Determinants of judgments of explanatory power: Credibility, Generalizability, and Causal Framing

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

Leandra Bucher (bucher@uni-wuppertal.de)
Department of Biological and General Psychology, Bergische Universität Wuppertal,
Max-Horkheimer-Str. 20, 42119 Wuppertal, Germany

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

Abstract

This study investigates how judgments of explanatory power are affected by (i) the prior credibility of a potential explanation, (ii) the causal framing used to describe the explanation, and (iii) the generalizability of the explanation. We found that the prior credibility of a causal explanation plays a central role in explanatory reasoning: first, because of the presence of strong main effects on judgments of explanatory power, and second, because of the gate-keeping role prior credibility has for other factors. Highly credible explanations were not susceptible to causal framing effects. Instead, highly credible hypotheses were sensitive to the generalizability of an explanation. While these results yield a more nuanced understanding of the determinants of judgments of explanatory power, they also illuminate the close relationship between prior beliefs and explanatory power and the relationship between abductive and probabilistic reasoning.

Keywords: Explanation; Prior credibility; Causal framing; Generalizability; Abduction

Introduction

Explanation is a central concept in human psychology. It supports a wide array of cognitive functions, including reasoning, categorization, learning, inference, and decision-making (Lombrozo, 2006; Keil & Wilson, 2000; Keil, 2006). When presented with an explanation of why a certain event occurred, how a certain mechanism works, or why people behave the way they do, both scientists and laypeople have strong intuitions about what counts as a good explanation. Yet, more than sixty years after philosophers of science began to elucidate the nature of explanation (Hempel & Oppenheim, 1948; Hempel, 1965; Salmon, 1989), the determinants of judgments of explanatory power remain unclear.

In this paper, we present three experiments on factors that may affect judgments of explanatory power. Motivated by a large body of theoretical results in epistemology and philosophy of science, as well as by a growing amount of empirical work in cognitive psychology (for respective surveys see Woodward, 2014; Lombrozo, 2012), we examined how judgments of explanatory power are affected by (i) the prior credibility of a potential explanation, (ii) the causal framing used to describe the explanation, and (iii) the generalizability of the explanation.

First we hypothesized that the prior credibility of a causal explanation predicts judgments of explanatory power. Thus, throughout all three experiments, we manipulated the prior credibility of different explanations, and examined the effects of this manipulation on explanatory judgments.

Our focus on the prior credibility of causal explanation was motivated by the fact that most philosophical and psychological analyses of explanatory power agree that powerful explanations provide information about credible causal relationships. Credible causal information facilitates the manipulation and control of nature (Pearl, 2000; Woodward, 2003; Strevens, 2008) and plays distinctive roles in human psychology (Lombrozo, 2011; Sloman & Lagnado, 2015). For example, credible causal information guides categorization (Carey, 1985; Murphy & Medin, 1985; Lombrozo, 2009), supports inductive inference and learning (Holyoak & Cheng, 2011; Legare & Lombrozo, 2014; Walker et al. 2014), and calibrates metacognitive strategies involved in problem-solving (Chi et al., 1994; Aleven & Koedinger, 2002).

Our second, related hypothesis was that presenting an explanatory hypothesis in causal terms predicts judgments of its explanatory power. Thus, we wanted to find out whether people's explanatory judgments are sensitive to causal framing effects.

The importance of this issue should be clear in the light of the fact that magazines and newspapers very often, even when it's not warranted, describe scientific explanations in terms of causal language (e.g., 'Processed meat causes cancer' or 'Economic recession leads to xenophobic violence') with the aim of capturing readers' attention and boosting their sense of understanding (Entmann 1993; Scheufele & Scheufele, 2010). By combining prior credibility and causal framing as predictors of judgments of explanatory power, Experiment 1 and 2 examined the

impact of causality on the explanatory power of scientific hypotheses.

With Experiment 3, we tested the hypothesis that the generalizability (or scope) of a hypothesis determines its explanatory power. While the generalizability of scientific results is an obvious epistemic virtue that figures in the evidential assessments made by scientists, its relation to explanatory power is less clear. Previous psychological findings about the role of generalizability in explanatory reasoning are mixed. Read & Marcus-Newhall (1993) found that generalizability predicts explanatory judgments. Preston & Epley (2005) showed that hypotheses that apply to a wide range of observations are judged as more valuable. However, these studies involved no uncertainty about whether or not a causal effect was actually observed (cf., Khemlani, Sussman, & Oppenheimer, 2011). So, whether or not generalizability is a robust determinant of explanatory judgment remains unclear.

In summary, bringing together different strands of research from philosophy and psychology, our study asks: How do the credibility, causal framing, and generalizability of a hypothesis influence judgments of explanatory power?

The pattern of our experimental findings supports the hypothesis that the prior credibility of a causal explanation plays a central role in explanatory reasoning: first, because of the presence of strong main effects on judgments of explanatory power, and second, because of the gate-keeping role it has for other factors. Highly credible explanations were not susceptible to causal framing effects. Instead, highly credible hypotheses were sensitive to the effects of factors which are usually considered relevant from a normative point of view like the generalizability of an explanation.

Overview of the experiments and pre-tests

To warrant the validity of the experimental material, we conducted a series of pre-studies, where participants evaluated different levels of causal framing, credibility, and generalizability. Materials which corresponded to high, low, and neutral levels of these three factors were implemented in the vignettes of our three experiments, either as independent variables or as control variables.

Material evaluation and main experiments were both conducted online on Amazon Mechanical Turk, utilizing the Qualtrics Survey Software. We only allowed workers with an approval rate > 95% and with a number of HITs approved > 5000 to submit responses. Instructions and material were presented in English.

Causal Framing

A sample of N = 44 participants (mean age 30.5 years, *SD* = 7.3, 28 male) from America (n = 27) and other countries rated eight brief statements, expressing relations between X and Y of the type “X co-occurs with Y”; “X is associated with Y”, and so on. Participants judged how strongly they agreed or disagreed that a certain statement expressed a causal relation between X and Y. Judgments were collected on a 7-point scale with options: "I strongly disagree" (-3), "I disagree", "I slightly disagree", "I neither agree nor disagree" (0), "I slightly agree", "I agree", "I strongly agree" (3). Based on participants’ ratings, we selected three types of statements for our main experiments: statements with a neutral causal framing (“X co-occurs with Y”), with a weak causal framing (“X is associated with Y”), and with a strong causal framing (“X leads to Y" and "X causes Y”).

Prior Credibility

We identified the prior credibility of different hypotheses by asking a new sample of N = 42 participants (mean age 30.7 years, *SD* = 7.5, 16 male) from America (n = 29) and other countries to rate a list of 24 statements. Participants judged how strongly they disagreed or agreed that a certain hypothesis was credible. For all hypotheses, we used the phrasing "... co-occurs with..." to avoid the influence of causal framing. Based on participants’ ratings, we selected four statements to use in our main experiments: two were highly credible, two were highly incredible (Table 1).

Table 1: The four hypotheses rated as least credible and as most credible.

Credibility	Hypothesis
Low	Eating pizza co-occurs with immunity to flu.
Low	Drinking apple juice co-occurs with anorexia.
High	Well-being co-occurs with frequent smiling.
High	Consuming anabolic steroids co-occurs with physical strength.

Generalizability

This pre-study included two questionnaires, which were administered to two different groups of participants. One questionnaire presented descriptions of the samples used in scientific studies, which varied with regard to the *number* of people involved. The other questionnaire presented sample descriptions that varied with regard to the type of people in the sample.

Forty-two participants (mean age 33.5 years, *SD* = 10.8, 27 male) from America (n = 38) and other countries were presented with a list of six statements about a sample of a certain number of participants, e.g. "The study investigates five people"; "The study investigates 500 people".We found that the perceived generalizability of a study increased with the number of people in the sample of the study.

A new group of N = 41 participants (mean age 33.0 years, *SD* = 9.7, 26 male) from America (n = 36) and other countries was presented with a list of nine statements about samples of particular types of people, e.g. "The study investigates a group of people who sit in a park"; "The study investigates a group of people who work at a university". However, focusing on the *number* instead of the *type* of people in the sample allowed for a neater distinction between narrowly and widely generalizable results. Therefore, we characterized generalizability as a function of the number of participants in the main vignettes of the experiment.

Vignettes of the Main Experiment

All experiments were performed using a 2x2 (within-subject) design with explanatory power as dependent variable and prior credibility of the hypothesis being one of the independent variables. The other independent variable was either causal framing, or generalizability.

Participants were presented with four short reports about fictitious research studies. Two of these reports involved highly credible hypotheses, the other two reports involved incredible hypotheses. Two reports showed a high level of the other independent variable, while the other two reports showed a low level of that variable.

Each vignette in our experiments followed the same format as in this sample vignette.

Consuming anabolic steroids leads to physical strength

A recent study by university researchers investigated the link between consuming anabolic steroids and physical strength. The researchers studied 240 persons. The level of physical strength was higher among participants who regularly consumed anabolic steroids than among the participants who did not regularly consume anabolic steroids. Family health history, age, and sex, which were controlled by the researchers, could not explain these results. The study therefore supports the hypothesis that consuming anabolic steroids leads to physical strength.

In all experiments, we varied the level of prior credibility of a hypothesis. In Experiment 1 and 2, we also varied the causal framing and interchanged "leads to" with "causes" and "is associated with", while we kept generalizability at its control. In Experiment 3, we varied the sample size (=generalizability) and controlled for causal framing by using the predicate "co-occurs with" in the headline and the conclusion. Participants were asked to rate our dependent variable: the explanatory power of the stated hypothesis for the results of the study.

Experiment 1 and 2. Credibility x Causal Framing

Participants, Design, and Material

Two-hundred-three participants (mean age 34.7 years, *SD* = 10.5; 121 male) from America (n= 130), India (n = 67) and other countries completed Experiment 1 for a small monetary payment. A new sample of two-hundred-eight participants (mean age 34.56 years, *SD* = 9.97; 124 male) from America (n = 154), India (n = 43), and other countries completed Experiment 2 for a small monetary payment.

In both experiments, participants were presented with four short reports about fictitious research studies along the lines of the above vignette. Across vignettes, we manipulated the causal framing of the relationship between hypothesis and evidence as well as the choice of the hypothesis (credible vs. incredible). Generalizability was controlled for by setting it to its medium value (i.e., 240 participants). Two of the four reports involved highly credible hypotheses, the other two involved incredible hypotheses. Similarly, two of these reports used weak causal framing (Experiment 1 and 2: "X is associated with Y") while the other two used strong causal framing (Experiment 1: "X leads to Y", Experiment 2: "X causes Y"). In other words, Experiment 1 used implicit causal language and Experiment 2 used explicit causal language, while the experiments were identical with respect to design, materials, and procedure.

To account for the possible influence of the content of a particular report, we counterbalanced the allocation of weak and strong causal framing conditions to the credibility conditions across the items, and created two versions of the experiments. The order of reports was individually randomized for each participant.

Participants judged each report in terms of the explanatory power of the hypothesis it described. Specifically, participants considered the statement: "The researchers' hypothesis explains the results of the study", and expressed their judgments on a 7-point scale with the extremes (-3) "I strongly disagree" and (3) "I strongly agree", and the center pole (0) "I neither disagree nor agree".

Analysis and Results

Separate two-way ANOVAs were calculated with the factors Credibility (low, high) and Causal Framing (weak, strong). ANOVA of Experiment 1 (implicit causal language) revealed a main effect of Credibility, F (1, 202) = 84.5; $p < .001$; $\eta_{part}^2 = 0.30$. There was no main effect of Causal Framing ($p = .37$), and no interaction ($p = .08$). Pair-wise comparisons showed that incredible hypotheses were rated significantly lower than credible hypotheses, independently of the value of Causal Framing (incredible hypotheses: $M = 0.26$; $SEM = 0.10$; credible hypotheses: $M = 1.14$; $SEM = 0.09$; t-test: $t(202) = -9.2$; $p < 0.001$; $d = 0.67$). The results of Experiment 1 therefore indicate that the prior credibility of a hypothesis was a strong predictor of judgments of explanatory power (Figure 1). Instead, framing a hypothesis with implicit causal language did not have effects on explanatory judgment.

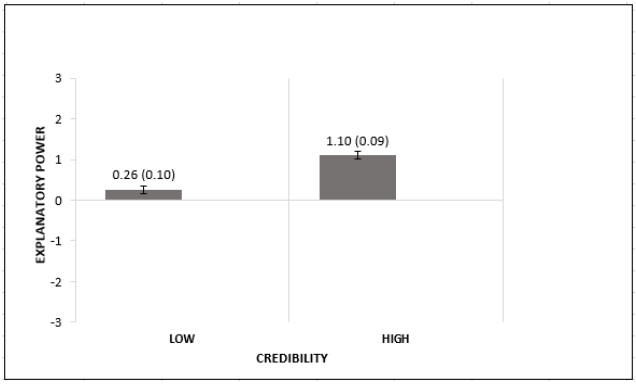

Figure 1: Explanatory power ratings for credible and incredible statements in Experiment 1. Error bars show standard errors of the mean, and are expressed numerically, in parentheses next to the mean value.

ANOVA of Experiment 2 (explicit causal language) revealed main effects of Credibility (F (1, 207) = 286.9; p <.001; η_{part}^2 = 0.58) and Causal Framing, F (1, 207) = 31.0; p <.001; η_{part}^2 = 0.13, as well as a significant interaction Credibility x Causal Framing, F (1, 207) = 37.6; p <.001; η_{part}^2 = 0.15. Figure 2 shows the effect sizes and the interaction between both factors as well as the relevant descriptives.

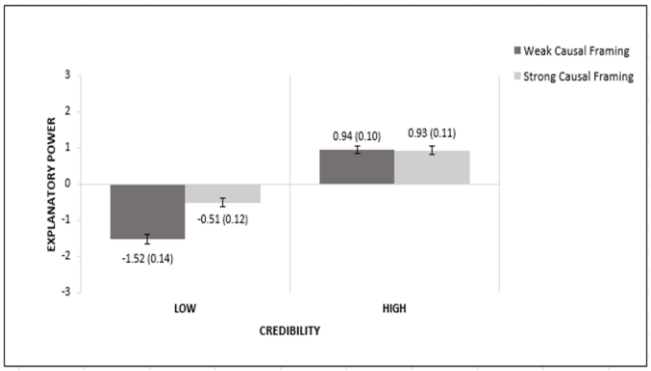

Figure 2: How explanatory power ratings vary with regard to Credibility and Causal Framing (Experiment 2). Error bars show standard errors of the mean and are expressed numerically, in parentheses next to the mean value.

The results of Experiment 2 confirm that the prior credibility of a hypothesis is a strong predictor of judgments of the hypothesis' explanatory power. Incredible hypotheses received negative explanatory power ratings, credible hypotheses receive positive ratings. The results also showed that explicit causal framing can increase ratings of explanatory power, but only for incredible hypotheses. While this effect may lead explanatory judgment astray, in most practical cases of explanatory reasoning, people are interested in the explanatory power of hypotheses which they find, at least to a certain extent, credible. As Figure 2 shows, there was no effect of causal framing on explanatory power in this important case.

This pattern of results confirms that the prior credibility of a hypothesis plays a gate-keeping-role in explanatory reasoning: only credible causal hypotheses qualify as explanatorily valuable. By contrast, implicit or explicit causal framing plays a small to negligible role in influencing judgments of explanatory power.

Experiment 3: Credibility x Generalizability

Participants, Design, and Material

Two-hundred-seven participants (mean age 33.4 years, SD = 9.1; 123 male) from America (n = 156), India (n = 37) and other countries completed Experiment 3 for a small monetary payment.

The experiment resembled Experiment 1 and 2. Four vignettes, each of which included a headline and five sentences, presented credible and incredible hypotheses. The relation between hypothesis and evidence was expressed by using the causally neutral wording "X co-occurs with Y". The critical manipulation concerned the sample descriptions used in the vignettes, which expressed either narrow or wide generalizability of the study's result. For narrowly generalizable results, the second sentence of a report indicated that the sample of the study encompassed around 5 people (e.g. "The researchers studied 6 people"). For widely generalizable results, the sample included about 10,000 people (*wide* generalizability condition, e.g. "The researchers studied 9891 people").

To control for the possible influence of the content of a particular report, we counterbalanced the allocation of narrow and wide generalizability conditions to the credibility conditions across the items, and created two versions of the experiments. The order in which reports were presented to the participants was individually randomized for each participant.

Participants were asked to carefully assess each report with regard to Explanatory Power. Participants' ratings were collected on 7-point scales, with the extreme poles (-3) "I strongly disagree" and (3) "I strongly agree", and the center pole (0) "I neither disagree nor agree".

Analysis and Results

The ratings were analyzed with a two-way ANOVA with the factors Credibility (low, high) and Generalizability (narrow, wide). ANOVA revealed significant main effects of Credibility, F (1, 206) = 83.830; p <.001; η_{part}^2 = 0.289; and Generalizability, F (1, 206) = 29.593; p < .001; η_{part}^2 = 0.126, and no interaction Credibility x Generalizability (p = .085, n.s.).

As with Experiment 1 and 2, credible hypotheses achieved significantly higher ratings than incredible hypotheses (incredible hypotheses: M = -0.01; SEM = 0.10; credible hypotheses: M = 0.95; SEM = 0.08; t-test: t(206) = -9.2; p < .001; d = 0.72). Furthermore, reports with wide generalizability achieved significantly higher ratings compared to reports with narrow generalizability (narrow: M = 0.21; SEM = 0.10; credible hypotheses: M = 0.73; SEM = 0.08; t-test: t(206) = -5.4; p < .001; d = 0.40). Figures 3 and 4 show the main effects for both variables.

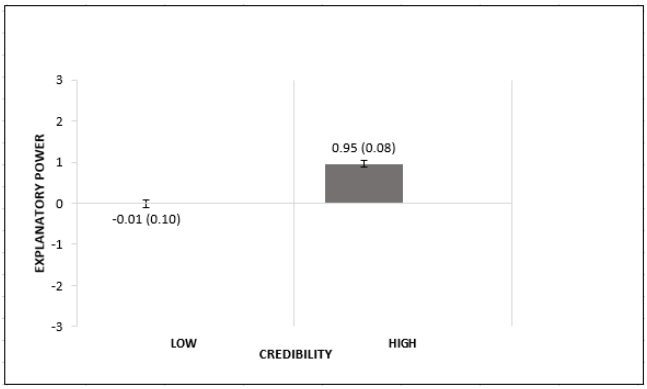

Figure 3: Explanatory power ratings as a function of Credibility. Error bars show standard errors and are also expressed numerically, next to the mean value.

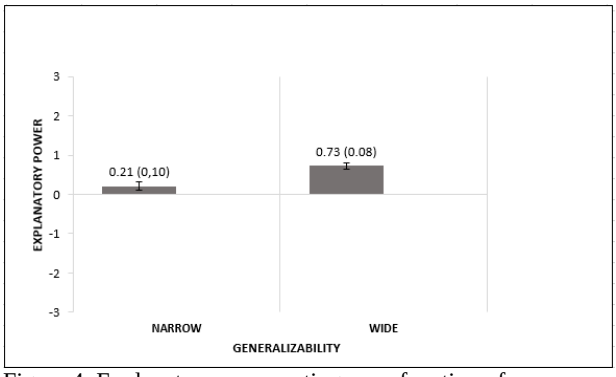

Figure 4: Explanatory power ratings as a function of Generalizability. Error bars show standard errors and are also expressed numerically, next to the mean value.

Discussion

We examined the impact of three factors---prior credibility, causal framing, and generalizability---on judgments of explanatory power. In a series of three experiments, we varied both the subjective credibility of an explanation and one of the other factors: causal framing and generalizability. In Experiments 1 and 2 we found that the impact of causal language on judgments of explanatory power was small to negligible. Experiment 3 showed that generalizable explanations with wider scope positively affected judgments of explanatory power.

Across all experiments, we found that the prior subjective credibility of a hypothesis had a striking effect on how participants assessed explanatory power. In particular, the credibility of an explanatory hypothesis had an important gate-keeping function: the impact of generalizability on explanatory power was more significant when credibility was high. On the other hand, the high credibility of a hypothesis controlled for the potentially misleading effect of causal framing on explanatory judgment.

This pattern of findings is consistent with existing psychological research demonstrating that people resist endorsing explanatory hypotheses that appear unnatural and unintuitive, given their background common-sense understanding of the physical and of the social world (Bloom & Weisberg 2007). Our findings are also consistent with the idea that stable background personal ideologies (often referred to as "worldview") can reliably predict whether people are likely to reject well-confirmed scientific hypotheses (Lewandowsky et al., 2013; Colombo, Bucher, & Inbar, 2016).

So, scientific hypotheses that are inconsistent with our prior, background, common-sense beliefs or in tension with personal ideologies are likely to be judged as implausible, and may not be endorsed as good explanations unless they are supported by extra-ordinary evidence gathered by some trustworthy source. On the other hand, for hypotheses that fit our prior, background belief or ideology, we often focus on information that, if the candidate explanatory hypothesis is true, would boost its goodness (Klayman & Ha 1987).

This kind of psychological process of biased evidence evaluation and retention might have led participants to give the highest ratings of explanatory power, across different experiments, when, in addition to a credible hypothesis, the report was widely generalizable. In comparison, the impact of causal framing was negligible in these cases. This result confirms that a good explanation has to be credible and widely generalizable, and that credible, widely generalizable explanations are not subject to misleading causal framing effects.

The interplay we observed between prior credibility and explanatory power is also relevant to understanding the relationship between abductive and probabilistic reasoning. Highly credible hypotheses were sensitive to the effects of factors which are usually considered explanatory virtues like the generalizability of an explanation.

In abductive reasoning, explanatory considerations are taken to boost the credibility of a target hypothesis while inducing a sense of understanding (Lipton, 2004). Previous psychological studies investigated the effect on people's assessments of explanatory power of factors like simplicity (Lombrozo, 2007; Bonawitz & Lombrozo, 2012) and coherence (Koslowski *et al*. 2008). Our results advance this body of literature by suggesting that the generalizability of a hypothesis will boost the acceptability of the hypothesis, when the hypothesis has a high prior subjective credibility.

High prior credibility may also insulate an explanation from causal framing effects, which may produce a deceptive sense of understanding leading to erroneous explanatory judgments (Rozenblit & Keil, 2002; Trout, 2002).

Overall, our experiments show that explanatory power is a complex concept, affected by considerations of prior credibility of a (causal) hypothesis, and its generalizability. These factors also figure prominently in (normative) philosophical theories of explanation. For instance, the D-N model (Hempel, 1965) stresses the generality of the proposed explanation, and the causal-mechanical account (Woodward, 2003) requires a credible causal mechanism.

On the other hand, the multitude of relevant factors in explanatory judgment explains why it has been difficult to come up with a theory of abductive inference that is both normatively compelling and descriptively accurate: after all, it is difficult to fit diverse determinants of explanatory judgment into a single unifying framework. In that spirit, we hope that our results will promote an interdisciplinary conversation between empirical evidence and philosophical theorizing, and about the "prospects for a naturalized philosophy of explanation" in particular (Lombrozo 2011, 549; Schupbach, 2015; Colombo, 2016).

Acknowledgements

The project was financially supported 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.).

References

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive science*, *26*(2), 147-179.

Bloom, P., & Weisberg, D. S. (2007). Childhood origins of adult resistance to science. *science*, *316*(5827), 996-997.

Bonawitz, E.B. & Lombrozo, T. (2012). Occam's rattle: children's use of simplicity and probability to constrain inference. *Developmental Psychology, 48,* 1156-1164.

Carey, S. (1985). *Conceptual Change in Childhood.* Plenum, Cambridge, MA

Chi, M.T.H., de Leeuw, N., Chiu, M., & 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

Colombo, M., Postma, M., & Sprenger, J. (2016). Explanatory Judgment, Probability, and Abductive Inference. In Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J.C. (Eds.). *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 432-437) Austin, TX: Cognitive Science Society.

Colombo, M., Bucher, L., & Inbar, Y. (2016). Explanatory Judgment, Moral Offense, and Value-Free Science. An Empirical Study. *The Review of Philosophy and Psychology*, 7, *743–763.*

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of communication*, *43*(4), 51-58.

Hempel, C.G..(1965). Aspects of Scientific Explanation. In: *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* The Free Press, pp. 331-496.

Hempel, C.G. & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science* 15: 135-75.

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

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

Khemlani, S. S., Sussman, A. B., & Oppenheimer, D. M. (2011). Harry Potter and the sorcerer's scope: latent scope biases in explanatory reasoning. *Memory & Cognition*, *39*(3), 527-535.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211-228.

Koslowski, B., Marasia, J., Chelenza, M., & Dublin, R., (2008) Information Becomes Evidence when an Explanation Can Incorporate it into a Causal Framework. *Cognitive Development*, 23: 472–487.

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. (2012). Explanation and abductive inference. In K. J. Holyoak & R. G. Morrison (eds.): *Oxford Handbook of Thinking and Reasoning*, 260–276. Oxford, UK: Oxford University.

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

Lombrozo, T. (2009). Explanation and Categorization: How 'Why?' Informs 'What?' *Cognition*, *110*, 248-253.

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

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences, 10,* 464–470.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological review*, *92*(3), 289.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Preston, J. & Epley, N. (2005). Explanations Versus Applications: The Explanatory Power of Valuable Beliefs. *Psychological Science 10,* 826-832.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science* 26, 521–562.

Salmon, W. (1989). *Four Decades of Scientific Explanation*, Minneapolis: University of Minnesota Press.

Scheufele, B. T., & Scheufele, D. A. (2010). Of spreading activation, applicability, and schemas. *Doing News Framing Analysis: Empirical and Theoretical Perspectives, New York, Routledge*, 110-134.

Schupbach, J. N. (2015). Experimental Explication. *Philosophy and Phenomenological Research*. doi: 10.1111/phpr.12207

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*(2), 212-233.

Walker, C.M., Lombrozo, T., Legare, C., & Gopnik, A. (2014). Explaining prompts children to privilege inductively rich properties. *Cognition*, *133*, 343-357.

Woodward, J. (2014). Scientific Explanation. In E.N. Zalta (Ed.) The Stanford Enclycopedia of Philosophy. URL = <https://plato.stanford.edu/entries/scientific-explanation/>