UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How people detect incomplete explanations

Permalink

https://escholarship.org/uc/item/1vg5g5qc

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors

Korman, Joanna

Khemlani, Sangeet

Publication Date

2018

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

How people detect incomplete explanations

Joanna Korman and Sangeet Khemlani
{joanna.korman.ctr, sangeet.khemlani}@nrl.navy.mil
Navy Center for Applied Research in Artificial Intelligence
US Naval Research Laboratory, Washington, DC 20375 USA

Abstract

In theory, there exists no bound to a causal explanation – every explanation can be elaborated further. But reasoners rate some explanations as more complete than others. To account for this behavior, we developed a novel theory of the detection of explanatory incompleteness. The theory is based on the idea that reasoners construct mental models of causal explanations. By default, each causal relation refers to a single mental model. Reasoners should consider an explanation complete when they can construct a single mental model, but incomplete when they must consider multiple models. Reasoners should thus rate causal chains, e.g., *A causes B* and *B causes C*, as more complete than "common cause" explanations (e.g., *A causes B* and *A causes C*) or "common effect" explanations (e.g., *A causes C* and *B causes C*). Two experiments validate the theory's prediction. The data suggest that reasoners construct mental models when generating explanations.

Keywords: explanatory reasoning, incompleteness, causal reasoning, mental models

Introduction

Suppose that you begin to sneeze on a hike through the woods. Here is one explanation for your experience:

1a. Being outside caused you to breathe in pollen.
b. Breathing in pollen caused sneezing.

On the one hand, the explanation may seem complete. On the other hand, every explanation can be elaborated further: assiduous readers may wonder what caused you to be outside in the first place. Their curiosity suggests that reasoners carry out a process to detect whether an explanation is incomplete. No theory of causal reasoning exists that accounts for the process, and so the present paper proposes a novel theory of how reasoners assess explanatory completeness.

Some philosophers of science hold that the notion of a "complete" explanation is nonsensical. Hempel, for instance, observed that an explanation can be judged complete "only if an explanatory account…had been provided for all of its aspects", but that the notion of completeness was "self-defeating" because any explanation can have "infinitely many aspects" (Hempel, 1965/2002). And other theorists concur: for instance, Rescher argued that "the finitude of human intellect" demands that we do not equate the adequacy of an explanation with how complete it is (Rescher, 1995, p. 8; see also Josephson, 2000; Railton, 1981, p. 239).

However, while explanatory completeness may be an intractable notion in the abstract, the finitude of human intellect does not prevent reasoners in daily life from judging whether some explanations are more complete than others. Pioneering work by Miyake (1986) showed that when people explain a particular phenomenon (e.g., how a sewing machine works), they often vacillate between feeling, on the one hand, that their understanding of the phenomenon is satisfactory and complete, and on the other, that their understanding is in need of elaboration. Moreover, Miyake's investigations demonstrated that when constructing explanations, there comes a point at which no further elaboration is possible, either because the relevant information is uncertain or unavailable, or because reasoners may fail to recognize what they do not know. As Keil (2006) observes, the overwhelming complexity of the world puts highly detailed explanations of phenomena beyond the reach of individuals, and so people have no choice but to get by with incomplete, partial explanations. More recently, when Zemla and colleagues (2017) asked participants to evaluate natural explanations, they found that assessments of an explanation's incompleteness predicted judgments of the explanation's quality – the more incomplete an explanation was judged, the worse it was perceived (r = -.65). But, as Zemla et al.'s analysis suggests, explanatory completeness and quality can diverge: it may be possible to generate explanations that are complete but of poor quality. Likewise, it is routine in scientific investigation to generate convincing but tentative explanations, i.e., those that explain available facts but whose internal mechanisms leave relevant causal relations unspecified. Contemporary astronomers, for instance, posit the existence of an as-yet-unobserved ninth planet to explain why the Solar System wobbles away from its center (Batygin & Brown, 2016). Such an explanation is "good" insofar as it accounts for many different observations, but it is incomplete without an articulation of what the planet is made of and how it affects its nearest celestial bodies. Hence, assessments of completeness can diverge from assessments of quality: an explanation's quality depends on corroboratory evidence, while an explanation's completeness depends on identifying and connecting relevant causal relations.

Detecting incompleteness with mental models

Detecting explanatory completeness is an online process that requires reasoners to mentally represent an explanation and assess its structure for potentially unspecified causal relations. In what follows, we present a novel theory that accounts for the mental representations reasoners use to assess some explanations as relatively more complete than others. The theory is based on the idea that humans build small-scale mental simulations – mental models – when they reason. Its central prediction is that reasoners should systematically distinguish complete from incomplete

explanations when they are unable to build an integrated explanatory mental model.

The mental model theory – the "model" theory, for short – posits that people reason on the basis of small, discrete mental representations of possibilities. The theory applies to reasoning in a variety of domains, including explanatory reasoning (Johnson-Laird, Girotto, & Legrenzi, 2004; Khemlani & Johnson-Laird, 2011, 2012), and reasoning about causal, spatiotemporal, and abstract relations (Goldvarg & Johnson-Laird, 2001; Goodwin & Johnson-Laird, 2005). The theory makes three central claims: first, mental models are representations of a conjunction of iconic possibilities (Khemlani, Byrne, & Johnson-Laird, in press). Iconicity implies that the structure of a model corresponds to the structure of what it represents (see Peirce, 1931-1958, Vol. 4). But models can also include abstract symbols, e.g., the symbol for negation (Khemlani, Orenes, & Johnson-Laird, 2012). Second, reasoners distinguish *mental models* – which are initial, incomplete representations that represent only what is true of a given description – from *fully-explicit models* that represent both what is true while keeping track of what is false in a given description. The theory posits two primary processes of inference: the first, an intuitive construction process, rapidly builds and scans initial mental models, but it is subject to various heuristics and biases. The second, a slower, deliberative process, revises the initial models into fully-explicit models, and it can eliminate systematic errors in reasoning (see, e.g., Khemlani & Johnson-Laird, 2017). A third assumption of the theory is that reasoners tend to be parsimonious: the more models that are required to solve a problem, the harder that problem will be, and most reasoners spontaneously draw conclusions based on a single mental model.

The model theory explains how reasoners represent and make inferences from causal relations (Johnson-Laird & Khemlani, 2017; Khemlani, Barbey, & Johnson-Laird, 2014), which underlie causal explanations. It posits that people understand a causal relation as a set of possibilities. For instance, the full meaning of a causal relation with an unknown outcome, such as "going outside causes *X*", refers to a conjunction of three separate models of possibilities, depicted in this schematic diagram:

```
   outside      X
¬  outside      X
¬  outside    ¬ X
```

where '¬' denotes negation. Each row in the diagram represents a different temporally ordered possibility, e.g., the first row represents the possibility in which the person goes outside first and then *X* occurs. The statement rules out the situation in which the person goes outside and *X* does not occur. The model theory accordingly posits that basic causal relations are interpreted deterministically (Frosch & Johnson-Laird, 2011).

A strong prediction of the theory is that when prompted to list the possibilities consistent with a given causal statement, reasoners should list the three possibilities above. Several studies corroborate the prediction (e.g., Bello, Wasylyshyn, Briggs, & Khemlani, 2017; Goldvarg & Johnson-Laird, 2001; Khemlani, Wasylyshyn, Briggs, & Bello, under review). In daily life, however, people do not reason based on the full meanings of causal statements. Instead, they rely on a single mental model to represent "going outside causes *X*", e.g.,

```
outside     X
```

A single model permits rapid inferences because reasoners need to maintain only one possibility in memory. It also permits the rapid construction of a causal chain of events. The theory posits, for instance, that when reasoners comprehend the causal description in (1a-b), they should build an initial model of (1a) first, e.g.,

```
outside     breathing-pollen
```

and then they should integrate a model of (1b) with the model of (1a), e.g.,

```
outside     breathing-pollen     sneezing
```

to create a single model of the phenomenon. One advantage to representing the causal sequence as a single possibility is that reasoners can scan the possibility to rapidly draw temporal inferences, e.g.,

2a. The person breathed in pollen before sneezing.
b. The person sneezed after he went outside.

The inference in (2a) comes about as a result of scanning the possibility from right to left. The inference in (2b) reflects a scan of the possibility from left to right. Hence, an explanatory mental model in the form of an integrated representation of a causal possibility is a productive representation in that it yields sensible temporal and causal inferences.

We propose the *principle of explanatory completeness*, which extends the model theory of causal reasoning to account for how reasoners detect incompleteness. The principle defines a complete causal explanation as a mental model of a single possibility that represents one or more causes and one or more effects. One of the effects constitutes the explanandum, i.e., the thing to be explained. In contrast, incomplete explanations are those that refer to two or more models of possibilities that may or may not share causes and effects. When generating explanations, reasoners should spontaneously construct complete explanatory models instead of incomplete ones.

The principle posits that reasoners should deem an explanation complete if it can be represented by a single causal mental model, e.g., of (1a-b):

```
outside     breathing-pollen     sneezing
```

The principle yields a novel prediction: causal descriptions known as "common cause" and "common effect"

explanations (Read, 1988; Rehder & Hastie, 2004; Salmon, 1978) should be considered less complete than "causal chains" (e.g., 1a-b). For instance, consider the "common effect" explanation in (3):

3. Being outside caused sneezing and having a cold caused sneezing.

The model theory predicts that mental models of the description in (3) should be iconic, i.e., they should reflect the structure of what they represent. Since the description in (3) concerns two separate, unrelated causes, reasoners should construct two separate models to represent the statement, e.g.:

```
outside          sneezing
having-a-cold    sneezing
```

One model of (3) would not suffice, because it would represent the situation in which being outside *and* having a cold caused sneezing.

An analogous argument holds for the "common cause" in (4):

4. Being outside caused sneezing, and being outside caused frostbite.

It requires reasoners to represent two distinct models:

```
outside          sneezing
outside          frostbite
```

Some reasoners may spontaneously consult background knowledge to infer whether the two possibilities can be reconciled, but, failing that, the principle of explanatory completeness predicts that the two possibilities should be considered incomplete.

We describe two experiments that tested the principle of explanatory completeness. The experiments compared causal chains with common effect explanations, as in (3), and common-cause explanations, as in (4), and they served to provide a definitive test of the model theory of explanatory completeness. The principle of explanatory completeness predicts that only causal chains should be represented by a single mental model, and so causal chain structures should be considered more complete than the latter two structures.

Experiment 1

Experiment 1 tested how reasoners assess the completeness of explanations. The model theory predicts that they should consider explanations in the form of causal chains (e.g., *A causes B* and *B causes C*) to be more complete than those in the form of common cause (e.g., *A causes B* and *A causes C*) or common effect structures (e.g., *A causes C* and *B causes C*).

Method

Participants. 51 participants completed the experiment for monetary compensation through Amazon Mechanical Turk. Fifty of the participants were native English speakers, and all but six had taken one or fewer courses in introductory logic.

Task. The experiment invited participants to think of themselves as teachers who had to evaluate their students' explanations for why a particular novel event, *C*, took place. Participants evaluated whether a putative causal explanation for *C* was complete by using a slider bar to indicate a number on a Likert scale from 1 (definitely incomplete) to 5 (definitely complete).

Materials. The content of the events (A, B, and C) was drawn from four separate domains (natural, biological, social, and mechanical). Each set of materials was a collection of candidate properties or behaviors of a novel entity. These properties and behaviors were designed such that any one property or behavior could serve as a cause or a resulting effect of any other. The materials are available at https://osf.io/3ezb5. For instance, one set of materials concerned a mechanical device used in factories called a "Zindo," and so some participants were instructed to evaluate students' explanations for *why the Zindo narrows an aperture*. For example, participants might have evaluated the following explanation:

> Releasing a valve [A] causes the Zindo to engage a pump [B].
> Engaging a pump [B] causes the Zindo to narrow an aperture [C].

On each problem, the experiment was programmed to randomly assign the three properties or behaviors (releasing valve, engaging a pump, and narrowing an aperture) to the event positions (*A*, *B*, and *C*) according to the structures of the problems in the study.

Design and procedure. Participants carried out eight problems altogether. Half of the problems concerned explanations that yielded one model (causal chains), and the other half concerned explanations that yielded multiple models. Two of the four causal chain problems comprised two premises, e.g., *A causes B* and *B causes C*, while the other two comprised a single premise, e.g., *A causes C*, where *A, B,* and *C* stand for various properties and behaviors of imaginary entities.

The other half of the problems concerned explanations that should yield multiple models, i.e., explanations that the theory construes as incomplete. Two of the four multiple-model problems concerned common cause explanations and the other two concerned common effect explanations. Common cause problems provided explanations adhering to the schematic: *A causes B* and *A causes C*, while common effect problems adhered to the schematic: *A causes C* and *B causes C*. The theory predicts that reasoners should be unable to construct an integrated mental model from common cause and common effect explanatory structures, and so they should be judged relatively less complete.

Participants acted as their own controls and carried out all eight problems in a fully repeated measures design. The

experiment was implemented in using the "nodus-ponens" experimental framework for Node.js (Khemlani, 2017). Participants completed two practice trials (one yielding a single mental model, another requiring multiple models), and they received the rest of the problems in a randomized order.

Results and discussion

Figure 1 presents the completeness ratings participants gave for each of the three types of problems presented in Experiment 1. As a whole, participants rated those problems that were predicted to yield one model – i.e., causal chains – as more complete than those predicted to yield multiple models (M = 3.13 vs. M = 2.93), but the difference between the two groups was unreliable (Wilcoxon test, z = 1.52, p = .13). We suspect that the lack of reliability between the two groups was a result of a confound in the design: only causal chains appeared as one-premise problems. When the analysis was restricted to only problems that comprised two premises, participants provided higher ratings for causal chains (M = 3.56), which can be represented with a single explanatory mental model, than for common cause (M = 2.84, Wilcoxon test, z = 3.67, p < .0001, Cliff's δ = .72) or common effect problems (M = 3.02, Wilcoxon test, z =2.64, p = .008, Cliff's δ = .54), both of which require multiple models. The pattern corroborates the principle of explanatory completeness.

Participants gave much lower ratings for causal chain problems containing only a single premise (M = 2.71) than for those containing two premises (z = 3.96, p < .0001, Cliff's δ = .85). Indeed, their ratings for one-premise causal chain problems did not differ reliably from common cause or common effect problems (z = 1.62, p = .11, Cliff's δ = .23). The result ran counter to the principle of explanatory completeness. But it did accord with the results of previous studies, which showed that people prefer explanations that concerned both causes and their effects to explanations that concerned either causes or effects alone (Legrenzi & Johnson-Laird, 2005). For instance, participants in Legrenzi and Johnson-Laird's (2005) study had to select from a set of plausible explanations for why a package was not received. They rated this explanation:

The package went astray because it had the wrong address.

as more probable than this one:

The package went astray.

The results of Experiment 1 are sensible in light of Legrenzi and Johnson-Laird's (2005) finding, but they imply that a) the results of Experiment 1 are confounded, because one-model and two-model problems were unbalanced with respect to the number of premises they contained, and that b) the principle of explanatory completeness has a boundary condition: explanations that consist of single explanatory cause, e.g., explanations of the form *A causes C*, should be considered incomplete. Experiment 2 addressed both issues by dropping shorter descriptions from the design.

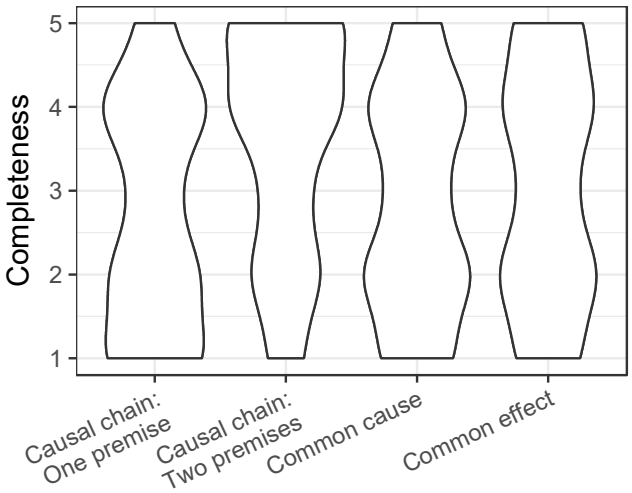

Figure 1. Violin plot of participants' responses to the three conditions in Experiment 1. The width of each shape is proportional to participants' response frequencies.

Experiment 2

Experiment 1 provided partial evidence that people consider explanations that require a single model as more complete than those requiring multiple models. However, the study showed that people tend to prefer more elaborated causal chains (e.g., *A causes B* and *B causes C*) to shallow causal chains (e.g., *A causes C*). Indeed, one-premise causal chains fail to provide an explanation containing both a novel effect and a cause that preceded that effect (Legrenzi & Johnson-Laird, 2005).

Experiment 2 sought to replicate and extend the findings of Experiment 1 by presenting participants with only two-premise problems. Hence, Experiment 2 eliminated a confound of Experiment 1, and it served as a stronger test of the principle of explanatory completeness: if an explanation's perceived completeness depends on specific features of representations – the number of possibilities represented – and not just the presence or absence of an effect and its preceding cause, then reasoners' explicit judgments of completeness should depend on those representations.

Method

Participants. 50 participants completed the experiment for monetary compensation through Amazon Mechanical Turk. All of the participants were native English speakers, and all but eight had taken one or fewer courses in introductory logic.

Preregistration and data availability. The predicted effects were pre-registered through the OSF platform: https://osf.io/sx38c/register/564d31db8c5e4a7c9694b2c0.

The data, experimental code, and materials for Experiment 2 are available at: https://osf.io/3ezb5/.

Design and procedure. The design and procedure for Experiment 2 were identical to that of Experiment 1, except that Experiment 2 included only two-premise problems across all conditions. As in Experiment 1, half of the problems described causal chains, and the other half were split evenly between common-cause and common-effect structures.

Results and discussion

Figure 2 presents the completeness ratings participants gave for each of the three types of problems presented in Experiment 2. As in Experiment 1, participants provided higher ratings for causal chains, which can be represented with a single explanatory mental model (M = 3.26), than for common cause (M = 2.64, Wilcoxon test, z = 3.19, p = .001, Cliff's δ = .62) or common effect problems (M = 2.83, Wilcoxon test, z = 2.05, p = .04, Cliff's δ = .43).

To control for the variance contributed by materials and individual participants, data were subjected to a generalized linear mixed model regression analysis that treated participants' judgments of completeness as the outcome variable and the three types of problem as a fixed effect, and it controlled for material and participant noise. The analysis revealed that the problem types reliably predicted judgments of completeness (B = -0.49, p < .001), further corroborating the theory's prediction.

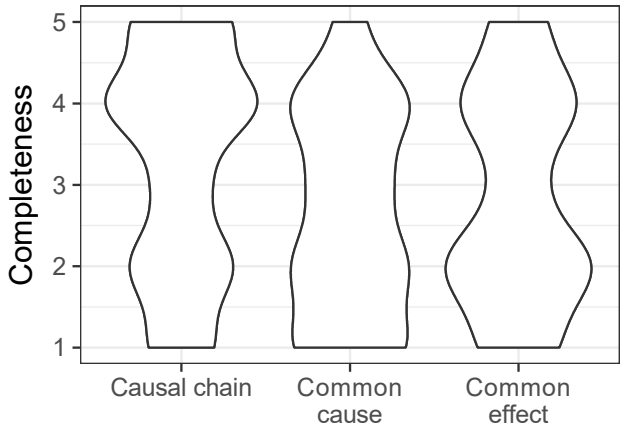

Figure 2. Violin plot of participants' responses to the three conditions in Experiment 2. The width of each shape is proportional to participants' response frequencies.

General discussion

Reasoners construe "complete" explanations as being better than incomplete explanations (Zemla et al., 2017). And, early studies showed that construing an explanation as incomplete allows reasoners to target questions in order to fill in the explanation's gaps (Miyake, 1986). But on any objective notion of completeness, all explanations are incomplete. So what makes people judge an explanation as more or less complete? We argue that, in contrast to philosophical accounts of explanatory completeness, this phenomenon is fundamentally psychologistic: that is, it can only be understood on the basis of subjective, cognitive constraints. We extended the model theory of causal reasoning to explain completeness judgments: it posits that a complete explanation refers to a representation of a single possibility, whereas an incomplete explanation refers to representations of multiple possibilities.

The theory uniquely predicts that reasoners should consider explanations of the form of simple causal chains to be more complete than explanations in the form of "common cause" and "common effect" structures (Read, 1988; Rehder & Hastie, 2004; Salmon, 1978). Two experiments confirmed the theory's prediction, and they suggest that reasoners construct iconic mental representations of causal relations when they generate and evaluate explanations (Khemlani & Johnson-Laird, 2011).

No psychological theory of causal reasoning is fixed in stone, and any of them can be adapted to yield the present prediction. Indeed, explanatory causal chains have been examined in previous work on causal islands (Johnson & Ahn, 2015), explanatory simplicity (Pacer & Lombrozo, 2017), and explanatory coherence (Thagard, 1989). Yet no account prior to the present one has focused on how and why reasoners distinguish complete from incomplete explanations. Unlike other theories of causal reasoning, the model theory precisely characterizes the increased representational burden that incomplete explanations impose on reasoners: it is easier for people to reason from a single possibility than it is to maintain multiple possibilities and to draw inferences from them (Johnson-Laird, 1983; Khemlani & Johnson-Laird, 2017). The model theory thus makes the unique prediction that reasoners should detect incomplete explanations when the burden of representing more than one possibility is present.

Other accounts focus less on how reasoners maintain multiple representations and instead on alternative mechanisms to explain causal inference. For instance, some theorists argue that causal reasoning can be best characterized by causal Bayesian networks (Sloman, 2005; Sloman, Barbey, & Hotaling, 2009). To explain the present data, such an account would need to be extended with mechanisms that maintain, align, and compare multiple networks at a time. A complete causal net would refer a single, integrated network, whereas an incomplete causal net would refer to multiple networks whose interdependent links are expected but unspecified. No such theory has been proposed, but it is a reasonable extension of other researchers' proposals (see Ali, Chater, & Oaksford, 2011). A limitation of this idea, however, is that causal networks might treat causal chains, common cause structures, and common effect structures as equivalently complete, because all three structures refer to a single, integrated network. As the present experiments show, reasoners distinguish between causal chains and other kinds of structures: chains are deemed more complete.

Acknowledgments

This work was supported by an NRC Research Associateship Award to J.K. and funding from the Office of Naval Research to S.K. We thank Paul Bello, Monica Bucciarelli, Ruth Byrne, Felipe de Brigard, Tony Harrison, Laura Hiatt, Phil Johnson-Laird, Robert Mackiewicz, and Greg Trafton for advice. We also thank Kalyan Gupta, Kevin Zish, and Knexus Research Corp.

References

Ali, N., Chater, N., & Oaksford, M. (2011). The mental representation of causal conditional reasoning: Mental models or causal models. *Cognition, 119,* 403-418.

Batygin, K., & Brown, M. E. (2016). Evidence for a distant giant planet in the solar system. *The Astronomical Journal*, 151, 22.

Bello, P., Wasylyshyn, C., Briggs, G., & Khemlani, S. (2017). Contrasts in reasoning about omissions. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Frosch, C.A., & Johnson-Laird, P.N. (2011). Is everyday causation deterministic or probabilistic? *Acta Psychologica*, 137, 280–291.

Goldvarg, Y., & Johnson-Laird, P.N. (2001). Naive causality: a mental model theory of causal meaning and reasoning. *Cognitive Science*, 25, 565−610.

Goodwin, G.P., & Johnson-Laird, P.N. (2005). Reasoning about relations. *Psychological Review*, 112, 468-493.

Hempel, C. (2002). Two models of scientific explanation. In Yuri Balashov & Alexander Rosenberg (eds.), *Philosophy of Science: Contemporary Readings* (pp. 45-55). London: Routledge. (Original work published 1965)

Johnson, S., & Ahn, W. (2015). Causal networks or causal islands? The representation of mechanisms and the transitivity of causal judgment, *Cognitive Science*. 1-36.

Johnson-Laird, P.N. (1983). *Mental models. Cambridge: Cambridge University Press*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P.N., & Khemlani, S. (2017). Mental models and causation. In M. Waldmann (Ed.), *Oxford Handbook of Causal Reasoning* (pp. 1-42). Elsevier, Inc.: Academic Press.

Johnson-Laird, P.N., Legrenzi, P., & Girotto, V. (2004). How we detect logical inconsistencies. *Current Directions in Psychological Science*, *13,* 41-45.

Josephson, J. R. (2000). Smart inductive generalizations are abductions. In *Abduction and induction* (pp. 31-44). Springer Netherlands.

Keil, J. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227-54.

Khemlani, S. (2017). nodus-ponens: A full-stack framework for running high-level reasoning and cognitive science experiments in Node.js. Retrieved from: www.npmjs.com/package/nodus-ponens.

Khemlani, S., Barbey, A., & Johnson-Laird, P. N. (2014). Causal reasoning with mental models. *Frontiers in Human Neuroscience*, 8, 849.

Khemlani, S., Byrne, R.M.J., & Johnson-Laird, P.N. (in press). Facts and possibilities: A model-based theory of sentential reasoning. Manuscript in press at *Cognitive Science*.

Khemlani, S. & Johnson-Laird, P.N. (2011). The need to explain. *Quarterly Journal of Experimental Psychology*, *64*, 2276-88.

Khemlani, S., & Johnson-Laird, P. N. (2012). Hidden conflicts: Explanations make inconsistencies harder to detect. *Acta Psychologica, 139,* 486–491.

Khemlani, S., & Johnson-Laird, P.N. (2017). Illusions in reasoning. *Minds & Machines*, 27, 11-35

Khemlani, S., Orenes, I., & Johnson-Laird, P.N. (2012). Negation: a theory of its meaning, representation, and use. *Journal of Cognitive Psychology*, 24, 541-559.

Khemlani, S., Wasylyshyn, C., Briggs, G. & Bello, P. (under review). Mental models and omissive causation. Manuscript under review.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, *10*, 151-177.

Pacer, M. & Lombrozo, T. (2017). Ockham's razor cuts to the root: Simplicity in causal explanation. *Journal of Experimental Psychology: General*, 146, 17651-1780.

Peirce, C. S. (1931-1958). *Collected papers of Charles Sanders Peirce*. 8 vols. C. Hartshorne, P. Weiss, and A. Burks, (Eds.). Cambridge, MA: Harvard University Press.

Railton, P. (1981). Probability, explanation, and information. *Synthese*, 48, 233–256.

Read, S. (1988). Conjunctive explanations: The effect of a comparison between a chosen and a nonchosen alternative. *Journal of Experimental Social Psychology,* 24, 146-162.

Rehder, B., & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition, 91*, 113-153.

Rescher, N. (1995). Satisfying reason: Studies in the theory of knowledge. Kluwer Academic Publishers: Dordrecht.

Sloman, S. A. (2005). *Causal models: How people think about the world and its alternatives.* Oxford University Press, USA.

Sloman, S.A., Barbey, A.K. & Hotaling, J. (2009). A causal model theory of the meaning of "cause," "enable," and "prevent." Cognitive Science, 33, 21-50.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences,* 12, 435-467.

Zemla, J.C., Sloman, S., Bechlivanidis, C., & Lagnado, D.A. (2017). Evaluating Everyday Explanations. *Psychonomic Bulletin and Review*, 24, 1488-1500.