Explanatory Judgment, Probability, and Abductive Inference

Matteo Colombo (m.colombo@tilburguniversity.edu)
Tilburg Center for Logic, Ethics and Philosophy of Science (TiLPS), Warandelaan 2, 5037 AB Tilburg
Tilburg, The Netherlands

Marie Postma-Nilsenová (m.nilsenova@tilburguniversity.edu)
Tilburg Center for Cognition and Communication (TiCC), Warandelaan 2, 5037 AB Tilburg
Tilburg, The Netherlands

Jan Sprenger (j.sprenger@tilburguniversity.edu)
Tilburg Center for Logic, Ethics and Philosophy of Science (TiLPS), Warandelaan 2, 5037 AB Tilburg
Tilburg, The Netherlands

Abstract

Abductive reasoning assigns special status to the explanatory power of a hypothesis. But how do people make explanatory judgments? Our study clarifies this issue by asking: (i) How does the explanatory power of a hypothesis cohere with other cognitive factors? (ii) How does probabilistic information affect explanatory judgments? In order to answer these questions, we conducted an experiment with 671 participants. Their task was to make judgments about a potentially explanatory hypothesis and its cognitive virtues. In the responses, we isolated three constructs: *Explanatory Value*, *Rational Acceptability*, and *Entailment*. Explanatory judgments strongly cohered with judgments of causal relevance and with a sense of understanding. Furthermore, we found that Explanatory Value was sensitive to manipulations of statistical relevance relations between hypothesis and evidence, but not to explicit information about the prior probability of the hypothesis. These results indicate that probabilistic information about statistical relevance is a strong determinant of Explanatory Value. More generally, our study suggests that abductive and probabilistic reasoning are two distinct modes of inference.

Keywords: Human Reasoning; Abduction; Explanatory Judgment; Explanatory Value; Probability.

Introduction

Explanatory judgments are central to abductive reasoning, that is: inferring to the available hypothesis that can best explain the evidence. Understanding what determines the explanatory power of a hypothesis is of crucial importance for researchers in different fields: for cognitive psychologists, who study the principles of human reasoning (e.g., Oaksford & Chater 2000); AI researchers, who use abduction for belief revision, knowledge representation and fault diagnosis (e.g., Paul 2000); for philosophers of science and epistemologists, who investigate the rationality of abductive inferences, and their relation to probabilistic inference (Lipton 2004; Douven 2011).

Expanding on a growing body of literature in the philosophy and cognitive psychology of explanation, we investigated (i) how explanatory power relates to other features of a hypothesis with respect to an explanandum, such the hypothesis' degree of confirmation, its causal relevance, logical relations, and the sense of understanding it provides; and (ii) how probabilistic information affects explanatory judgment.

Regarding the first question, empirical research in cognitive and developmental psychology has shown that explanation plays several intertwined roles in human cognition (Lombrozo 2011). Explanatory relationships, particularly causal relationships, guide categorization (Carey 1985, 201 ff; Lombrozo 2009), support inductive inference and learning (Holyoak & Cheng 2011; Legare & Lombrozo 2014), and calibrate metacognitive strategies involved in problem-solving (Chi, De Leeuw, Chiu, & LaVancher 1994).

Explanatory considerations can also contribute to the credibility of a hypothesis. Koehler (1991) reviewed much of the work on how explanation influences subjective probabilities, and argued that merely focusing on a hypothesis as if it were the true explanation of observed data is sufficient to boost the subjective probability assigned to that hypothesis. Explanation has also been demonstrated to influence how probabilities are assigned to one proposition in the light of another. Sloman (1994) found that a proposition boosted the probability assigned to another proposition if they shared an explanation.

In particular, cognitive virtues such as the simplicity (Lombrozo 2007; Bonawitz & Lombrozo 2012) and generalizability (Preston & Epley 2005) of a potentially explanatory hypothesis have been shown to be plausible determinants of explanatory judgment. For example, Bonawitz & Lombrozo (2012) and Lombrozo (2007) quantified the simplicity of an explanation in terms of the number of causes it cites, and provided pre-school and adult participants with information about the base rate of each cause. They found that both children and adults relied on the simplicity of a hypothesis as a cue commensurate to base rate information in the face of uncertainty. These results indicate that explanatory judgment is a function of both simplicity and probabilistic information.

While simplicity, and generalizability may be two determinants of explanation, the *sense of understanding* bestowed by a hypothesis seems to be a robust outcome of good explanations. Rozenblit & Keil (2002), for example, considered the phenomenon of the illusion of explanatory depth, whereby people believe they understand the world in greater detail, coherence, and depth than they actually do. A sense of understanding was found to be significantly stronger for explanatory knowledge relative to other knowledge domains. Along similar lines, the connection between inference and explanatory understanding has been criticized by Trout (2002), who argues that a sense of understanding is frequently deceptive, produced by overconfidence and hindsight bias, and in any case not a good indicator for an actually valuable explanation.

In this body of literature in cognitive psychology, explanations are typically presented as *causal* explanations. This choice coheres with philosophical literature on explanation that stresses that good explanations have to provide information about the causes that bring about the explanandum phenomenon (e.g., Salmon 1971/1984; Woodward 2003; Strevens 2008). However, other philosophical models, such as the covering-law model of explanation, focus on the logical relations between explanans and explanandum, such as logical implication (Hempel 1965). In order to have a better understanding of what determines explanatory judgments, it is therefore natural to compare them to judgments of causality, logical implication, and probability.

Regarding our second question about the relationship between probability and explanation, much of the attention has concentrated on the relationship between abductive and Bayesian inference. For example, van Fraassen (1989) argued that abductive and probabilistic inference are incompatible. If we believe more strongly in good explanations than warranted by their posterior probability, we will violate Bayes' rule and be subject to probabilistic incoherence. However, if abductive inferences have to agree with probabilistic inferences as governed by the axioms of probability and by Bayes' Theorem, it is unclear whether they have independent significance, or whether they are just redundant with respect to probabilistic inference.

To address van Fraassen's challenge, many authors have maintained that explanatory judgments *guide* probabilistic inference, or act as a heuristics for them. For instance, Lipton (2004) argued that explanatory judgments may give a good *descriptive* account of our inferential practices, or help to determine the ingredients of probabilistic inferences (e.g., Henderson 2014). Other authors have turned their attention to explicating explanatory power in probabilistic terms (McGrew 2003; Schupbach & Sprenger 2011; Crupi & Tentori 2012), which could help to better explore whether and how abductive reasoning can be embedded into an overarching probabilistic framework.

Notably, a substantial part of the aforementioned studies and recent literature in the psychology of explanation (Keil & Wilson 2000; Lombrozo 2011, 2012) has been concerned with the question of how explanatory power determines probability judgments. Instead, the reverse question has not gained as much attention. With our study, we hope to advance understanding of how probabilistic information determines explanatory judgments.

We constructed vignettes where participants were given information about the priors and likelihoods of a potential explanation, and were asked to make judgments on its explanatory power, posterior probability, and acceptability, as well as on its logical and causal relation to the explanandum. After eliciting these judgments, we used a principal component analysis to examine how they cluster together. We found that a small set of constructs, which we dubbed *Explanatory Value*, *Rational Acceptability*, and *Entailment*, accounted for most of the variation in the participants' judgments, and that explicit information about statistical relevance relations, but not about base rates, determined judgments of the **Explanatory Value** of a hypothesis. This result suggests that abductive and probabilistic reasoning are distinct modes of inference, and that abduction is often probabilistically incoherent.

In what follows, we first describe the methods we used in the experiment designed to collect our data. We then summarize our findings and put them into a broader perspective.

Experiment

Methods

Participants 744 students at Tilburg University (The Netherlands) participated in our study, and 671 completed it (383 male, M_{age} = 21.5 (*SD* = 2.3)). They were randomly assigned to one version of an experimental vignette. Participants were all proficient English speakers, and participated on a voluntary basis, either for free or in exchange for course credit.

Design and Material We employed three distinct types of vignettes with the same logical structure, but different content. Participants were presented with one version of a vignette, where two possible events were related to two possible explanations for that event.

> **Vignette 1:** *There are two urns on the table. Urn A contains 67% white and 33% black balls, Urn B contains only white balls. One of these urns is selected. You don't know which urn is selected, but you know that the chance that Urn A is selected is 25%, and that the chance that Urn B is selected is 75%. From the selected urn a white ball is taken at random.*
> *Please now consider the hypothesis that Urn A has been chosen.*

Vignette 2 and 3 had more concrete content, closer to cases of ordinary reasoning.

> **Vignette 2:** *Again and again, Ruud has knee problems when playing football. The doctors give him two options: knee surgery or a conservative treatment. If Ruud chooses to go into surgery, he cannot play football for half a year; if he chooses the conservative treatment, there is a 33% chance that he can play again after one month; otherwise (with a chance of 67%) he has to rest longer. You don't know which option Ruud chooses, but you believe that the chance that he chooses surgery is 75%—and that the chance that he chooses the conservative treatment is 25%. A month later a joint friend tells you that Ruud is still unable to play football.*

Please now consider the hypothesis that Ruud has chosen for the conservative treatment.

Vignette 3: *Louise arrives by train in Twin City. Twin city has two districts: West Bank and East Bank. In West Bank, there is only one taxi company, namely Green Taxi Ltd., and all their cabs are green. Green Taxi Ltd. also owns 67% of all cabs in East Bank. The other cabs in East Bank are owned by The Red Taxi Inc., all their cabs are red. Louise does not know which part of the city the train is entering, but judging from her knowledge of Twin City she assumes that there is a 75% chance that she is in West Bank (and a chance of 25% that she is in East Bank). At some point, Louise sees a green cab from the train.*
Please now consider the hypothesis that Louise is in East Bank.

After reading the vignette, participants assessed seven items (the construct names in italics were not provided to the participants) on a Likert scale ranking from 1 ("do not agree at all") to 7 ("fully agree"). For Vignette 1, the seven items read as follows:

Logical Implication The hypothesis logically implies that a white ball has been taken out.

Causal Relevance The hypothesis specifies the cause that a white ball has been taken out.

Confirmation The hypothesis is confirmed by the fact that a white ball has been taken out.

Posterior Probability The hypothesis is probable given that a white ball has been taken out.

Explanatory Power The hypothesis explains that a white ball has been taken out.

Understanding The hypothesis provides understanding why a white ball has been taken out.

Truth The hypothesis is true.

The distinction between explanatory power and explanatory value just serves to keep apart the name of a response variable and a broader cognitive factor associated with a hypothesis.

After filling in this questionnaire, the participants could provide information about the way they made their judgments, and we collected some demographic data.

Vignettes were, in a between-subjects design, varied in two dimensions, corresponding to two independent variables:

IV 1: Statistical Relevance The degree of statistical relevance between the explanans and the explanandum, with the four values "strong/weak disconfirm" and "strong/weak confirm".

IV 2: Prior Probability The prior probability of the hypothesis under consideration (.25, .5, or .75).

For all three types of vignettes, all possible $4 \times 3 = 12$ combinations of the values of these variables were realized in the experiment. For example, in the case of Vignette 1, Statistical relevance was manipulated by changing the color of the ball drawn from the urn and/or by changing the explanatory hypothesis (from Urn A to Urn B), leading to four different conditions ordered according to the degree of confirmation they provide.

Procedure Participants completed the questionnaire on a university PC or their own computer in the digital environment of LimeSurvey installed on a local server. The use of LimeSurvey guaranteed that the data could be protected and provided with a time stamp, and information about the IP address of the respondent. The experiment was self-paced, and took on average approximately 10 minutes to complete.

Results

Prior to the analysis of the effects of vignette manipulation, we explored the interdependencies of the seven items in the response questionnaire. To recall, the participants were asked to judge several features of the hypothesis with respect to the evidence: logical implication, causal relevance, explanatory power, increase in understanding, confirmation, posterior probability and truth. By analyzing the interdependencies with the help of the Pearson zero-order correlation coefficient, we determined whether some of these seven concepts tap on the same dimension and could be identified with each other.

Table 1: Zero-order correlations for 7 items (N = 671), all correlations with $p < .01$.

	1	2	3	4	5	6	7
1. *Log. Implication*	-	.38	.22	.32	.46	.30	.12
2. *Causality*		-	.45	.39	.56	.63	.37
3. *Confirmation*			-	.56	.35	.47	.63
4. *Post Probability*				-	.37	.51	.46
5. *Explan. Power*					-	.60	.28
6. *Understanding*						-	.36
7. *Truth*							-

The correlations are presented in Table 1. The analysis revealed that all of the variables correlated at least with .3 with several other variables, but at most .63. These medium-sized correlations show that the participants did not conflate cognate concepts (e.g., causal relevance and explanatory power) with each other, which would be reflected in correlation coefficients greater than .7. At the same time, the response variables were sufficiently related to each other to motivate a Principal Component Analysis: that is, a decomposition of the seven response variables into constructs that could account for most of the variation in the data.

Principal Component Analysis The factorability of the 7 items was examined with a Principle Component Analysis (PCA). The Kaiser-Meyer-Olkin measure of sampling adequacy was .82 and the Bartlett's test of sphericity was significant ($\chi^2(21)$ = 1790.77, $p < .0001$). The initial eigenvalues showed 51% of variance explained by the first factor, 16% explained by the second factor, and 10% explained by the third factor. A visual inspection of the scree plot revealed a 'leveling off' of eigenvalues after the three factors, therefore, a three factor solution using the oblique rotation was conducted, with the three factors explaining 77% of the variance. All items had primary loadings over .7 Table 2 presents the factor loading matrix (loadings under .30 suppressed). In the remainder, we restrict our analysis to these three factors.

Table 2: Factor loadings and communalities based on a principle component analysis with oblimin rotation for 7 items (N = 671).

	1	2	3	Communality
Log. Implication			.94	.94
Causality	.86			.74
Confirmation		-.84		.77
Post Probability		-.72		.67
Explan. Power	.81			.73
Understanding	.87			.78
Truth		-.88		.75

The names for these factors were derived from the clustering shown in Table 2. Factor 1, **Explanatory Value**, clustered explanatory power together with related features of a hypothesis, such as causality and enhancement of understanding (Dieks & DeRegt, 2005; Strevens 2008). Factor 2, **Rational Acceptability**, captured those features that hang together with the acceptability of a hypothesis: probability, confirmation, and truth. The strong correlations between these features were not surprising: confirmation raises posterior probability, which is in turn an indicator of the truth of a theory. Finally, Factor 3 captured the strength of the logical relation between hypothesis and evidence. Since no other response variable was loaded on this factor, it figured as **Entailment**, showing the link to the response variable Logical Implication.[1]

Tests of Experimental Manipulation We conducted two analyses of variance (ANOVAs) to test the effects of the independent variables, Statistical Relevance and Prior Probability, on Explanatory Value, Rational Acceptability, and Entailment, respectively.[2]

Table 3: Estimated Marginal Means and SE of Explanatory Value by Statistical Relevance and Prior Probability (N = 671). SD = Strong Disconfirm, WD = Weak Disconfirm, WC = Weak Confirm, SC = Strong Confirm.

Explanatory Value				
	Statistical Relevance			
Prior	*SD*	*WD*	*WC*	*SC*
Low	2.02 (.22)	3.02 (.21)	4.72 (.20)	4.81 (.19)
Medium	1.95 (.21)	3.05 (.20)	4.66 (.21)	4.58 (.20)
High	1.88 (.24)	3.12 (.21)	4.63 (.22)	4.64 (.20)

First, we tested the effects of the experimental manipulation on *Explanatory Value*. There was a significant main effect of *Statistical Relevance*, $F(3,659)$ = 118.53, $p < .001$, $\eta^2{}_p = .35$, but no effect of *Prior Probability*, $F(2,659)$ = 0.20, $p = .822$. There was no interaction effect between *Statistical Relevance* and *Prior Probability*, $F(6,659)$ = 0.12, $p = .994$—see Table 3 for the descriptives. A pair-wise comparison with Bonferroni correction of the levels of *Statistical Relevance* showed a significant difference between all levels ($p < .001$), with the exception of Weak and Strong Confirm.

Table 4: Estimated Marginal Means and SE of Rational Acceptability by Statistical Relevance and Prior Probability (N = 671). Abbreviations like in Table 3.

Rational Acceptability				
	Statistical Relevance			
Prior	*SD*	*WD*	*WC*	*SC*
Low	1.93 (.18)	2.93 (.18)	3.45 (.17)	5.39 (.16)
Medium	2.05 (.18)	3.10 (.17)	3.74 (.18)	5.62 (.17)
High	1.95 (.20)	3.20 (.18)	3.48 (.18)	5.82 (.17)

Second, we examined the effects of the experimental manipulation on *Rational Acceptability*. There was again a significant main effect of *Statistical Relevance*, $F(3,659)$ = 223.76, $p < .001$, $\eta^2{}_p = .51$, but no effect of *Prior Probability*, $F(2,659)$ = 1.68, $p = .188$. There was no interaction effect between *Statistical Relevance* and *Prior Probability*, $F(6,659)$ = 0.81, $p = .463$—see Table 4 for the descriptives. A pair-wise comparison with Bonferroni correction of the levels of *Statistical Relevance* showed a significant difference between all levels ($p < .001$), with the exception of Weak Disconfirmation and Weak Confirmation ($p = .005$).

[1]The internal consistency for two of the three scales (the third scale only consisted of one item) was examined using Cronbach's alpha, resulting in alpha .82 for Factor 1 and .79 for Factor 2. Composite scores were calculated for each of the three factors using the mean of the items with primary loadings on each factor. The descriptive values for the newly constructed scales were M = 3.65, SD = 1.91 for Explanatory Value, M = 3.63, SD = 1.87 for Rational Acceptability, and M = 3.75, SD = 2.40 for Logical Implication.

[2]A prior analysis of the effect of the vignette on the three dependent variables revealed that Explanatory Value (but not Rational Acceptability and Entailment) was also affected by the vignette manipulation. For clarity of exposition, the statistics is not included here because the size and direction of the significant main effects and the interaction effect remained the same when vignette was included as a factor.

Table 5: Estimated Marginal Means and SE of Entailment by Statistical Relevance and Prior Probability (N = 671). Abbreviations like in Table 3.

Entailment				
	Statistical Relevance			
Prior	*SD*	*WD*	*WC*	*SC*
Low	2.37 (.27)	2.98 (.26)	5.92 (.26)	4.14 (.25)
Medium	2.33 (.26)	3.00 (.26)	5.93 (.26)	3.69 (.26)
High	1.88 (.30)	3.09 (.27)	5.92 (.27)	3.32 (.26)

Finally, we analyzed the effects of the experimental manipulation on *Entailment*. Similarly to the previous two dependent variables, there was again a significant main effect of *Statistical Relevance*, $F(3,659) = 105.40$, $p < .001$, $\eta^2{}_p = .32$, but no effect of *Prior Probability*, $F(2,659) = 1.23$, $p = .292$. There was no interaction effect between *Statistical Relevance* and *Prior Probability*, $F(6,659) = 0.76$, $p = .598$—see Table 5 for the descriptives. A pair-wise comparison with Bonferroni correction of the levels of *Statistical Relevance* showed a significant difference between all degrees ($p < .001$) with the exception of Weak Confirm and Weak Disconfirm ($p = .007$). Since Entailment measures the impact of the hypothesis on the explanandum rather than vice versa, it is logical that the order deviates from the previous two tables, with Weak Confirm obtaining the highest score.

Discussion

Our study investigated the coherence of judgments of explanatory power with other cognitive factors that affect the evaluation of a hypothesis in token explanations of singular events. With the help of a principal component analysis of participants' responses, we identified several constructs under which the response variables could be subsumed. We also investigated how probabilistic information affected these constructs, and *Explanatory Value* (the most salient construct) in particular. We observed a neat distinction between judgments of Explanatory Value and posterior probability. More generally, participants' judgments on the seven response variables (i.e., Logical Implication, Causality, Explanatory Value, Understanding, Posterior Probability, Confirmation, and Truth) were aligned along three dimensions: *Explanatory Value* (loaded with the response variables Causality, Explanatory Power, and Understanding), *Rational Acceptability* (Posterior Probability, Confirmation and Truth) and *Entailment* (Logical Implication). Participants' judgments on these three factors were strongly affected by changes in statistical relevance, specifically by manipulations of the likelihood of the target hypothesis. Instead, the prior probabilities of the candidate explanatory hypothesis, presented as objective base rates, affected participants' judgments in none of those three dimensions. These results suggest a number of conclusions, and further questions for research.

First of all, explanatory value is a distinct feature of a hypothesis that relates to several other processes such as deductive and inductive reasoning, causal reasoning, and reasoning about truth (Lombrozo 2012). Isolating "causality", "understanding" and "explanatory power" ($\rightarrow$ Explanatory Value) from "confirmation", "posterior probability" and "truth" ($\rightarrow$ Rational Acceptability), our principal component analysis coheres with previous results about the tight connections between explanatory power, causality, and a sense of understanding (Lipton 2004; Keil 2006 Lombrozo 2007; Trout 2002). It also suggests that reasoners distinguish between the concepts of *Explanatory Value*, *Rational Acceptability*, and *Logical Entailment*. However, the number of components behind *Explanatory Value* is obviously limited by the number of our dependent variables. One question for future research is whether *Explanatory Value* presents the same components when other dependent variables are added, or when vignettes with different content or structure are evaluated.

Our results bear on the relationship between abduction and probabilistic inference in at least two ways. On the one hand, they show that the rational acceptability of a hypothesis does not always track its explanatory value. This indicates that abduction cannot simply be reduced to Bayesian inference, at least not from a descriptive point of view. On the other hand, *Explanatory Value* is strongly determined by explicit, numeric information about likelihoods. This finding is consistent with the well-documented phenomenon of base rate neglect (Tversky& Kahneman 1982). It also indicates that in situations where causal detail is kept sparse and the explanandum corresponds to a singular token event, abductive and probabilistically reasoning may pull into different directions: the former is insensitive to information about base rates. This conclusion presents a challenge for those who want to defend the rationality of abductive inference as a way of coming closer to the truth. However, when base rates are presented in a different format, reasoners often take them into account (Gigerenzer& Hoffrage 1995), and it remains an open question whether in those circumstances, explanatory inference can be fully described in probabilistic terms. To explore this possibility, further empirical research is needed about the determinants of explanatory value across a broader range of situations, as well as further theoretical explications of explanatory power (Schupbach 2011; Pacer, Lombrozo, Griffiths, Williams, & Chen 2013). We believe that this exchange between normative theorizing about the nature of explanation and empirical evidence about the psychology of explanatory judgment is the key to our understanding of explanatory judgments and abductive reasoning (Colombo 2016).

Acknowledgments

The project was financially supported by the Netherlands Organisation for Scientific Research (NWO) under Vidi grant #276-20-023 (J.S), by the European Research Council (ERC) under Starting Investigator Grant #640638 (J.S.) and by the Deutsche Forschungsgemeinschaft (DFG) in the Priority Program 1516 "New Frameworks of Rationality" (M.C., J.S.). The authors wish to thank Leandra Bucher, Vincenzo Crupi, Tania Lombrozo, Henrik Singmann, and the members of the aforementioned DFG program for helpful discussion and feedback.

References

Bonawitz, E.B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental psychology, 48,* 1156.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Chi, M.T., De Leeuw, N., Chiu, M.H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science, 18,* 439–477.

Colombo, M. (2016). Experimental Philosophy of Explanation Rising. The case for a plurality of concepts of explanation. *Cognitive Science.* doi: 10.1111/cogs.12340

Crupi, V., & Tentori, K. (2012). A second look at the logic of explanatory power (with two novel representation theorems). *Philosophy of Science, 79,* 365–385.

De Regt, H., & Dieks, D. (2005). A Contextual Approach to Scientific Understanding. *Synthese, 144,* 137–170.

Douven, I. (2011). Abduction. *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.), URL = `<http://plato.stanford.edu/abduction/>`.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review, 102,* 684–704.

Hempel, C.G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: Free Press.

Henderson, L. (2014). Bayesianism and Inference to the Best Explanation. *British Journal for the Philosophy of Science, 65,* 687–715.

Holyoak, K.J., & Cheng, P.W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology, 62,* 135–163.

Keil, F.C. (2006). Explanation and understanding. *Annual Review of Psychology, 57,* 227–254.

Keil, F.C., & Wilson, R.A. (2000). *Explanation and Cognition.* Cambridge, MA: MIT Press.

Koehler, D.J. (1991). Explanation, Imagination, and Confidence in Judgment. *Psychological Bulletin, 110,* 499–519.

Legare, C.H., & Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *Journal of experimental child psychology, 126,* 198–212.

Lipton, P. (2004). *Inference to the Best Explanation.* Second edition. London: Routledge.

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

Lombrozo, T. (2009). Explanation and categorization: How "why" informs "what?". *Cognition, 110,* 248–253.

Lombrozo, T. (2011). The instrumental value of explanations. *Philosophy Compass, 6,* 539–551.

Lombrozo, T. (2012). Explanation and abductive inference. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford Handbook of Thinking and Reasoning*. Oxford: Oxford University.

McGrew, T. (2003). Confirmation, Heuristics, and Explanatory Reasoning. *British Journal for the Philosophy of Science, 54,* 553–567.

Oaksford, M., & N. Chater (2000). *Bayesian Rationality.* Oxford: Oxford University Press.

Pacer, M., Lombrozo, T., Griffiths, T., Williams, J., & Chen, X. 2013. Evaluating computational models of explanation using human judgments. In *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, pp. 498–507.

Preston, J., & Epley, N. (2005). Explanations Versus Applications: The Explanatory Power of Valuable Beliefs. *Psychological Science, 16,* 826–832.

Rozenblit, L. R., & Keil, F. C. (2006). The Misunderstood Limits of Folk Science: An Illusion of Explanatory Depth. *Cognitive Science, 26,* 521–562.

Salmon, W. (1971/1984). Statistical Explanation. Reprinted in Salmon (1984): *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Schupbach J. (2011). Comparing Probabilistic Measures of Explanatory Power. *Philosophy of Science, 78,* 813–829.

Schupbach, J., & Sprenger J. (2011). The Logic of Explanatory Power. *Philosophy of Science, 78,* 105–127.

Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition, 52,* 1–21.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Cambridge/MA: Harvard University Press.

Trout, J.D. (2002). Scientific Explanation and the Sense of Understanding. *Philosophy of Science, 69,* 212–233.

Van Fraassen, B.C. (1989). Laws and Symmetry. New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Woodward, J. (2003). *Making Things Happen*. Oxford: Oxford University Press.