

Harry Potter and the sorcerer's scope: latent scope biases in explanatory reasoning

Sangeet S. Khemlani · Abigail B. Sussman · Daniel M. Oppenheimer

© The Psychonomic Society 2010

Abstract What makes a good explanation? We examine the function of *latent scope*, i.e., the number of unobserved phenomena that an explanation can account for. We show that individuals prefer narrow latent scope explanations—those that account for fewer unobserved effects—to broader explanations. In [Experiments 1a–d](#), participants found narrow latent scope explanations to be both more satisfying and more likely. In [Experiment 2](#) we directly manipulated base rate information and again found a preference for narrow latent scope explanations. Participants in [Experiment 3](#) evaluated more natural explanations of unexpected observations, and again displayed a bias for narrow latent scope explanations. We conclude by considering what this novel bias tells us about how humans evaluate explanations and engage in causal reasoning.

Keywords Latent scope · Explanations · Causal reasoning

Introduction

The ability to explain observed behaviors and phenomena is a hallmark of human rationality (Harman, 1965). Explanations serve to illuminate the past and anticipate the future (Craik, 1943; Keil, 2006), and are central to the way we communicate our understanding of the world (Lombrozo, 2007). In recent years, cognitive scientists have begun to examine how explanations constrain category learning and judgment (Ahn, Marsh, Luhmann & Lee, 2002; Murphy & Allopenna, 1994), facilitate conceptual development (Murphy, 2002), and

resolve inconsistencies (Johnson-Laird, Girotto & Legrenzi, 2004; Khemlani & Johnson-Laird, 2010).

So what makes a good explanation? Theorists since William of Ockham have emphasized the role of simplicity in scientific explanation (Newton, 1953/1686; Peirce, 1998; Thornburn, 1918), and recent research has uncovered cognitive biases toward simpler explanations (Chater, 1996; Lagnado, 1994; Lombrozo, 2007). For instance, Lombrozo (2007) asked people to evaluate various explanations of hypothetical medical symptoms of aliens. Participants preferred simple explanations to more complex ones. They thought that a single condition that caused multiple symptoms was a more satisfying explanation of the alien's illness than a conjunction of two other conditions, even though the conjunction could also exhaustively account for the symptoms.

The abundance of data supporting such preferences for simplicity has led some researchers to argue that simplicity is a fundamental cognitive principle (Chater & Vitanyi, 2003). Yet in practice, simplicity may not be the only construct that drives explanatory reasoning. Lombrozo anticipates this, and argues that individuals may use “other explanatory virtues like consistency, scope, and fruitfulness” when assessing different explanations (2007, p. 252). While these other factors might provide a more complete picture of the psychology of explanation, to date there has been little research to determine whether they actually play a role in how reasoners generate or evaluate explanations.

Explanatory scope and latent scope

Philosophers of science often draw analogies between explanations and scientific theories, and they view explanatory scope as a way to distinguish between opposing theories.

S. S. Khemlani (✉) · A. B. Sussman · D. M. Oppenheimer
Department of Psychology, Princeton University,
Princeton, NJ 08540, USA
e-mail: khemlani@princeton.edu

Theories that explain more observed phenomena are said to have broader explanatory scope and are therefore considered stronger explanations of the data than alternative accounts¹. For instance, Thagard (1992, p. 74) notes that “other things being equal, we should prefer a hypothesis that explains more than alternative hypotheses. If hypothesis H1 explains two pieces of evidence and H2 explains only one, then H1 should be preferred to H2.”

Recent evidence appears to corroborate the prediction that explanations with broader scope are often preferred. For instance, Read and Marcus-Newhall (1993) presented participants with information diagnostic of a particular medical condition, e.g., that Cheryl felt tired, gained weight, and had an upset stomach. The authors found that people preferred simpler explanations that account for all of the symptoms (e.g., that Cheryl is pregnant) to more complex combinations of explanations that individually account for only a single symptom (e.g., that Cheryl has mononucleosis, has stopped exercising, and has a virus). Likewise, Preston and Epley (2005) asked participants to consider novel beliefs, other people’s beliefs, and their own cherished religious beliefs, and found that the beliefs that can be used to explain a variety of observations were considered more meaningful, important, and personally relevant than those that can be applied more narrowly. However, despite these studies, the notion of explanatory scope has received less attention in psychology than in philosophy (but see also White, 1997). Part of the difficulty is that explanatory scope assumes that individuals are certain about the data under investigation, i.e., that they compare explanations against a set of observed phenomena. Yet reasoners are often forced to evaluate explanations about uncertain, unobserved phenomena.

The existence of uncertainty suggests an important psychological construct that has not yet been studied: *latent scope*. While explanatory scope considers the number of observed phenomena that a given theory can account for, *latent scope* can be thought of as the number of distinct effects for which an explanation *could* account, but which have not yet been observed. For instance, contrast two explanations for the observation that Lois painted her nails in the shower: (a) she is obsessive–compulsive, or (b) she is afraid of spilling nail polish on her antique bathroom rug. The first explanation has broader latent scope. Being obsessive–compulsive could account for a wide range of possible behaviors: it could explain why she washes her hands many times each day, repeatedly checks to make sure

the door is locked, or studies for hours to get a perfect score on an exam. The alternative explanation can account for a much narrower set of possible behaviors, namely, the steps that Lois might take to keep nail polish from spilling on her antique rug.

Philosophical treatments of explanatory scope, as well as psychological findings (Lombrozo, 2007; Read & Marcus-Newhall, 1993), suggest that people might prefer broader latent scope explanations. For instance, in Lombrozo’s studies, single explanations that could account for multiple observed events were more compelling than multiple smaller scope explanations. As such, people might prefer hypotheses that can generalize well to whatever new data come along.

Alternatively, narrow latent scope explanations might be preferred because of their close match to the observed data. There are fewer predictions made by the explanation about which the observer is uncertain. This may yield a bias towards narrower scope explanations.

We tested whether people preferred narrow scope explanations using both rigorously controlled stimuli (Experiments 1 and 2) and more ecologically valid naturalistic stimuli (Experiment 3). To foreshadow the results, participants robustly preferred narrow latent scope explanations.

Experiments 1a–d

Method

Participants

Thirty participants in Experiment 1a, 31 in 1b, 30 in 1c, and 26 in 1d were recruited through Mechanical Turk, an online platform hosted through Amazon.com (for a discussion on the validity of results from this platform, see Paolacci, Chandler & Ipeirotis, 2010) in exchange for monetary compensation.

Design and procedure

The materials used a background story based on the *Harry Potter* fantasy novels (Boyle, 2004; Rowling, 1997), which served to keep participants engaged and motivated to complete the task. Participants received the following instructions:

In this experiment, you will be presented with several cases of individuals who have strange things happening to them. Their symptoms were undoubtedly caused by Death Eaters’ spells, and it’s your job to figure out which spell or spells caused them.

Their task was to decide which of the “spells” (i.e., explanations) caused the “symptoms” (i.e., effects) that

¹ The earliest treatment of explanatory scope arguably can be attributed to William Whewell (1840), who termed the notion ‘consilience’. The concept has also been referred to as ‘breadth’ (Thagard, 1992), ‘unification’ (Kitcher, 1981), ‘versatility’ (White, 1997), and ‘power’ (Preston & Epley, 2005). The term ‘scope’ appears to come from Kuhn (1977).

were observed for a particular patient. The problems appeared with contents relevant to *Harry Potter*, e.g.:

- Delimenta causes lumps and spots
- Homorula causes lumps, spots, and bumps
- Nothing else is known to cause lumps, spots, or bumps
- Daryl has lumps; we don't know whether or not Daryl also has spots or bumps

In the above case, Delimenta can explain one fewer effect than Homorula, so we refer to these spells as the narrow and broad latent scope explanations, respectively. Participants received four experimental problems in which both explanations could account for the observed effect (as in the above example, both Delimenta and Homorula could have caused lumps) and four control problems in which only the broad latent scope explanation could account for the observed effect (e.g. *Daryl has bumps*, which could only be caused by Homorula) (see Table 1 for a list of all the problems used in these studies). The control conditions ensured that participants were paying attention, understood the task, and were answering coherently when there was an objectively 'correct' answer. The order in which the eight problems appeared was randomized for each participant.

The names of the spells, symptoms, and people used in each problem were assigned randomly.

In **Experiment 1a**, participants were then asked: "What is the most satisfying explanation for Daryl's symptoms?" (c.f. Lombrozo, 2007). Participants responded by typing out their responses into a text box provided on the screen, which gave them the flexibility to suggest multiple spells, unmentioned spells, or to remain noncommittal, if they so desired.

Experiment 1b sought converging evidence using an alternative dependent measure. Participants provided two likelihood ratings: one for how likely it was that the narrow latent scope spell was cast and another for how likely it was that the broad latent scope spell was cast. They responded by registering their response on a 7-point scale such that +3 meant "very likely" and -3 meant "very unlikely".

Moreover, to ensure that participants correctly interpreted the information provided to them, in **Experiment 1b** participants were told that they were evaluating the afflicted individual's health from a separate room by using a magical spell (the "Analytica" spell) which revealed some, but not all, of the afflicted Wizard's symptoms. This instruction made explicit that it was inappropriate to infer the absence of a symptom from the uncertainty of its presence.

Table 1 Problems for Experiments 1a–d and 2

Control premises	Experimental premises
A causes X. B causes X and Y. Nothing else is known to cause X or Y. Y occurred; we don't know whether or not X occurred.	A causes X. B causes X and Y. Nothing else is known to cause X or Y. X occurred; we don't know whether or not Y occurred.
A causes X. B causes X and Y. Nothing else is known to cause X or Y. X and Y occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X occurred; we don't know whether or not Y or Z occurred.
A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X, Y, and Z occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. Y occurred; we don't know whether or not X or Z occurred.
A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. Z occurred; we don't know whether or not X or Y occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X and Y occurred; we don't know whether or not Z occurred.

In **Experiments 1a** and **1d**, participants answered the question "What is the most satisfying explanation?" In **Experiment 1b**, they answered the questions, "How likely is A?" and "How likely is B?", and in **1c**, they answered either the control question, "How likely is Y?" where Y indicates a symptom that is caused by both spells, or else they answered the experimental question "How likely is Z?" where Z indicates a symptom that is caused by only the broad latent scope spell. The problems used in **Experiment 1d** described the premises using the modified verb "always causes" instead of "causes", and participants also received a premise of the form "Both A and B are cast around [1–5]% of the time." In **Experiment 2**, participants answered the questions "How likely is A?" and "How likely is B?" Instead of the premise, "Nothing else is known to cause X, Y, or Z", they received the less ambiguous premise, "Only A and B cause X, Y, and Z." Participants in **Experiment 2** also received an additional premise detailing the frequencies of A and B

If participants comprehended the modified instructions used in [Experiment 1b](#), then the results rule out the possibility that they interpreted the uncertainty of a symptom as its absence. [Experiment 1c](#) used the same dependent measure of likelihood ratings to assess whether the instructions about the Analytica spell had their intended effect. Participants' task was to estimate the likelihood of a symptom that was not explicitly exhibited by the afflicted Wizard in each problem. For [Experiment 1c](#), problems were those in which participants assessed symptoms that could have been caused by either spell, and for experimental problems, participants assessed symptoms that could have been caused by only the broad latent scope spell. Thus, given the example above, participants could have been asked to evaluate how likely it was that Daryl had lumps (a control question) or else how likely it was that he had bumps (an experimental question).

[Experiment 1d](#) sought to rule out the possibility that participants interpreted a premise like "Delimenta causes lumps and spots" as a modal assertion, which can be paraphrased as, "Delimenta *can cause* lumps and spots". To eliminate this possibility, we qualified the verb by adding 'always' in the formulation of the premises, e.g., "Delimenta *always causes* lumps and spots." Further, the study examined whether latent scope biases could be explained by tacit assumptions about base rates, i.e., whether narrow latent scope spells were preferred because they were assumed to happen more often. To eliminate any potential inferences about base rates, participants were told that both spells occurred approximately X% of the time, where X was a number randomly chosen between 1 and 5. Additionally, a novel dependent measure was used. Participants made a forced choice of which of the two spells was more likely to have been the cause of the observed symptoms.

Results and discussion

In [Experiment 1a](#), each response was rated by one of the experimenters and a separate rater who was blind to the hypothesis. The raters agreed on 97.5% of responses, and conflicts were resolved by consensus. [Table 2](#) provides the percentages of responses that favored the narrow latent scope, the broad latent scope, or neither explanation. For

Table 2 Percentages of responses on which the narrow or broad latent scope explanation was chosen as most satisfying for control and experimental problems in [Experiment 1a](#)

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	95	7
Narrow	5	72
Neither	0	21

control problems, 29 out of 30 participants chose the broad scope explanation as satisfying more often than the narrow latent scope explanation (Binomial test, $p < .0001$), and broad scope explanations were favored on 95% of control problems, demonstrating that the task was comprehensible and participants answered sensibly.

For experimental problems in which both the narrow and broad latent scope explanations could account for the observed effects, 25 out of 30 participants chose the narrow latent scope explanation as satisfying more often than the broad scope explanation (Binomial test, $p < .0005$); narrow latent scope explanations were favored on 72% of experimental problems.

For [Experiment 1b](#), [Table 3](#) provides average likelihood estimates for the narrow and broad latent scope explanations as a function of whether they solved control or experimental problems. There was a reliable interaction between problem type and explanation type (Wilcoxon test, $z = 4.69$, $p < .0001$).

Broad latent scope explanations were rated more likely than narrow latent scope explanations for control problems (2.17 vs. -1.33, Wilcoxon test, $z = 4.69$, $p < .0001$), indicating that participants understood the task. As predicted, their preferences switched on experimental problems: they rated narrow latent scope explanations more likely than broader ones (0.76 vs. 0.37, Wilcoxon test, $z = 2.10$, $p < .05$)². The results again supported a latent scope bias whenever no information was present to support one explanation over the other (i.e., for the experimental problems).

In [Experiment 1c](#), participants did not interpret the uncertainty of a symptom's presence as implying its absence. Likelihood estimates of unobserved effects were higher for control problems than experimental problems (1.25 vs. 0.03, Wilcoxon test, $z = 3.35$, $p < .001$), and 25 out of 30 participants exhibited this pattern (Binomial test, $p < .0005$). Likelihood estimates of unobserved effects for experimental problems were not reliably different from chance (Wilcoxon test, $z = .20$, $p = .84$). These data suggest that participants did not interpret the unobserved effects as implying their absence; instead, they understood that effects could be present without being mentioned explicitly.

In [Experiment 1d](#), participants chose the narrow latent scope explanation 3% of the time for control problems and 80% of the time for experimental problems (Wilcoxon test, $z = 4.04$, $p < .0001$), and this pattern was seen across 24 of

² As one reviewer observed, one shortcoming of [Experiment 1](#) is that the inclusion of control conditions encouraged participants to prefer the broad latent scope spell or disease on half the problems. Thus, a simple explanation of the narrow latent scope preference is that participants sought to balance the number of broad and narrow latent scope responses, and so they chose narrow latent scope responses when either spell could occur. We ruled this out by re-running [Experiment 1b](#) without any control problems, and replicated the main finding: 15 Internet participants rated narrow latent scope spells more likely than broader ones (1.32 vs 0.52, Wilcoxon test, $z = 2.69$, $p < .01$). Indeed, the effect appears stronger than when control problems were included.

Table 3 Average likelihood estimates of narrow and broad latent scope explanations for control and experimental problems in [Experiment 1b](#). Estimates range from +3 (very likely) to -3 (very unlikely)

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	2.17	0.37
Narrow	-1.33	0.76

our 26 participants (Binomial test, $p < .0001$). Participants chose the narrow latent scope explanation for experimental problems reliably more than what would be expected by chance (Wilcoxon test, $z = 3.42$, $p < .001$). Even when explicit qualifiers and base rates were added, and for a third dependent measure, the narrow latent scope bias was robust.

[Experiments 1a–d](#) showed that, all else being equal, reasoners preferred narrow latent scope explanations to broader ones when reasoning about uncertain and incomplete information. However, participants' responses may have been driven by implicit likelihood estimates of the two spells. While [Experiment 1d](#) demonstrated the narrow latent scope bias despite equalizing the base rates of the spells, the equalization of base rates may not have convinced participants to disregard base rate information outright. [Experiment 2](#) was designed to test more explicitly whether participants take frequency information into account by systematically varying base rates.

[Experiment 2](#) also provided further evidence that participants did not interpret the uncertainty of a symptom with its absence. While the symptoms used in the previous studies were observable to the naked eye (e.g., “bumps”, “lumps”, and “spots”), [Experiment 2](#) used a medical decision-making domain in which symptoms were unobservable without medical tests. The new domain also brought participants from the magical world of Harry Potter to a controlled decision-making setting that more closely resembled explanatory reasoning in the real world.

Experiment 2

In [Experiment 2](#), instead of reasoning about spells causing symptoms, participants reasoned about diseases leading to abnormal biochemical levels. Biochemicals are inherently difficult to observe unaided, and so we presented participants with the following instructions:

Each biochemical requires a different blood test. While you've ordered blood tests for each biochemical that might be observed in the patient, sometimes the lab gets backed up and not all the results have arrived in your office before you have to make your diagnosis. So, do the best with the information you have.

For example, patients may learn that the disease Vellereum always leads to abnormal alanine levels, and that Pythium always leads to abnormal alanine and valine levels. Moreover, they may learn that a patient had abnormal alanine levels, but that the information about valine had not come back from the lab.

In addition to a novel domain, participants were given information about the frequency with which the two diseases occur, using a method similar to Lombrozo's (2007). For example:

You see about 750 patients a year; about 75 patients are diagnosed with Vellereum and 38 patients are diagnosed with Pythium.

Reasoners who take into account base rate information should note that Vellereum occurs roughly twice as often as Pythium. The study manipulated the relative frequencies of the two candidate diseases, and therefore provided a stronger test of whether latent scope biases could be attributed to assumptions about base rates.

Method

Participants

Twenty-five participants were recruited using the same population as in the previous studies, and completed the study for monetary compensation.

Design

On each problem, participants were presented with a narrow latent scope explanation and a broad latent scope explanation. Participants received four control problems in which only one explanation could account for some observed effects and four experimental problems in which both explanations could account for the observed effects. The problems were structured similar to those shown in [Table 1](#), but they utilized a background story based on a medical decision-making domain. Participants were asked to evaluate how likely the two explanations were, and to register their responses on a 7-point scale in which +3 meant “very likely” and -3 meant “very unlikely”.

Five different base rate conditions were presented to participants. The relative base rates of the two explanations were manipulated such that the ratio between the frequencies of the narrow latent scope explanation to the broad latent scope explanation was 1:2, 2:3, 1:1, 3:2, or 2:1. The base rate conditions rotated so that all of the problems and base rate conditions were viewed the same number of times across the study as a whole. The order in which the eight problems appeared, the names of the diseases, biochemicals, and patients used in each problem were assigned randomly.

Results and discussion

Table 4 provides average likelihood estimates for the narrow and broad latent scope explanations for control and experimental conditions. There was a reliable interaction between problem type and explanation type (Wilcoxon test, $z = 2.55$, $p < .05$). As expected, broad latent scope explanations were rated more likely than narrow ones for control problems, indicating that participants were comfortable with the task (0.80 vs. 0.13, Wilcoxon test, $z = 2.10$, $p < .05$)³. As in the previous studies, preferences switched on experimental problems: participants rated narrow latent scope explanations to be more likely than broader ones (0.85 vs. 0.40, Wilcoxon test, $z = 2.82$, $p < .005$).

Base rate conditions had no reliable effect on the likelihood estimates of broad latent scope diseases (Friedman analysis of variance, $\chi^2 = 7.05$, $p = .13$) or narrow latent scope diseases (Friedman analysis of variance, $\chi^2 = 3.23$, $p = .52$). Figure 1 shows the distribution of likelihood estimates for broad and narrow scope diseases across the five base rate conditions. Base rates did not appear to change the way participants evaluated the explanations.

Despite these results, we do not make strong claims about the effects (or lack thereof) of base rate information. Instead, we highlight that two attempts to examine base rates, one in which the base rates were equalized (Experiment 1d) and the other in which they were directly manipulated (the present study), revealed a latent scope bias. This suggests that the effects of latent scope are probably not due to participants' inferences about base rates.

In the studies presented so far, the latent scope bias occurred when participants reasoned about artificial problems. The problems, however, were unnatural in that they made explicit the effects of each explanation. In real life, this knowledge is implicit or inferred, and is not available for direct inspection. Does a preference for specific explanations extend to more complex, naturalistic contexts? Experiment 3 examined this issue.

Table 4 Average likelihood estimates of narrow and broad latent scope explanations for control and experimental problems in Experiment 2. Estimates range from +3 (very likely) to -3 (very unlikely)

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	0.84	0.40
Narrow	0.15	0.85

³ We suspect that the domain of Harry Potter engaged participants more than did the domain of medical decision-making, which would explain their relative reticence to endorse broad latent scope explanations for control problems in Experiment 2 (0.80) versus Experiment 1b (2.17).

Experiment 3

Method

Participants

Thirty-seven participants were recruited using the same population as in the previous studies, and completed the study for monetary compensation.

Design

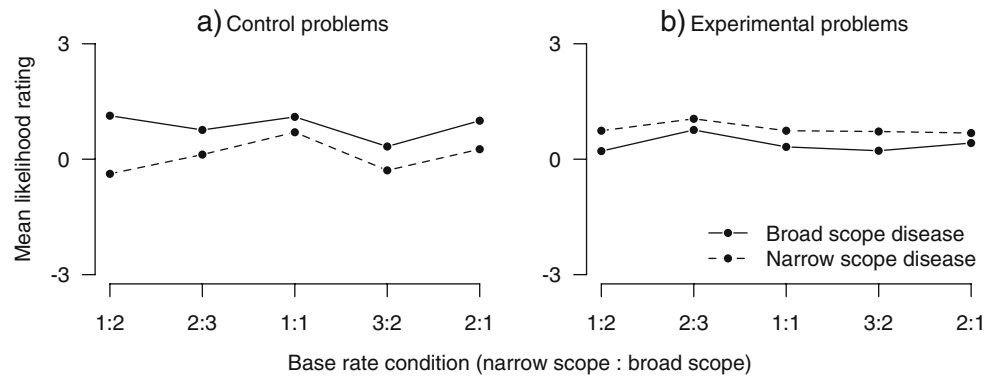
Participants saw ten descriptions of unusual events, and for each description they received four candidate explanations: two broad latent scope and two narrow latent scope. To avoid introducing experimenter bias through stimulus generation, the unusual events and possible explanations were not generated by the researchers, but were instead assembled based on two previous rounds of data collection: the first asked participants to come up with unusual events that might be observed in the world and the second asked participants to generate possible explanations for a subset of those events. A norming study was conducted to ensure that participants viewed narrow latent scope explanations to indeed be narrow, and the broad latent scope explanations to indeed be broad (for details on the norming study and a complete list of items used in Experiment 3, see the Appendix).

The order in which the events and corresponding explanations appeared was randomized for each participant. Participants were shown an event and asked: "What is the most satisfying explanation for this event?" For example, they saw the statement: "George dyed his hair black and then shaved his head" followed in a random order by two broad latent scope explanations (i.e. explanations that could explain the event in question as well as a wide variety of others): "George is going through a mid-life crisis" and "George is going through a fraternity initiation" and two narrow latent scope explanations (i.e. explanations that could explain the event in question but very few others): "George didn't like his new hair color" and "George discovered lice". Participants choose the most satisfying explanation from among these four options.

Results and discussion

For each event, we considered whether participants chose a narrow or broad latent scope explanation. Overall, narrow latent scope explanations were chosen as the most satisfying 64% of the time, which was reliably more often than broad latent scope explanations (Wilcoxon test, $z = 4.16$, $p < .0001$). This preference was reliable for six of the ten events presented, it was in the same direction but unreliable for three of the events, and it was reversed for only one event.

Fig. 1 Average likelihood estimates of narrow and broad latent scope explanations for control (*panel a*) and experimental (*panel b*) problems in Experiment 2 as a function of base rate condition. Estimates range from +3 (very likely) to -3 (very unlikely)



Because the study made use of naturalistic materials, it is susceptible to many criticisms. For instance, the events or explanations may have differed in probability; the probability of the event given the explanation may have been higher for narrower explanations. Thus, taken in isolation it would be difficult to draw strong conclusions from this study. However, in conjunction with the rigorously controlled designs of Experiments 1 and 2, it allows for greater generalizability of the effect. It demonstrates that a small latent scope bias is not restricted to controlled laboratory problems that make latent explanatory scope explicit. Experiment 3 corroborated the hypothesis that people prefer narrow latent scope explanations to broader ones by using novel events and explanations provided by a separate set of participants. This suggests that the tendency to prefer narrow latent scope explanations is robust.

General discussion

Latent explanatory scope refers to the number of unobserved effects for which an explanation could account. Using three distinct dependent measures (free response, Likert scale, and forced choice), we showed a preference for narrow latent scope explanations to broader ones when both explanations could account for the observed effects, even when participants were explicitly given equal base rates. The broad latent scope explanations in Experiment 1 were artificial; they accounted for only one or two more effects than did narrower alternatives. Nevertheless, reasoners were biased to choose the narrow latent scope explanations. This effect was replicated in Experiment 2, which additionally eliminated several alternative explanations for the bias. Experiment 3 revealed that the bias toward narrow latent scope explanations could be found in more natural contexts.

At first blush, our findings appear incompatible with Read and Marcus-Newhall's (1993) results, as well as Preston and Epley's (2005) data. The authors of those studies found that broad scope explanations are preferable to those with narrower scope. However, the crucial

difference between their studies and ours is that we examined how participants would react when they reason about incomplete, unknown information. When participants reason about uncertainty, they may engage in qualitatively different reasoning processes. For instance, they may engage in counterfactual reasoning by imagining what *would* happen if one of the spells had occurred. This type of consideration is not necessary when information is complete.

Our results may also appear to stand in contrast with Lombrozo's (2007) findings that participants prefer simpler explanations, because simplicity and scope seem to be quite dissimilar from one another. However, it may be the case that considerations of simplicity and latent scope are the result of the same cognitive mechanism. In Lombrozo's problems, participants were asked to compare combinations of different causal explanations. In our studies, participants evaluated two causes based on the combinations of their effects. When judging the strength of an explanation, people may prefer explanations that minimize the comparisons they are invited to make. Thus, for our studies they may have minimized the number of unobserved effects, and for Lombrozo's they may have minimized the presence of unwarranted causes.

Why are narrow latent scope explanations more satisfying in the minds of our participants? The result could be explained if reasoners believed that the uncertainty of an effect constituted evidence for its absence. In that case, they might select the spell that exhibits the most positive evidence. However, Experiments 1b-d and 2 rule out the possibility that uncertainty is equated with absence. Whatever strategy our participants employed had to take into account more than just the amount of positive evidence available.

We argue instead that participants exhibited a tradeoff between observed and unobserved information. In other words, they preferred the explanation that predicted the most observed effects and the fewest unobserved effects. Under this account, broad latent scope explanations were penalized for predicting more unobserved effects than narrow latent scope explanations. Indeed, as Experiments 1d and 2 show, such a tradeoff appears to take priority over probabilistic considerations.

One limitation of our studies is that we treated each effect of a cause as either present or else uncertain. However, the degree of uncertainty for a particular effect may modulate participants' scope bias. Future studies should manipulate the uncertainty of unobserved effects by telling participants, e.g., that there is a 65% chance that the patient has abnormal valine levels.

In sum, we investigated how reasoners evaluated explanations and found that they generally preferred narrower explanations to broader ones (pace Preston & Epley, 2005; Read & Marcus-Newhall, 1993). Our studies were unique in assuming that reasoners use uncertain information to make explanatory inferences. The preference we discovered is therefore best characterized as a bias toward narrow *latent* scope, because participants appeared to evaluate the number of effects for which an explanation *could potentially* account. Such a bias may not be limited to explanatory evaluation; it is possible that when people generate explanations, they produce narrower explanations sooner and more often than broader ones. Likewise, since explanations influence categorization, communication, and concept development, these findings have implications across cognition.

Acknowledgments This research was supported by a National Science Foundation Graduate Research Fellowship awarded to the first author. We are grateful for substantive suggestions and criticisms from Shane Blackman, Phil Johnson-Laird, Joanne Kane, Tania Lombrozo, David Mackenzie, Joe McGuire, Ani Momjian, Anuj Shah, Steven Sloman, Bobbie Spellman, Cristina Vinado, Dara Wathanapaisal, Jiaying Zhao, and three anonymous reviewers.

Appendix: norming study and stimuli used for experiment 3

Norming study To generate the materials for Experiment 3, we conducted a norming study in which a separate group of participants rated explanations as having broad or narrow latent scope by characterizing the types of events that were caused by each explanation. Each explanation was rated approximately 25 times by different participants, and participants did not see unexpected events when evaluating explanations. For each explanation, they were presented with the question, "What kind of events might be caused by the fact that [explanation tested]?" and registered their response on a 5-point scale such that 1 meant "a few similar events" and 5 meant "many different events". Average ratings are reported in parentheses below. When responses were pooled across events and explanation types (narrow or broad latent scope), these ratings confirmed that broad latent scope explanations were viewed as intended, i.e., judged to cause a wider variety of events ($M = 3.48$, $SD = .84$) than narrow latent scope explanations ($M = 2.28$, $SD = .74$, Wilcoxon test, $z = 7.56$, $p < .0001$).

Materials used in Experiment 3

Unexpected event	Narrow latent scope explanation #1	Narrow latent scope explanation #2	Broad latent scope explanation #1	Broad latent scope explanation #2
Lance burned his pants after Sunday brunch	Lance dropped his after-meal cigarette (2.19)	Lance stained his pants beyond repair (2.56)	Lance was mentally ill (3.59)	Lance drank too much (3.25)
George dyed his hair black and then shaved his head	George didn't like his new hair color (2.36)	George discovered lice (2.41)	George is going through a mid-life crisis (3.45)	George is going through a fraternity initiation (2.95)
Daniel rode a tortoise to work yesterday	Daniel is protesting traffic congestion (2.45)	Daniel is promoting a new turtle exhibit at the zoo (2.38)	Daniel is on acid (3.04)	Daniel is on a reality TV show (2.59)
Jennifer tried to correct a mistake in an article she wrote on-line with White-Out	Jennifer had already printed the article (2.50)	Jennifer doesn't have basic computer knowledge (3.0)	Jennifer is suffering from Alzheimer's (3.37)	Jennifer is stupid (3.97)
Lois painted her fingernails in the shower	Lois thinks the shower steam causes nails to dry more quickly (1.97)	Lois is afraid of spilling polish on her bathroom rug (2.04)	Lois is obsessive-compulsive (3.48)	Lois is efficient (3.83)
Natalie dumped maple syrup all over her clean laundry	Natalie was testing a new stainguard (1.91)	Natalie heard that maple syrup can get out stains (2.21)	Natalie is irrational (4.22)	Natalie is clumsy (3.22)
At her wedding, Jenny wore yellow rain boots with her white wedding dress	Jenny was marrying the owner of a rain boot company (2.32)	Jenny was wearing yellow rain boots when she met the groom (1.83)	Jenny was dared (3.38)	Jenny had lost a bet (3.19)
Anna gave Larry a stocking full of paperclips for Christmas	Larry always borrows paper clips from Anna (1.57)	Their weekly poker game uses paperclips as currency (2.09)	Anna doesn't like Larry (3.33)	It was a gag gift (3.75)
The church built a handicapped access ramp with Legos	The church is located in Legoland in San Diego (1.87)	The church wants to attract families with young children (3.21)	The church did it as a publicity stunt to attract attention (3.64)	The church wanted to make it into the Guinness Book of World Records (3.00)
The ski resort made red and green colored snow	The resort was designating different ski runs (2.78)	The resort wanted to be festive for Christmas (3.13)	An employee did it as a joke (3.62)	The resort was doing it for publicity (3.76)

References

- Ahn, W., Marsh, J., Luhmann, C., & Lee, K. (2002). Effect of theory-based feature correlations on typicality judgments. *Memory & Cognition, 30*, 107–118.
- Boyle, F. (2004). *An unofficial muggle's guide to the wizarding world: Exploring the Harry Potter universe*. Toronto: ECW Press.
- Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review, 103*, 566–581.
- Chater, N., & Vitanyi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Science, 7*, 9–22. doi:10.1016/S1364-6613(02)00005-0.
- Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.
- Harman, G. H. (1965). The inference to the best explanation. *Philosophical Review, 74*, 88–95.
- Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review, 111*, 640–661. doi:10.1037/0033-295X.111.3.640.
- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227–254. doi:10.1146/annurev.psych.57.102904.190100.
- Khemlani, S., & Johnson-Laird, P. N. (2010). Explanations make inconsistencies harder to detect. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the cognitive science society*. Austin: Cognitive Science Society.
- Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*, 507–531.
- Kuhn, T. S. (1977). *Objectivity, value judgment, and theory choice. In the essential tension: Selected studies in scientific tradition and change*. Chicago: University of Chicago Press.
- Lagnado, D. (1994). *The psychology of explanation: A Bayesian approach*. Masters Thesis. Schools of Psychology and Computer Science, University of Birmingham.
- Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology, 55*, 232–257. doi:10.1016/j.cogpsych.2006.09.006.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 904–919.
- Newton, I. (1953/1686). *Philosophiae naturalis principia mathematica*. Reprinted in H. Thayer (Ed.) *Newton's philosophy of nature*. New York: Hafner.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*, 411–419.
- Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings, 1893–1913*. Bloomington: Indiana University Press.
- Preston, J., & Epley, N. (2005). Explanations versus applications: The explanatory power of valuable beliefs. *Psychological Science, 18*, 826–832.
- Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65*, 429–447.
- Rowling, J. K. (1997). *Harry Potter and the philosopher's stone*. Scholastic.
- Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.
- Thornburn, W. (1918). The myth of Occam's razor. *Mind, 27*, 345–353. doi:10.1093/mind/XXVII.3.345.
- Whewell, W. (1840). *The philosophy of the inductive sciences: Founded upon their history*. London: John W. Parker.
- White, P. (1997). Explanatory versatility and exclusivity as principles of causal judgment. *The American Journal of Psychology, 110*, 159–175.