Supplementary materials for: "Simplicity as a Cue to Probability: Multiple roles for Simplicity in Evaluating Explanations"

Study S1

An additional study was run to better understand the results of Study 1 – in particular, the finding that simpler explanations were assigned higher likelihoods than complex ones – and to understand why this differed from previous research that found the opposite effect: higher likelihoods for complex explanations (Johnson et al., 2019). This discrepancy was unexpected, since the scenarios and questions used in Study 1 were quite similar to those used by Johnson et al. (2019, Study 1). The current study therefore examined whether subtle differences in the scenario wording across studies might have led to these different effects on likelihood judgments. Specifically, the current study compared likelihood judgments when scenarios were either worded as in our Study 1, or as in Johnson et al.'s Study 1.[1]

In addition to examining likelihood judgments, the current study also examined people's independence assumptions in these scenarios. As discussed in the main text (and in Johnson et al., 2019), one reason people might report higher likelihoods for simple explanations is if they assumed that symptoms tend to co-occur when caused by a single disease (i.e., the symptoms are not conditionally independent given the disease), but that the symptoms are unrelated when caused by two separate diseases (i.e., the symptoms *are* conditionally independent given the pair of diseases). According to the rules of probability, these independence assumptions should affect the likelihood of the pair of symptoms occurring under each explanation. For example, suppose people think that there is a roughly equal probability of each individual symptom occurring when one has a given disease (say, an 80% chance). Under the complex two-disease explanation, if these probabilities are seen as independent, then the likelihood of both symptoms occurring should be 80% * 80% = 64%. On the other hand, under the simple single-disease explanation, if the symptoms are assumed to co-occur above chance levels, then the likelihood of both symptoms occurring should be greater than 64%. This increase is because the co-occurrence assumption means that if someone has the first symptom, they now have a greater than 80%

---

[1] We also ran a preliminary study that varied both the scenario wording (as in the current study) and the likelihood rating format (either eliciting likelihoods separately for the simple and complex explanations, as in Study 1, vs. asking likelihoods on a single relative likelihood scale that compared the simple and complex explanations, as in Johnson et al.'s (2019) Study 1). This preliminary study found a significant effect of scenario wording on the relative likelihoods of simple vs. complex explanations ($p < .001$), with our wording leading to higher likelihoods for simple compared to complex explanations ($p < .001$), and Johnson's wording leading to non-significantly higher likelihoods for complex compared to simple explanations ($p = .16$). In contrast, there was no effect of likelihood rating format ($p = .71$). The current study therefore further examined the effects of scenario wording using a larger sample size to increase power.

chance of having the second symptom. Indeed, assuming this type of dependence for the single-disease symptoms might be reasonable, if, for example, people think that having a more severe case of a disease (as indicated by having one symptom) means that someone is also more likely to have the other symptoms – that is, they assume there is a common underlying mechanism (such as disease severity) that similarly drives the two effects (Park & Sloman, 2013). Because of the potential importance of these independence assumptions in explaining our results in Study 1, the current study explicitly asked participants to report on these assumptions, and compared them across different scenario wordings.

This study was fully preregistered, including the hypotheses, design, analysis plan, sample size, and exclusion criteria.

Methods

Participants. Participants were 571 adults recruited from the US through Prolific (age: *M* = 37, *SD* = 13; gender: 377 women, 184 men, 10 additional or multiple responses). Participants were excluded from the task if they failed to correctly answer all of the scenario comprehension questions by their second attempt, or if they failed to pass additional attention check questions.

Materials and Procedures. Participants were randomly assigned to one of two scenario wording conditions, which either used our scenario wording from Study 1, or Johnson et al.'s (2019) scenario wording from their Study 1. (These wording conditions will be referred to as VL, or JVK, respectively, referring to the authors' initials). As in our Study 1, all participants completed three trials, where each trial involved reading a scenario, and then answering some questions about it. For each trial, participants first read a disease scenario describing the same causal structure used in the main experiments – i.e., where a single disease (the simple explanation) could cause two symptoms, or two diseases (the complex explanation) could each cause one symptom. The exact wording of this initial scenario varied based on condition; see Table S1 for examples of both types. As in Study 1, after reading each scenario, which remained available throughout the trial, participants were given up to two chances to answer a set of three comprehension questions, and then reported the likelihood of the pair of symptoms occurring under each of the two explanations.

At the end of the third trial, participants also reported their independence assumptions for the symptoms in that third scenario. Asking this only once at the end of the three trials ensured that participants' likelihood responses were not altered by being forced to explicitly think about their independence assumptions. Participants reported these assumptions for both the simple and complex explanation, by responding to two questions with the following format: "Suppose there were two aliens. Each of them has [Tritchet's syndrome/both Morad's disease and a Humel infection]. One of these aliens has sore minttels, the other does not. Which alien (if either) do you think is more likely to have purple spots?" Responses were provided on a continuous scale, where -5 = "The alien WITHOUT sore minttels", 0 = "Both equally likely", 5 = "The alien WITH sore minttels." Positive values here thus indicate that the symptoms are thought to co-occur, 0 indicates that the symptoms are thought to be independent, and negative values indicate the symptoms are negatively dependent or anti-correlated, given the information about the alien’s disease(s).

The three scenarios used for VL’s scenario wording were identical to those used in Study 1. The three scenarios used for JVK’s scenario wording were identical to those used by Johnson et al. (2019, Study 1), with three out of four of their scenarios used here. In order to exactly replicate Johnson et al.’s stimuli, the creatures, disease names and symptoms sometimes differed from those used in our scenarios. In both conditions, all three scenarios were presented in a random order for each participant. Across participants, we counterbalanced whether the simple or the complex explanation was described and rated first.

Table S1: Example Scenario Wording

VL (based on Study 1 in the current work)	JVK (based on Johnson et al., 2019, Study 1)
You are visiting a group of aliens that lives on planet Zorg. The alien, Treda, has two symptoms: **Treda’s minttels are sore and Treda has developed purple spots.** **Tritchet’s syndrome** can cause **sore minttels** and can cause **purple spots**. **Morad’s disease** can cause **sore minttels**, but never causes purple spots. A **Humel infection** can cause **purple spots**, but never causes sore minttels. Nothing else is known to cause an alien’s minttels to be sore or the development of purple spots.	There is a population of aliens that lives on planet Zorg. Sometimes the aliens have medical problems such as sore minttels or purple spots. **Tritchet’s syndrome** can cause sore minttels. **Tritchet’s syndrome** can cause purple spots. **Morad’s disease** can cause sore minttels. A **Humel infection** can cause purple spots. Nothing else is known to cause an alien’s minttels to be sore or the development of purple spots.

Results

The first question we examined was whether the scenario wording affected participants’ likelihood ratings, perhaps explaining the discrepancy between our results in Study 1 and those found in Johnson et al. (2019). In order to test this, the two likelihood ratings were converted to a difference score (simple minus complex) measuring relative likelihoods given the simple vs.

complex explanation. These relative likelihoods were then predicted from the scenario wording (0.5 = VL, -0.5 = JVK). We found that scenario wording did significantly alter the relative likelihoods given to simple vs. complex explanations ($B = 13.72$, $p < .001$). As shown in Figure S1, for VL's scenario wording, the effect from Study 1 replicated: participants reported that simple explanations had higher likelihoods than complex explanations ($B = 7.36$, $p < .001$). In contrast, the scenarios using JVK's wording replicated Johnson et al.'s previous findings, with participants reporting that complex explanations had higher likelihoods than simple explanations ($B = -6.36$, $p < .001$). These results suggest that differences in scenario wording may account for the discrepancy between our findings and previous work.

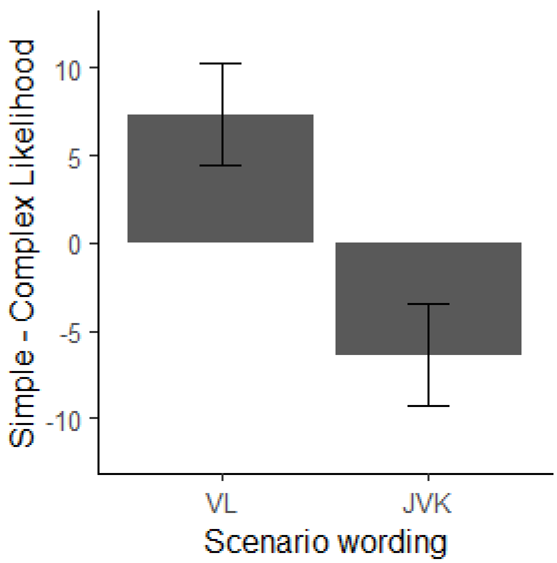

Figure S1: Relative likelihood judgments (simple – complex) for both scenario wordings. VL = the wording from Study 1 in the current work, JVK = Johnson et al.'s (2019, Study 1) wording. 95% CIs shown.

One possible driver of the likelihood effects with VL's scenario wording, and perhaps the differences between the two wordings, might be participants' assumptions about the conditional independence of the two symptoms. In particular, VL's wording may have implied that symptoms co-occurred with the single disease, but were more independent with the two diseases, while, with JVK's wording, perhaps independence/co-occurrence assumptions were similar for both explanations. This was tested by predicting independence ratings from scenario wording, explanation type (0.5 = simple, -0.5 = complex), and the interaction of these. These results showed that scenario wording significantly interacted with explanation type to predict independence ratings ($B = 0.56$, $p = .01$); see Figure S2. With VL's scenario wording, as predicted, symptoms were thought to co-occur given the simple explanation ($M = 0.34$, $SD =$

2.30) but were seen as less likely to co-occur, and indeed, as somewhat anti-correlated, given the complex explanation (*M* = -0.82, *SD* = 2.65; difference: *B* = 1.16, *p* < .001). On the other hand, with JVK's wording, independence assumptions for the two explanations were more similar. In particular, symptoms were seen as largely independent given the simple explanation (*M* = 0.04, *SD* = 2.25), while symptoms were again seen as somewhat anti-correlated given the complex explanation (*M* = -0.56, *SD* = 2.38; difference: *B* = 0.60, *p* < .001). The co-occurrence/independence assumptions implied by VL's wording thus may help explain why participants assigned higher likelihoods to simpler explanations in our scenarios, and the different assumptions across scenario wordings may account for some of the differences in the likelihood effects observed with the two wordings.

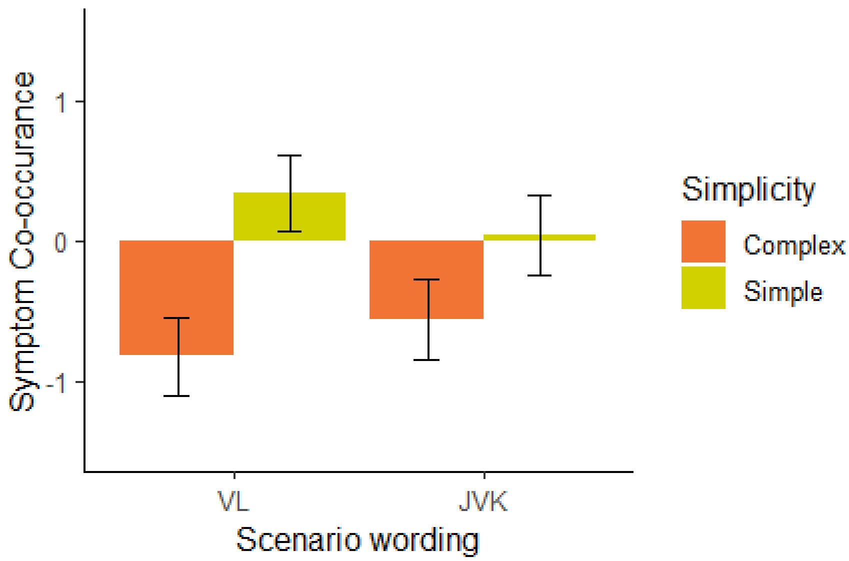

Figure S2: Conditional independence judgments for the two symptoms given either the simple or complex explanation, and both scenario wordings. Positive values indicate symptoms are thought to co-occur, 0 indicates symptoms are thought to be independent, and negative values indicate symptoms are thought to be negatively dependent or anti-correlated. 95% CIs shown.

To further test the idea that the tendency to assign higher likelihoods to simpler explanations, when this occurs, might be caused by these independence assumptions (i.e., assuming symptoms co-occur more when caused by one vs. two diseases), we examined the relationship between individual participants' likelihood judgments and independence judgments. Specifically, we examined the correlation between participants' relative likelihood ratings with their relative co-occurrence/independence ratings (computed as the simple minus the complex rating, where higher values mean that symptoms were judged to co-occur more under simple vs. complex explanations). A significant positive correlation was found between relative likelihoods and relative independence judgments (*r* = 0.15, *p* < .001), indicating that the more participants thought that symptoms co-occur under simple more than complex explanations, the more they assigned these simple explanations higher likelihoods than complex explanations. This further supports the idea that participants' independence assumptions played a key role determining how simplicity was used when inferring likelihoods.

Discussion

Together, these results help us understand how people use simplicity as a cue to likelihoods. In particular, they shed light on why people assigned higher likelihoods to simpler explanations in our studies, and why this effect may have differed from previous research that found the reverse effect (Johnson et al., 2019). The current study suggests that likelihood effects in our studies may have been driven by assumptions about the independence/co-occurrence of the symptoms given different explanations – here, the assumption that symptoms tend to co-occur more when caused by one vs. two diseases. This is consistent with mathematical and philosophical arguments implying that simplicity preferences may often be explained by the higher likelihoods assigned to simpler explanations, but that this crucially depends on additional assumptions about the explanations involved (Sober, 2015).

The more similar independence assumptions for the two explanations in Johnson et al.'s scenarios also partly explain why our effects (higher likelihoods for simple explanations) were not observed in Johnson et al.'s work. However, these more similar independence assumptions should just lead to more similar likelihoods for the two explanations, and do not explain why Johnson et al.'s effects went in the reverse direction, with likelihoods seen as higher for complex explanations vs. simple ones. To explain this, Johnson et al. (2019, see also Johnson et al., 2014), proposed that their results were due a 'complexity heuristic' – i.e. a general tendency for people to think that more complex explanations have higher likelihoods of causing their effects. The current study suggests that people may indeed use this complexity heuristic, but that its effects can be overwhelmed by other factors (e.g., independence assumptions induced by our scenario wording) that can push judgments in the opposite direction.

While the current study showed that the specifics of the scenario wording could alter participants' independence assumptions and likelihood judgments, it is still unclear what it is about the wording that led to these results. One possibility is that these differences stem from how the single disease's symptoms were described. Specifically, in Johnson et al.'s scenarios, it said, "**Tritchet's syndrome** can cause sore minttels. **Tritchet's syndrome** can cause purple spots," vs. in our scenarios, it said, "**Tritchet's syndrome** can cause **sore minttels** and can cause **purple spots**." Describing these symptoms in two separate sentences, vs. a single sentence, may have led people to think of these symptoms as more distinct, independent effects, vs. as a single, co-occurring unit, or the result of a common mechanism. If this is the case, it suggests that people's use of simplicity as a cue to likelihoods may be highly sensitive to subtle cues, such as these shifts in language use. More generally, this further suggests that the use of simplicity as a cue to likelihoods often functions not as an inflexible heuristic, but one that tracks various nuances of the context, and the assumptions people draw based on this.

References

Johnson, S., Jin, A., & Keil, F. (2014). Simplicity and goodness-of-fit in explanation: The case of intuitive curve-fitting. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36).

Johnson, S., Valenti, J. J., & Keil, F. C. (2019). Simplicity and complexity preferences in causal explanation: An opponent heuristic account. *Cognitive Psychology*, *113*, 101222.

Park, J., & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology*, *67*(4), 186–216.

Sober, E. (2015). *Ockham's Razors: A User's Manual*. Cambridge University Press.