UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Opponent Uses of Simplicity and Complexity in Causal Explanation

Permalink
https://escholarship.org/uc/item/5gw1g9zx

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Authors
Johnson, Samuel G. B.
Valenti, J. J.
Keil, Frank C.

Publication Date
2017

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Opponent Uses of Simplicity and Complexity in Causal Explanation

Samuel G. B. Johnson[1], J. J. Valenti[2], & Frank C. Keil[1]
(sgbjohnson@gmail.com, jvalen23@jhu.edu, frank.keil@yale.edu)
[1]Department of Psychology, Yale University, New Haven, CT 06520 USA
[2]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218 USA

Abstract

People often prefer simpler explanations because they have higher prior probability. However, simpler explanations are not always normatively superior because they often do not fit the data as well as complex explanations. How do people negotiate this trade-off between prior probability (favoring simplicity) and goodness-of-fit (favoring complexity)? Here, we argue that people use *opponent heuristics*—relying on simplicity as a cue to prior probability but complexity as a cue to goodness-of-fit (Study 1). We also examine factors that lead one or the other heuristic to predominate in a given context. Study 2 finds that people have a stronger simplicity preference in deterministic rather than stochastic contexts, while Study 3 finds that people have a stronger simplicity preference for physical rather than social causal systems. Together, we argue that these cues and contextual moderators act as powerful constraints that help to specify otherwise ill-defined hypothesis comparison problems.

Keywords: Causal reasoning; explanation; probabilistic reasoning; heuristics; judgment under uncertainty.

Introduction

The principle of parsimony has a long and venerable pedigree. It has been discussed since at least Aristotle, who wrote in his *Physics* that "nature operates in the shortest way possible," and it has since become one of the core tools in our argumentative arsenal as scientists. Of course, this principle was given its most famous formulation given by William of Occam, who advised against "multiplying entities beyond necessity."

Simplicity is not only a core notion in science and philosophy, but may well be an organizing principle of cognition (Chater & Vitányi, 2003). People prefer simpler causal explanations (Lombrozo, 2007), category assignments (Pothos & Chater, 2002), and perceptual organizations (van der Helm & Leeuwenberg, 1996), and more easily learn simple concepts (Feldman, 2000).

This principle is not arbitrary. Other things equal, simpler explanations are more likely to be true because they have higher prior probability. Consistent with this analysis, Lombrozo (2007) found that people use simplicity as a heuristic for estimating prior probabilities. In her experiments, participants performing simulated medical diagnoses would not accept a complex explanation over a simple one unless the prior probabilities favored the complex explanation by a factor of 4. Further, participants who had a simplicity bias had distorted memories of the disease base rates, recalling the simpler explanations as having had higher prior probabilities than they in fact did. Thus, people's preference for simple explanations, though sometimes stronger than normatively warranted, appears to track the probabilistic logic favoring simpler explanations.

Yet, simplicity has its limits because a simple and a complex explanation do not always fit the data equally well. There is generally a U-shaped curve in how simple an explanation ought to be. Too complex, and the explanation has a lower prior probability and overfits the data; too simple, and it does not account for the nuance of the phenomenon (Forster & Sober, 1994). How, if at all, does cognition perform this trade-off?

We propose that people use *opponent heuristics* to compare a simpler versus a more complex explanation. This view incorporates Lombrozo's (2007) insight that people use simplicity to estimate prior probability—the $P(H_i)$ terms in Bayesian hypothesis comparison—but couples it with the idea that people also use *complexity* to estimate likelihoods—the $P(E|H_i)$ terms that measure the goodness-of-fit of the evidence to the data.

For example, if a patient is sneezing and has a stomach ache, the patient could have a cold. This explanation is simple, but fits the data imperfectly. If we took a random sample of the population, a reasonably large fraction would have a cold at any given time—so this explanation has high prior probability. But among those people who *have* a cold, how many of them would both be sneezing and have a stomach ache? The *facts* here are complex, and this simple explanation does not fit very well.

In contrast, the patient could have both allergies and a stomach virus. This explanation is more complex, but fits the data neatly. In a random sample of the population, a fairly small number would have both allergies and a stomach virus. Yet, many of those who *do* have both diseases would likely be suffering from both sneezing and a stomach ache. Even though the prior probability of this complex explanation is low, it fits the data very well.

In this case, simplicity seems to be associated with our estimate of prior probability and complexity seems to be associated with our estimate of likelihood. Of course, this explanation was engineered to produce these intuitions by relying on specific beliefs we have about these diseases. The opponent heuristic account proposes that people also use simplicity and complexity as cues in cases where they cannot estimate probabilities directly from background knowledge. Study 1 tests this possibility.

Initial evidence for this idea comes from studies of intuitive curve-fitting—a superficially distinct but deeply related problem to causal explanation. For any set of scatterplot data, many different trend curves can be drawn to explain the data, but statistical theory can tell us which curve has the best predictive power, fitting as much of the underlying signal as possible while fitting little of the noise. Yet, people tend to choose curves that are more *complex* than they normatively should be, rather than curves that are too simple (Johnson, Jin, & Keil, 2014), as one would expect if people only have a simplicity heuristic but no complexity heuristic. Indeed, these curve-fitting studies uncovered direct evidence of a complexity bias, because participants judged the more complex to be literally closer fits to the data, even when the actual fit was held constant. This finding is also consistent with naturalistic studies of everyday verbal explanations drawn from an Internet corpus, for which the best explanations actually tend to be fairly complex (Zemla et al., 2017).

Why is this pair of heuristics useful? Simplicity is just the *absence* of complexity. How, then, can a *pair* of heuristics accomplish any more than a single heuristic, when these two heuristics rely on the same cue? While it may seem more parsimonious to assume that people merely use one cue in a U-shaped manner, it is difficult to specify, for any given problem, where the bend in this U should be. Contextual factors (along with background knowledge) must work to calibrate the strength of these two heuristics, in order to produce a unique solution in any given case. Although there is no reason to think that a context-sensitive dual heuristic solution will give an optimal answer, there *is* reason to think that it may bring the reasoned closer to the right part of the hypothesis space, compared to either heuristic working alone or to any cookie-cutter U-shaped response to simplicity that is not calibrated to the explanatory problem. The current studies look at two possible contextual factors that might modulate the strength of the two heuristics.

First, we consider the determinism of the causal system. In previous studies of simplicity (Lombrozo, 2007), explanations have been produced for deterministic causal systems. In such systems, it is *rational* to prefer simple explanations. If disease *A always* causes symptoms *X* and *Y*, while disease *B always* causes symptom *X* and disease *C always* causes symptom *Y*, the issue of likelihoods or goodness-of-fit simply does not come up: Disease *A perfectly* explains the evidence, and so do Diseases *B* and *C* together. The only issue is which explanation has the higher prior probability, and the simplicity heuristic tells us that, absent any other information, the answer is Disease *A*. Therefore, there is no reason to invoke a complexity heuristic to countervail against the presumption of a simple explanation.

In contrast, when the causal system is stochastic, the likelihoods become a more crucial part of the computation. If disease *A* sometimes causes *X* and sometimes causes *Y*, while disease *B* sometimes causes *X* and disease *C* sometimes causes *Y*, it is difficult to evaluate whether the evidence (symptoms *X* and *Y*) are made likelier by disease *A* or by diseases *B* and *C* combined: It depends on the nature of "sometimes." Yet, in the real world, it is the exception rather than the rule to have precise quantitative information about these likelihoods in stochastic systems. If people rely on a complexity heuristic in such cases, they would judge the likelihood of the evidence to be higher for an individual with two diseases than for an individual with one disease. Study 2 tests whether stochastic contexts therefore lead to a weaker simplicity preference.

Second, we consider the content domain of a causal system. People seem to have different beliefs about the causal textures of different domains. Whereas people tend to identify physical events as having relatively few causes, social events are often thought to have many causes (Strickland, Silver, & Keil, 2016). This suggests that people may calibrate their prior expectations to more complex explanations in the social domain, compared to the physical domain. Furthermore, people may even deploy different causal concepts across domains (Lombrozo, 2010). Whereas causal claims about physical systems appear to be evaluated in terms of transference and contact, social causal claims appear to be evaluated counterfactually. This too may reinforce the intuition that physical events typically result from highly specified causal factors, whereas social events result from more complex configurations of counterfactual conditions. Since such complex conditions can seldom be known, social systems are often highly unpredictable.

As a consequence of these domain-specific beliefs, people may rely on simplicity as a cue to prior probability to a differing degree across domains. Whereas simplicity is likely to be a potent heuristic for evaluating explanations of physical causation, it may be a weaker cue for evaluating explanations of social causation, if people have a meta-theory that assigns higher prior probabilities to complex social causal explanations, as compared to physical causal explanations. In addition, if social causal systems are seen as more stochastic, this would increase the importance of the complexity heuristic for evaluating explanations of social causation, as compared to physical causation. With a weaker simplicity heuristic and stronger complexity heuristic, people may have a less pronounced bias toward simple explanations in the social domain. Study 3 tests this idea.

Study 1

To a Bayesian, the key quantities required to compare two hypotheses are the relative prior probabilities of the hypotheses (the *prior odds*), and the relative fit of each explanation to the data (the *likelihood ratio*). Study 1 tests whether people use simplicity to estimate these quantities.

Study 1A seeks converging evidence for Lombrozo's (2007) claim that people assign higher prior probabilities to simple hypotheses. Study 1B tests whether this heuristic favoring simple explanations might be opposed by a heuristic that assigns higher likelihoods to more *complex* explanations: Do people believe that complex explanations are better fits to the data?

Method

Participants in all studies were recruited from Amazon Mechanical Turk. Each study included a series of check questions at the end, and participants were excluded from analysis if they answered more than 33% incorrectly.

Participants (N = 80, 9 excluded) were randomly assigned to Study 1A (making judgments about priors probabilities) or to Study 1B (making judgments about likelihoods). In both studies, participants completed four items about diseases, similar to the following problem:

> There is a population of elves that lives at Gelfert's Glacier. Sometimes the elves have medical problems such as feverish muffets or wrinkled ears.
>
> A **Yewlie infection** can cause feverish muffets.
> A **Yewlie infection** can cause wrinkled ears.
> **Hepz's disease** can cause feverish muffets.
> **Aeona's syndrome** can cause wrinkled ears.
>
> Nothing else is known to cause an elf's muffets to be feverish of the development of wrinkled ears.

On the same screen, participants completed a series of 10 true/false questions to ensure comprehension.

Participants in Study 1A were then asked to judge the relative prior probabilities ("Imagine that we randomly select an elf from Gelfert's Glacier. Which of the following types of elves do you think we are more likely to have selected?") on a 10-point scale, with one end corresponding to the simple explanation ("An elf who has a Yewlie infection only") and one end to the complex explanation ("An elf who has both Hepz's disease and Aeona's syndrome"). Participants in Study 1B were asked to judge the likelihoods ("Imagine an elf who has a Yewlie infection only, and another elf who has both Hepz's disease and Aeona's syndrome. Which elf do you think is more likely to develop both feverish muffets and wrinkled ears?") on the same scale.

Results and Discussion

Data for all studies were recoded so that negative numbers correspond to the simple explanation and positive numbers to the complex explanation.

Participants in Study 1A used a simplicity heuristic, indicating that a randomly selected elf was more likely to have one disease than two diseases [M = –2.19, SD = 1.78; $t(33) = 7.19$, $p < .001$, $d = 1.23$]. This is consistent with Lombrozo's (2007) studies, where overwhelming prior odds were required before participants would favor a complex over a simple explanation in deterministic cases.

However, the story was different for judgments of likelihoods or goodness-of-fit. Here, participants favored the *complex* explanation [M = 1.41, SD = 2.35; $t(36) = 3.65$, $p = .001$, $d = 0.60$]. This complexity bias in estimating likelihoods was substantial in magnitude (d = 0.60), though smaller than the simplicity bias in estimating priors (d = 1.23), at least for these stimuli.

These results shows that people do not blindly prefer simple explanations, but instead calibrate their preferences according to the question asked. Even though the problem did not include any probability information, participants used simplicity and complexity to estimate different probabilistic quantities in opposing ways.

Study 2

In any causal system where there is uncertainty about which explanation is correct, the prior probabilities of each explanation must be less than 1, since otherwise there is no reason to observe any data (as it will fail to move the posteriors). However, the *likelihoods* differ across deterministic and stochastic systems. In deterministic systems, the evidence is always produced with probability 1 by its causes, whereas in stochastic systems, these likelihoods are less than 1.

If explanatory heuristics exist in part because degrees of uncertainty are difficult to estimate and to use in computations, then a simplicity heuristic will always be a useful tool for estimating priors, since they are always uncertain. However, a complexity heuristic is only useful in stochastic systems, where the likelihoods are uncertain. Thus, both heuristics should be at work in stochastic systems (a simplicity heuristic pushing toward simpler explanations and a complexity heuristic pushing toward more complex explanations), whereas only the simplicity heuristic applies in deterministic systems (pushing toward simpler explanations, without an opposing force pushing toward more complex explanations). This leads to the prediction that people should especially favor simple explanations for deterministic systems.

Method

Participants (N = 80, 14 excluded) completed four items corresponding to the cover stories used in Study 1. For one of these items—in the 100% condition—the causal system was described as deterministic, in that the diseases always led to their symptoms (100% likelihood):

> **Tritchet's syndrome** always (*100% of the time*) causes both sore minttels and purple spots.
>
> **Morad's disaease** always (*100% of the time*) causes sore minttels, but the disease never (*0% of the time*) causes purple spots.
>
> When an alien has a **Humel infection**, that alien will always (*100% of the time*) develop purple spots, but the infection will never (*0% of the time*) cause sore minttels.

The other three items corresponded to the 90%, 80%, and 70% conditions, which differed only in the causal system being described as stochastic:

> **Tritchet's syndrome** often ([*80/65/50]% of the time*) causes both sore minttels and purple spots.
>
> **Morad's disaease** often (([*90/80/70]% of the time*) causes sore minttels, but the disease never (*0% of the time*) causes purple spots.
>
> When an alien has a **Humel infection**, that alien will often (([*90/80/70]% of the time*) develop purple spots, but the infection will never (*0% of the time*) cause sore minttels.

After reading this information, participants were asked about their favored explanation ("Which do you think is the most satisfying explanation for Treda's symptoms?") on a scale from 0 (Tritchet's syndrome only) to 10 (both Morad's disease and a Humel infection). The conditions were balanced across the cover stories using a Latin square, and items were completed in a random order.

Results and Discussion

Participants strongly preferred the simple explanation [$M = –3.81$, $SD = 1.95$; $t(65) = 15.84$, $p < .001$, $d = –1.95$] given deterministic (100%) likelihoods. This replicates Lombrozo's (2007) finding that people strongly favor simple explanations in deterministic causal systems.

The key question is whether this preference would differ in the stochastic conditions, where a complexity heuristic would be more likely at play for understanding the likelihoods. To keep the likelihood ratio objectively identical across conditions, the likelihood for the simple explanation must equal the product of the likelihoods for the components of the complex explanation (i.e., 90% × 90% ≈ 80%, 80% × 80% ≈ 65%, and 70% × 70% ≈ 50%). This calculation assumes that people believe diseases to cause their symptoms independently—an assumption that Lombrozo (2007) validated for her very similar stimuli.

As predicted by the opponent heuristic account, the simplicity bias was weaker in each of the three stochastic conditions, although participants still had a robust simplicity preference in each of them [$M = –3.00$, $SD = 2.68$, $t(65) = 9.09$, $p < .001$, $d = 1.12$ for the 90% condition; $M = –2.50$, $SD = 2.58$, $t(65) = 7.86$, $p < .001$, $d = 0.97$ for the 80% condition; $M = –2.48$, $SD = 2.45$, $t(65) = 8.24$, $p < .001$, $d = 1.01$ for the 70% condition]. The simplicity bias in the stochastic conditions, while large (d from 0.97 to 1.12), was smaller compared to the deterministic condition ($d = 1.95$; ps $> .012$), as predicted.

However, this design is subject to concerns about demand characteristics and difficulties with probabilities that are unrelated to the proposed mechanisms. In particular, the deterministic condition set all likelihoods to 100%, whereas the stochastic condition had to set different likelihoods for the simple explanation and for each component of the complex explanation. Could people have relied on a strategy such as comparing these numerical likelihoods (100% vs. 100% and 90% vs. 80% for complex vs. simple), favoring the complex explanation in the stochastic conditions merely because it was superficially associated with higher numbers?

If this were the case, people should be increasingly less biased toward the simple explanation as the difference between the simple and complex likelihoods increased. This difference increases not only between the deterministic and stochastic conditions, but also *across* the stochastic conditions (90% vs. 80%, 80% vs. 65%, and 70% vs. 50%). Thus, on this deflationary account there should be large gaps not only between the deterministic and stochastic conditions, but also among the stochastic conditions. In contrast, the opponent heuristic account predicts a qualitative shift between the deterministic condition and the stochastic conditions that introduce uncertainty into the likelihoods.

The data are more consistent with the latter prediction, as suggested by the similar effect sizes of the simplicity bias across the three stochastic conditions. There is a significant difference between the 100% and 90% conditions, where we shift from deterministic to stochastic [$t(65) = 2.61$, $p = .011$, $d = 0.32$]. However, the difference between the 90% and 80% conditions reaches only marginal significance [$t(65) = 1.88$, $p = .064$, $d = 0.23$] and the difference between the 80% and 70% conditions is nowhere near significant [$t(65) = 0.04$, $p = .97$, $d = 0.01$]. The deflationary account would predict equally large differences across these sets of conditions.

Thus, determinism may play a role in striking the balance between the simplicity and complexity heuristics. These results also resolve a puzzle about Lombrozo's (2007) findings. Given that people are reasonably well-calibrated in evaluating explanations in the real world, it is surprising to see such a striking simplicity bias as one finds in her studies, with prior odds of 4-to-1 required to override a simplicity preference when the evidence is perfectly consistent with either hypothesis. Study 2 found that in more ecologically realistic conditions, where the evidence is not perfectly predicted by any explanation, people are more likely to hedge their bets. People may thus make more accurate explanatory inferences in realistic, stochastic environments.

Study 3

A second contextual factor that may influence preferences of simple and complex explanations is a system's *content domain*. People believe that physical events have fewer causes than social events (Strickland, Silver, & Keil, 2016) and use causal concepts relying on physical transference for physical systems but complex counterfactual conditions for social systems (Lombrozo, 2010). Thus, Study 3 tests the possibility that people would use these expectations to calibrate their explanatory inferences, favoring simpler explanations in physical causal systems compared to social systems.

Method

Participants ($N = 479$, 89 excluded) read 12 items across four content domains (physics, biology, artifact, and social), which were deterministic for half of participants and stochastic for the other half. These items had the same format as the items used in Study 2, but the content was replaced with various items in physical (ultraviolet waves, subatomic particles), biological (disease, agriculture, dieting), artifact (robots, clocks, toys), and social (team dynamics, child behavior, and romantic attraction) causal systems. Participants then made explanatory judgments on the same scale as Study 2. Items were completed in a random order.

	Deterministic	Stochastic
Physical	–2.76 (2.10)	–2.15 (2.40)
Biological	–2.59 (2.19)	–2.15 (2.28)
Artifact	–2.32 (2.41)	–1.81 (2.53)
Social	–1.81 (2.71)	–1.22 (2.59)

Table 1: Means (SDs) in Study 3.

Results and Discussion

Table 1 shows the effects of both moderators (negative scores reflecting an overall simplicity preference). First, as in Study 2, participants favored the simple explanations more strongly for deterministic than for stochastic systems [$t(388) = 2.52$, $p = .012$, $d = 0.26$]. Thus, the shift seen in Study 2 was not unique to unfamiliar stimuli, or specific to reasoning about diseases. Rather, it is a general pattern used across many content domains.

Second, the ordering of the means across domains was consistent with predictions. Critically, participants had a much stronger simplicity preference in the physical than in the social domain [$t(389) = 8.62$, $p < .001$, $d = 0.38$]. The biological and artifact domains fell in between, with the strongest preference for the physical, followed by the biological, artifact, then social domains. (Keil, Lockhart, & Schlegel, 2010 find similar patterns in a different task.)

Together, the results of Studies 2 and 3 help to resolve the puzzle of how people could rely on a single cue—an explanation's simplicity—to do two logically independent jobs: estimating the prior and likelihood of an explanation. If contextual moderators can influence the weighting of the simplicity and complexity heuristics, then a reasoner could reach different conclusions about simplicity and complexity in different contexts, in ways which are broadly adaptive.

However, there are lingering puzzles about what determines the strength and even direction of simplicity and complexity preferences. For example, one might have expected inferences to more strongly favor the simple explanations than they did here, given the strong simplicity preferences found for the artificial items in Study 2. The more moderate inferences here may have occurred because the items were seen as more reflective of the real world—where true determinism is rare—leading participants to hedge their bets. Alternatively, participants here could be recruiting background knowledge, relying more on memory rather than reasoning. In that case, the strong simplicity preferences found for artificial items in Studies 1 and 2 may better reflect the underlying reasoning processes.

General Discussion

We set out to understand how people use simplicity to constrain their evaluation of explanations, making tractable an otherwise ill-defined computational problem. Usually, simplicity is a good cue for an explanation's prior probability (intuitively, simple causes require fewer stars to align in order to occur) while complexity is a good cue for an explanation's likelihood or fit to the evidence (since complex causes have more opportunities to cause each aspect of the evidence). Study 1 found direct evidence for both of these *opponent heuristics*, directly asking about participants' priors and likelihoods.

However, our explanatory strategies must be definite enough to provide both a unique answer for a given explanatory problem, but also flexible enough to provide different answers to different problems. The opponent heuristics strategy solves this dilemma by modulating the inference depending on context. Study 2 showed that people shift toward complex explanations in stochastic contexts (because such contexts render a complexity heuristic more computationally relevant), and Study 3 showed that people favor simple explanations to varying degrees across domains, in ways that track people's general expectations about the causal textures of these domains: People believe that physical systems are more linear, whereas social systems are more subject to branching, and people correspondingly favor simple explanations to a greater degree for physical systems.

Explanatory Logic. We view these opponent heuristics as one part of a broader *explanatory logic*—a set of heuristics and strategies that people use for evaluating explanatory hypotheses across a variety of processes in light of our cognitive and informational limitations (Johnson, 2017). Here, we focused on causal explanation and previous work has found similar effects in visual curve-fitting (Johnson, Jin, & Keil, 2014), both tasks requiring people to evaluate competing hypotheses (causes, trend lines) for available data (effects, data points). However, many other processes also take this form, including categorization (which category best explains the features?), theory-of-mind (which mental state best explains the behaviors?), language (which meaning best explains the utterance?), and memory (which past events best explain the details I recall?).

In ongoing work, we have been looking at simplicity heuristics in some of these other processes. For example, people can belong to several categories simultaneously—you can be a feminist bank teller, a Jewish woman, or a gay cognitive scientist. When explaining particular traits, people tend to favor social categorizations that invoke fewer categories simultaneously, but this bias is weaker when the categories are more loosely (i.e., stochastically) associated with the relevant features (Johnson, Kim, & Keil, 2016). Similarly, people favor mental-state explanations that invoke relatively fewer goals to explain a particular behavior, but this simplicity preference is weaker when the goals are more stochastically associated with the behaviors (Johnson, 2017). Thus, opponent simplicity heuristics appear to pervade cognition.

The Adaptive Value of Opponent Heuristics. Our empirical argument for opponent heuristics has required us to engineer situations where people make errors.

Nonetheless, we maintain that under more ecologically realistic conditions, these heuristics often serve us well and help to make explanatory reasoning possible.

If you have a well-specified prior distribution and likelihood function, then you can do no better than normative Bayesian inference. Our participants fell short of this standard, making inferences that were unreasonably biased toward simple explanations and influenced by normatively irrelevant factors.

Yet, in the real world, we often lack access to substantial information about probability distributions. We often are confronted with novel situations in which we cannot calculate but must simply guess, based on what little we can glean from the immediate problem and what minimal cues we can bring to bear from our previous experience. It may be true that people seldom encounter cases where they must diagnose an elf, deciding among unfamiliar diseases on the basis of make-believe symptoms, but it *is* true in real-world medical decision-making that we are often faced with highly limited information. Doctors have built up a corpus of statistical knowledge about some familiar diseases, and medical scientists may have some evidence to bring to bear on less familiar ones. Yet, *no one* has joint probability information about all combinations of diseases and symptoms. We must rely on iffy assumptions and fallible heuristics to make any real progress, even in a highly constrained problem domain such as medical diagnosis.

In other cases, probabilities may be even less evident. When making geopolitical forecasts, assessing the reasons for a friend's odd decision, or debating philosophical conundrums, there may be little relevant prior information, and it may be impossible to model the probabilities with any degree of confidence. This is known as *radical uncertainty* or *Knightian uncertainty* (Knight, 1921), and some thinkers argue that many sources of uncertainty are not quantifiable using probabilities (e.g., Levi, 1974; Mises, 2008/1949). In cases of Knightian uncertainty, the best we can do is to adopt rules that work reasonably well most of the time, much as David Hume has argued that our inductive habits are justified by habit rather than logic (Hume, 1977/1748). The use of simplicity and other explanatory heuristics appears to be one such adaptive habit.

This is not to claim that our explanatory habits are untethered to the world. On the contrary, simplicity is usually an excellent principle to use because there are often multiple explanations, varying in complexity, which fit the data equally well. In such cases, the priors generally *do* favor simple explanations, so a simplicity heuristic is reasonable. But when the explanations vary in likelihood, simplicity will lead us astray, as complex explanations are often better fits to the data. Opponent heuristics allow us to harness both of these general facts to our advantage, while avoiding computations that may be intractable and, in Knightian cases, even impossible.

References

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences*, *7*, 19–22.

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*, 630–633.

Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less *ad hoc* theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, *45*, 1–35.

Hochberg, J., & McAlister, E. (1953). A quantitative approach to figural "goodness." *Journal of Experimental Psychology*, *46*, 361–364.

Hume, D. (1977/1748). *An enquiry concerning human understanding*. Indianapolis, IN: Hackett.

Johnson, S.G.B. (2017). *Cognition as sense-making*. Unpublished doctoral dissertation.

Johnson, S.G.B., Jin, A., & Keil, F.C. (2014). Simplicity and goodness-of-fit in explanation: The case of intuitive curve-fitting. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Johnson, S.G.B., Kim, H.S., & Keil, F.C. (2016). Explanatory biases in social categorization. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Keil, F.C., Lockhart, K.L., & Schlegel, E. (2010). A bump on a bump? Emerging intuitions concerning the relative difficulty of the sciences. *JEP: General*, *139*, 1–15.

Kemp, C., Goodman, N.D., & Tenenbaum, J.B. (2010). Learning to learn causal models. *Cognitive Science*, *34*, 1185–1243.

Knight, F.H. (1921). *Risk, uncertainty, and profit*. Boston, MA: Hart, Schaffner, & Marx.

Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy*, *71*, 391–418.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232–257.

Lombrozo, T. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, *61*, 303–332.

Mises, L. (2008/1949). *Human action: A treatise on economics.* Auburn, AL: Ludwig von Mises Institute.

Pothos, E.M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, *26*, 303–343.

Shipley, E.F. (1993). Categories, hierarchies, and induction. In D.L. Medin (Ed.), *The psychology of learning and motivation: Vol. 30*. San Diego, CA: Academic Press.

Strickland, B., Silver, I., & Keil, F.C. (2016). The texture of causal construals: Domain-specific biases shape causal inferences from discourse. *Memory & Cognition*.

Zemla, J.C., Sloman, S., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review*.