide explanation in responding to the cause question [$M = 2.13$, $SD = 1.42$], these participants also tended to choose the option corresponding to that diagnosis [$M = 1.91$, $SD = 1.91$]. In fact, these choices were just as strong as their initial diagnoses [$t(112) = 1.56$, $p = .12$ and $t(100) = 1.33$, $p = .19$, respectively], indicating little evidence for less biased decisions than judgments, even though regression toward the mean would push judgments toward less bias.

We draw two conclusions from these results. First, even though decision-making leads to error-correction when made in the absence of an explicit judgment, errors can be “locked in” by first making an explicit judgment. That is, participants were no less biased in making decisions than they were in making judgments in this task, where their decisions followed explicit judgments. Second, despite the unbiased choices in previous studies, these results suggest a strong relationship between diagnostic causal reasoning and subsequent decisions that depend on those causal judgments: Analyses of individual participants revealed that those whose causal judgments were biased in one direction tended to likewise make decisions that were biased (just as much) in the same direction. Although choices were unbiased at the aggregate level in previous studies, likely due to adjustments caused by System 2 error-correction, choices are nonetheless strongly associated with their antecedent causal judgments.

This study is subject to the limitation that choices were made immediately after judgments on a very similar scale, which may lead to anchoring and other scale-use issues. Hence, future work should correct this problem by using less alignable scales, or by using an intermittent task to reduce carry-over effects. Nonetheless, the finding that there was no significant regression toward the mean between tasks suggests that anchoring-and-adjustment cannot be a complete explanation for the current results: There could indeed have been anchoring, but there was little or no adjustment.

General Discussion

Decisions are often predicated upon causal judgments. Yet, the heuristic mechanisms underlying causal judgments often lead to biased inferences. Would these biases translate into decision-making?

At least in the case of the bias against explanations making unverified predictions (Khemlani et al., 2011), the answer appears to be ‘no’. Although participants made biased judgments in choice-relevant inference problems (Exp. 1A), these biases were eliminated when making choices based on those inferences (Exp. 1B). These unbiased responses also carried over to causal judgments that were accompanied by information contextualizing the

choice as an economic decision (such as prices; Exp. 1C). The bias appeared to return when the stakes of the choice were greatly lowered (Exp. 1D), although the bias was of a smaller magnitude than it had been when the choice context was omitted altogether.

These results together suggest that choice contexts can attenuate or eliminate diagnostic reasoning biases. This effect is most likely attributable in part to increases in System 2 monitoring when the choice context is made salient (regardless of stakes), and in part due to accentuated monitoring caused by higher stakes.

Nonetheless, these results do not undermine the claim that choices depend on diagnostic reasoning processes. Indeed, Exp. 2 asked participants to make both a judgment and a decision, and found that participants who made biased judgments were also likely to make biased decisions, in the same direction. This finding indicates that participants’ decisions are based on antecedent judgments. In addition, in contrast to Exp. 1, the choices were *just as biased* as the judgments, suggesting that the act of making an inference can “lock in” the relevant decision, when the judgment and decision are separated in time.

These results contribute to debates concerning human rationality. On the one hand, our results affirm mainstream views in behavioral economics, which have a generally low opinion of human decision processes. This is true in two senses in the current work: First, inferences were biased in a decision-relevant context (Exp. 1A), and these biased judgments could be “locked in” to biased decisions when the judgment and decision were separated in time (Exp. 2). Second, when the decision was not preceded by an explicit judgment, the decision was inconsistent with its antecedent judgment, suggesting incoherence in the decision-making process, in violation of traditional normative models (e.g., Jeffrey, 1965).

Nonetheless, these results are hopeful in a different sense, and more friendly to classical views of human decision faculties. Economists are fond of critiquing lab experiments (including many behavioral economics studies) because they often fail to reflect the incentives present in the marketplace which can create pressure for more optimal behaviors (Levitt & List, 2007), especially at the aggregate level of institutions such as the stock market. Thus, the argument goes, suboptimal behavior in lab contexts can give way to more optimal behavior in economic contexts. Our results go a step further: Not only can market mechanisms potentially drive more rational behavior, but the psychological mechanisms underlying choice behavior appear to induce error-monitoring processes that can lead people to behave more rationally, even in the absence of economic incentives.

This work can be expanded upon in several ways. First, it should extend to other reasoning biases. For instance, people are biased in some contexts to favor overly simple explanations (Lombrozo, 2007), and in other contexts to favor overly complex explanations (Johnson, Jin, & Keil,

2014). Further, the size and direction of these biases can be influenced by normatively irrelevant factors. We have collected some preliminary data, suggesting that these biases are attenuated in choice contexts, just as scope biases were shown to be attenuated in the current results. Likewise, people tend to ‘digitize’ their beliefs, holding propositions to be either certainly true or certainly false, rather than coming in degrees (Johnson, Merchant, & Keil, 2015; Murphy & Ross, 1994). Perhaps people would be less likely to digitize in a choice context, leading to more normative (Bayesian) behavior.

Second, more work is needed to uncover the mechanisms underlying the bias attenuation. We have argued that System 2 error-monitoring is the most plausible explanation, but further work should pinpoint the mechanism and pinpoint the boundary conditions. Explicit responses (such as think-aloud protocols) would be one useful source of evidence, as would other manipulations designed to change the stakes and context (building on Exp. 1C and 1D). Such studies could help to pinpoint the conditions under which biases are and are not attenuated, with potential implications for real-world choice behavior and for debates on the limits of human rationality—as well as the limits on those limits.

References

- Chaxel, A., Russo, J.E., & Kerimi, N. (2013). Preference-driven biases in decision makers’ information search and evaluation. *Judgment and Decision Making*, *8*, 561–576.
- Fischer, P., & Greitemeyer, T. (2010). A new look at selective-exposure effects: An integrative model. *Current Directions in Psychological Science*, *19*, 384–389.
- Jeffrey, R.C. (1965). *The logic of decision*. New York, NY: McGraw-Hill.
- Johnson, S.G.B., Jin, A., & Keil, F.C. (2014). Simplicity and goodness-of-fit in explanation: The case of intuitive curve-fitting. *Proceedings of the 36th Conference of the Cognitive Science Society*.
- Johnson, S.G.B., Johnston, A.M., Toig, A.E., & Keil, F.C. (2014). Explanatory scope informs causal strength inferences. *Proceedings of the 36th Conference of the Cognitive Science Society*.
- Johnson, S.G.B., Kim, H.S., & Keil, F.C. (2016). Explanatory biases in social categorization. *Proceedings of the 38th Conference of the Cognitive Science Society*.
- Johnson, S.G.B., Merchant, T., & Keil, F.C. (2015). Predictions from uncertain beliefs. *Proceedings of the 37th Conference of the Cognitive Science Society*.
- Johnson, S.G.B., Rajeev-Kumar, G., & Keil, F.C. (2014). Inferred evidence in latent scope explanations. *Proceedings of the 36th Conference of the Cognitive Science Society*.
- Johnson, S.G.B., Rajeev-Kumar, G., & Keil, F.C. (2015a). *Sense-making under ignorance*. Manuscript under review.
- Johnson, S.G.B., Rajeev-Kumar, G., & Keil, F.C. (2015b). Belief utility as an explanatory virtue. *Proceedings of the 37th Conference of the Cognitive Science Society*.
- Johnston, A.M., Johnson, S.G.B., Koven, M.L., & Keil, F.C. (2015). Probability versus heuristic accounts of explanation in children: Evidence from a latent scope bias. *Proceedings of the 37th Conference of the Cognitive Science Society*.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.
- Khemlani, S.S., Sussman, A.B., & Oppenheimer, D.M. (2011). *Harry Potter* and the sorcerer’s scope: Latent scope biases in explanatory reasoning. *Memory & Cognition*, *39*, 527–35.
- Levitt, S.D., & List, J.A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*, 153–174.
- Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232–57.
- Murphy, G.L., & Ross, B.H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology*, *27*, 148–93.
- Rottman, B.M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*, 109–139.
- Slooman, S. (2005). *Causal models: How people think about the world and its alternatives*. Oxford, UK: Oxford University Press.
- Sussman, A.B., Khemlani, S.S., & Oppenheimer, D.M. (2014). Latent scope bias in categorization. *Journal of Experimental Social Psychology*, *52*, 1–8.
- Waldmann, M.R., & Holyoak, K.J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, *121*, 222–236.