Title
Causal relations from kinematic simulations

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Khemlani, Sangeet
Goodwin, Geoffrey P
Johnson-Laird, Phil

Publication Date
2015

Peer reviewed

Causal relations from kinematic simulations

Sangeet Khemlani[1], Geoffrey P. Goodwin[2], and Phil Johnson-Laird[3,4]
sangeet.khemlani@nrl.navy.mil, ggoodwin@psych.upenn.edu, phil@princeton.edu

[1]US Naval Research Laboratory, Washington, DC 20375 USA
[2]University of Pennsylvania, Philadepهia PA 19104 USA
[3]Princeton University, Princeton NJ 08540 USA
[4]New York University, New York, NY 10003, USA

Abstract

Reasoners distinguish between different types of causal relations, such as causes, enabling conditions, and preventions. Psychologists disagree about the representations that give rise to the different relations, but agree that mental simulations play a central role in inferring them. We explore how causal relations are extracted from mental simulations. The theory of mental models posits that people use a kinematic simulation to infer possibilities. It predicts that causes should be easier to infer than enabling conditions, and that the time it takes to infer a causal relation should correlate with the number of moves in a mental simulation. To test these two predictions, we adapted a railway domain designed to elicit mental simulations, and we devised problems in which reasoners had to infer causal relations from simulations of the movements of cars in this domain. Two studies corroborated the model theory's predictions. We discuss the results in light of recent theories of causation and mental simulation.

Keywords: causal reasoning, mental models, mental simulation, railway domain, enabling conditions.

Introduction

A man presses a button on his computer that switches on a missile control system. His partner, a woman, presses a button on the missile control system to launch the missile, and a few minutes later the missile is launched. Did the first man cause the missile to launch? Or did he merely enable its launch? Reasoners often have to make judgments of the distinction between causing and enabling, and in doing so they may rely on mental simulations of a sequence of events (as in the example above). Many people may conclude that the man caused the launch, because the temporal contiguity of two events is often all that is required to infer causality (Bramley, Gerstenberg, & Lagnado, 2013; Lagnado & Sloman, 2006; Rottman & Keil, 2012). In the case of the missile launch, the causal inference may be unwarranted, because the description is consistent with alternative possibilities, such as one in which the woman decides not to press the button she controls. Suppressing the initial causal inference requires mental simulation too, because reasoners may engage in a search for alternative possibilities consistent with the description (Frosch & Johnson-Laird, 2011).

Causation is controversial; it has vexed scholars for centuries, and psychologists disagree on its underlying mechanisms (Ahn & Bailenson, 1996; Cheng, 1997; Hilton & Erb, 1996; Sloman, 2005; White, 2014; Wolff, 2007). What is less controversial is the centrality of mental simulation in causal thinking and reasoning: people construct small-scale simulations of possibilities to make predictions of outcomes (Kahneman & Tversky, 1981), to understand mechanisms (Hegarty, 2004) and physical scenes (Battaglia, Hamrick, & Tenenbaum, 2013), to resolve inconsistencies and contradictions (Khemlani & Johnson-Laird, 2012, 2013; Park & Sloman, 2014), to deduce the consequences of algorithms (Khemlani et al., 2013), and to reason about counterfactual scenarios (Byrne, 2005; Galinsky & Moskowitz, 2000).

One challenge for theories of causality is how different causal relations are extracted from a mental simulation. People distinguish between causal relations such as *cause*, *enable*, and *prevent*. For example, they recognize that the meaning of,

1. The button *caused* the missile to launch.

is distinct from,

2. The button *enabled* the missile to launch.

Theorists have appealed to the transmission of force (Wolff, 2007), causal model structure (Sloman et al., 2009), and mental models of possibilities (Goldvarg & Johnson-Laird, 2001), to explain the differences in meaning between causal relations (for a review, see Khemlani, Barbey, & Johnson-Laird, 2014). But there exists no account of how causal relations are inferred from simulations, and as such no one has specified an algorithm that can carry out the task.

Our goal in the present article is to report studies that should help to solve this problem. We begin by illustrating how causal relations can be inferred from kinematic models by introducing the general tenets of mental model theory. We used a domain designed to elicit kinematic mental simulations, and so we introduce that domain and describe its characteristics. We then describe two studies in which participants' inferences about causal relations depended on the number of discrete steps in mental simulations. Finally, we evaluate the results in the context of current theories of causal reasoning.

Models of possibilities

The mental model theory – the "model" theory, for short – applies to reasoning across many domains, including reasoning based on sentential connectives, such as *if*, *or*, and

and (Johnson-Laird & Byrne, 1991), reasoning based on quantifiers (Khemlani, Lotstein, Trafton, & Johnson-Laird, 2015) and reasoning about temporal, spatial, causal, and abstract relations (Goodwin & Johnson-Laird, 2005). Three main principles underlie the theory (Johnson-Laird, 2006). First, mental models represent discrete *possibilities*: each model captures a distinct set of possibilities. Second, mental models are iconic as far as that is possible: the structure of the model corresponds to the structure of what it represents (see Peirce, 1931-1958, Vol. 4), and so kinematic models that unfold in time can represent a temporal sequence of events (Johnson-Laird, 1983; Khemlani et al., 2013). But, models can also include abstract symbols, e.g., the symbol for negation (Khemlani et al., 2012). Third, the model theory posits a principle of "truth": mental models represent only what is true and not what is false.

Inferences that require more models are more difficult than those that require fewer models. As a consequence, reasoners take longer to draw such inferences and are more likely to err, particularly by overlooking possibilities that render a given statement false. As such, they often represent only possibilities that render a statement true – their *mental* models – though they can flesh out those mental models to include additional possibilities to build *fully explicit* models. We illustrate the difference for the case of causal reasoning.

The model theory resolves the differences in interpretation between causal relations by distinguishing the sets of possibilities to which those relations refer (Goldvarg & Johnson-Laird, 2001), i.e., fully explicit models. A causal assertion such as (1) above refers to a conjunction of three separate models of possibilities, depicted in this schematic diagram:

```
  button   missile-launch
¬ button   missile-launch
¬ button ¬ missile-launch
```

Each row in the diagram represents a different temporally ordered possibility, e.g., the first row represents the possibility in which the button is pushed and the missile launches. In other words, the model theory posits that causality rules out those situations in which the button is pushed and the missile doesn't launch, as well as those situations in which the missile launch precedes the button push. In contrast, an enabling assertion, such as the one specified in (2), refers to a different conjunction of possibilities:

```
  button   missile-launch
  button ¬ missile-launch
¬ button ¬ missile-launch
```

i.e., to say that pushing the button enabled the missile to launch is to assert that the missile may or may not launch (the first two possibilities above). The enabling condition is inconsistent with the possibility in which the missile launches without the button being pushed. Reasoners list these possibilities for assertions such as (1) and (2) (Goldvarg & Johnson-Laird, 2001, Experiment 1). However, unless otherwise prompted to do so, reasoners build models in accordance with the principle of truth. As such, causes and enabling conditions have a single mental model representing one possibility:

```
button   missile-launch
```

Hence, individuals often fail to distinguish enabling from causing (Goldvarg & Johnson-Laird, 2001, Experiment 5).

When individuals observe or envisage a sequence of events, such as a button being pressed before a missile is launched, they can infer a causal relation between them. A correct inference depends on not only observing the factual mental model of the relation (e.g., the button being pressed, then the missile launching) but also envisaging the counterfactual possibilities to which the relation refers (i.e., the set of fully explicit models). If reasoners can envisage possibilities that correspond to the first conjunctive set above (e.g., the missile launching after the button press, and either launching or not in the absence of the button press) then they should infer that pressing the button *caused* the missile-launch. In contrast, inferring an enabling condition should be more complex. The difference in difficulty depends on the assumption that at least one of the following causal relations holds between the button and the missile launch: causes, enables, or prevents. Reasoners observe what happens given that the button is pressed. If the missile launches in this case, then they are likely to infer that *the button causes the missile launch*, and they will be correct if indeed it does. They will even be correct if they also consider the counterfactuals of what happens given that the button is not pressed. The only way that they are likely to infer that *the button enables the missile launch* is if they can envisage a counterfactual possibility in which the missile does not launch when the button is pressed. The asymmetry arises because the button press and the missile launch are both possible given either *causes* or *enables*, but the button press and missile not launching is possible only with *enables*. Hence, the theory predicts that when reasoners draw causal conclusions from mental simulations, it should be more difficult to infer enabling conditions, e.g., that pressing the button *enabled* the missile launch.

When individuals need to simulate a sequence of events in a kinematic model, the theory makes a direct prediction: the more events that occur in the sequence, the harder the inference should be – it should take longer and be more likely to yield an error.

In order to test the theory's predictions we adapted an experimental domain used to elicit kinematic mental simulations in order to study causal inferences. We describe that experimental domain in the next section.

Kinematic models in a railway domain

We sought to investigate how individuals without any formal training in logic, probability, or causal reasoning were able to infer causal relations by carrying out mental simulations. Accordingly, we developed a domain based on the railway environment shown in Figure 1. The environment is composed of a railway track and a siding, and recent studies demonstrate its ability to elicit kinematic simulations underlying deductive and abductive inferences (Khemlani et al., 2013). The environment is simple enough for children to understand and to reason about (Bucciarelli et al., under review). Cars are prohibited from moving from the siding directly to the right track and vice versa, and they are prohibited from moving from the right track back to the left track. In other words, there are only three legal moves in the environment: i) a move from the left track to the siding, ii) a move from the siding to the left track, and iii) a move from the left track to the right track. Multiple cars can be moved at once such that any move of a selected car applies to all cars in front of it. In Figure 1, if you moved the D car to the right track, then the E car would move along in front of it. Because both the siding and the left track function as stacks in an automaton, the environment in principle has the computational power of a universal Turing machine. To restrict the environment to a single stack, cars could move from the siding only to the output on the right track.

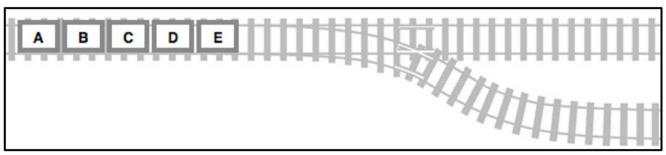

Figure 1. The railway domain with an example of an initial configuration in which a set of cars is on the left side of the track, the siding can hold one or more cars while other cars are moved to the right side of the track.

Consider the problem of reversing the order of five cars, ABCDE, i.e., to produce the sequence, EDCBA on the right track. The environment can be depicted in this diagram:

```
ABCDE[    ]
```

where the brackets denote the siding, the area to the left of the brackets denotes the left track, and the area to the right of the brackets denotes the right track. This sort of notation is used in a computer program that solves such problems and infers algorithms for solving them (Khemlani et al., 2013). To reverse the order of the train, a reasoner can move the cars as follows:

```
A[BCDE]      Move all but A to the siding.
 [BCDE]A     Move A to the left track.
B[CDE ]A     Move B to the left track…
 [CDE ]BA    …then to the right track.
C[DE  ]BA    Repeat for all the remaining cars.
 [DE  ]CBA
D[E   ]CBA
 [E   ]DCBA
E[    ]DCBA
 [    ]EDCBA
```

Reasoners can mentally simulate such a sequence of moves. Indeed, a reasoner who carried out the steps above to reverse the train ABCDE may have made causal inferences in passing. Consider the first move in the sequence:

```
ABCDE[    ]
    A[BCDE]
```

those who envision moving car B from the left track to the siding might recognize that doing so *caused* cars C, D, and E to move along with it. The inference can be drawn because the causation relation refers to three separate possibilities in a situation, such as Figure 1:

```
  B moves-to siding    C moves-to siding
¬ B moves-to siding    C moves-to siding
¬ B moves-to siding  ¬ C moves-to siding
```

In reversing the order of the cars, individuals can simulate the first possibility. They might also simulate counterfactual possibilities, e.g., if B hadn't moved to the siding, C may or may not have moved to the siding.

Reasoners can infer enabling conditions in a similar fashion. For example, reasoners can infer that moving B to the siding *enabled* A to move alone to the right track on a subsequent move. How might reasoners infer the relation? They may envisage that if B moves to the siding then A can move alone to the right track:

```
  B moves-to siding    A moves-to right track
```

Some reasoners may therefore infer that B's move causes A's move. But, the inference is erroneous. To make the correct inference, reasoners need to consider two counterfactual moves:

```
¬ B moves-to siding    A moves-to right track
¬ B moves-to siding  ¬ A moves-to right track
```

In other words, A could have remained on the left track even if B had moved to the siding; but, A could not have moved alone to the right track if B hadn't moved to the siding. The three models together suffice to infer the enabling relation: moving B to the siding enabled A to move to the right-track alone. But, the inference should be more difficult than inferring that A's move to the right track caused B to move there too.

We carried out two studies in which reasoners made such inferences about the railway environment. The model theory makes two main predictions about errors and latencies:

1. Inferences that one event *causes* another should be easier than inferences that one event *enables* another for the reasons we explain above.
2. The number of moves required to carry out the simulation should predict the difficulty of an inference.

Experiment 1

Experiment 1 aimed to test the two predictions, and thereby to corroborate the model theory's account of causal meanings and the role of kinematic models. On each trial, participants saw a picture of three cars on the railway track, such as one corresponding to the situation:

```
ABC[    ]
```

They then had to understand a supposition, such as: *Suppose B has just moved to the empty siding.* In this case, they have to simulate a single move: A[BC]. Finally, they had to answer a question, such as: *Did that move cause C to move to the siding?* In this case, the theory predicts that they should respond: Yes. The experiment manipulated the number of moves required in the simulation in order to respond to the question: 0, 1, 2, or 3, whether the question referred to "cause" or "enable", and whether the predicted answer should be "yes" or "no".

Method

Participants. Thirty-six students at the University of Pennsylvania completed the experiment for partial course credit. All of the participants were native English speakers, and none had had any prior training in formal logic.

Design. Participants acted as their own controls and carried out 16 problems in a fully repeated measures design, which manipulated the number of moves in a simulation (4), the causal or enabling relation (2), and the correct answer (2). The 16 problems in the study are in the Appendix. The study measured the accuracy of participants' responses to the questions and their latencies.

Procedure. The instructions explained the railway environment, and that moving a car moved all of the cars in front of it too. As a screening procedure, participants had to manipulate cars in the environment to reverse the order of a five-car train. All the participants passed this test. They were told that all of the descriptions of moves that they would receive were a result of making the fewest moves possible. They were also told (bold text in the original):

> "When we ask about whether one move *causes* another, we are concerned with whether the first move **makes the second move occur**. When we ask about whether one move *enables* a second, we are concerned with whether that second move **immediately becomes possible as a result of the first.**"

On each trial in the experiment, participants saw an image of the empty railway track. After a 1000 ms delay, cars in a specified arrangement appeared on the track. After another 2000 ms delay, a premise and a question appeared in the center of the screen, below the railway track, together with two buttons marked "Yes" and "No", which participants clicked with the mouse to record their responses. The experiment recorded the latency from the appearance of the premise and question to when participants clicked one of the buttons.

Results and discussion

The participants' inferences were more accurate for causal relations than for enabling relations (66% vs. 59%, Wilcoxon test, z = 1.74, p = .04, one-tailed, Cliff's δ = .07), which corroborated the model theory's first prediction. Figure 2 presents the proportion of correct responses (left panel) and the Winsorized latencies of all of the participants' responses, both correct and incorrect (right-panel) as a function of the number of moves in the simulation. Accuracies did not differ as a function of the number of moves in simulation (Page's trend test, z = .03, p = .98). But, latencies yielded a significant trend depending on the number of moves (Page's trend test, z = 3.98, p < .0007). The effect was more pronounced when isolating only correct responses (Page's trend test, z = 4.10, p < .0001). This result corroborated the model theory's second prediction.

An immediate concern is whether participants were in fact simulating moves in the environment, or whether the significant trend in latency is attributable to the number of words in each of the problems. To address the issue, we conducted a linear regression analysis on log-transformed latencies that included the number of words and the moves required as predictors. Both were significant predictors, but number of words had a lower regression coefficient (B = .09, p = .02) than moves required (B = .13, p = .007). We likewise conducted a hierarchical analysis in which two regression models were contrasted against one another: Model 1 regressed the latencies on the number of words alone, while Model 2 regressed latencies on the number of words in the problem as well as the number of moves required. An analysis of deviance showed a significant increase in model fit from Model 1 to Model 2 (Model 1 R^2 = .46 vs. Model 2 R^2 = .67, F = 10.23, p = .007).

The results corroborated the model theory's prediction that causes should be easier to envisage from simulation than enabling conditions. The results also corroborated the

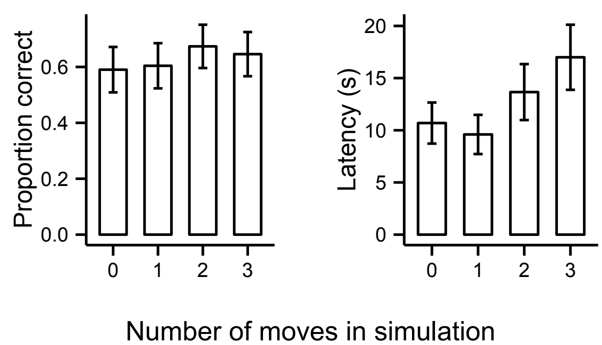

Figure 2. The proportion of correct responses (left-panel) and the response latencies from all responses (right-panel) in Experiment 1 as a function of the number of moves required to carry out a simulation.

prediction that latencies should correlate with the number of steps in a simulation needed to infer a causal relation. Both of the results suggest that participants extracted causal and enabling relations by carrying out and inspecting mental simulations. As the Figure shows, problems calling for zero moves in a simulation were not reliably easier than those calling for one move. A simple explanation is that individuals nevertheless envisaged the move described in the supposition even though its effects were depicted in the picture of the railway.

One shortcoming of the present experiment was that the tenses in the verb phrase of the stimuli were held constant across the problems. This constancy eliminated some confounds, but introduced an oddity: for problems that require at least one or more simulations, the premise used the past tense to refer to moves that had yet to occur. For example, consider the first 1-move problem. Participants saw the a track corresponding to:

```
ABC[    ]
```

and they were told:

Suppose B moved to the empty siding.
Did that move cause C to move to the siding? [Y/N]

The premise uses the past tense of "move" to refer to a situation to be simulated. A more natural formulation would use premises and questions that fit the temporal constraints of the simulation, e.g.,

Suppose B moves to the empty siding.
Does that move cause C to move to the siding?

Experiment 2 accordingly replicated Experiment 1 using the more natural present tense when appropriate.

Experiment 2

Experiment 2 used the same task and design as the previous study, however it used slightly modified materials. That is, the descriptions used the present tense to refer to arrangements in the environment that reasoners had to simulate.

Method

Participants. Twenty-one participants were recruited from the same subject pool as in the previous study, and they completed the study for course credit. All of them were native English speakers; all of them passed the screening described in Experiment 1; and none of them had received any training in formal logic.

Design and procedure. Same as Experiment 1.

Materials. Modifications to the materials are shown in the Appendix.

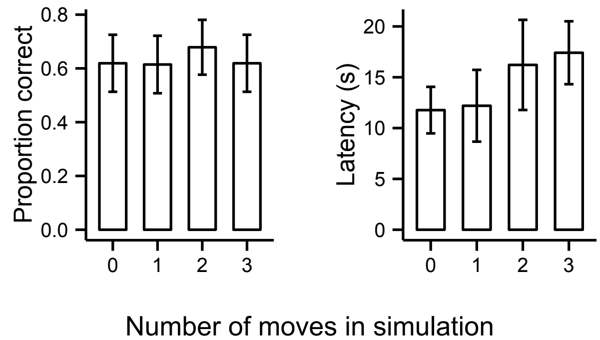

Figure 3. The proportion of correct responses (left-panel) and the response latencies (right-panel) in Experiment 2 as a function of the number of moves required to carry out a simulation.

Results and discussion

Experiment 2 measured participants' accuracies and latencies. Their responses were again more accurate for causal relations than for enabling conditions (69% vs. 57%, Wilcoxon test, $z = 1.71$, $p = .04$, one-tailed, Cliff's $\delta = .12$). Figure 3 presents the proportion of correct responses (left panel) and the Winsorized latencies of all of the participants' responses (right-panel) depending on the number of moves in the simulation. As in Experiment 1, accuracies did not reflect the number of moves in a simulation (Page's trend test, $z = .64$, $p = .52$), but latencies did for all responses and correct responses (Page's trend tests, $z > 2.95$, $p < .003$).

General discussion

Experiments 1 and 2 corroborated the model theory's two predictions: causes were easier to infer than enabling conditions, and the number of moves in a mental simulation predicted response latencies. The distinction in the meanings of "causes" and "enables" depends on a semantics for causal relations capable of building discrete representations, i.e., one that is deterministic (Frosch & Johnson-Laird, 2011; Khemlani et al., 2014). This distinction cannot be captured in a probabilistic account (pace, e.g., Suppes, 1970; Cheng & Novick, 1991).

Previous studies have suggested that mental simulation underlies reasoning about mechanical systems (Hegarty, 2004), about instabilities in physical systems (Battaglia et al., 2013), and about the consequences of algorithms (Khemlani et al., 2013). The present studies also demonstrate for the first time that the number of discrete steps in a simulation has a direct effect on the time that it takes individuals to make inferences. This result corroborates the use of kinematic mental models in reasoning.

Acknowledgements

We thank Paul Bello, Monica Bucciarelli, Ruth Byrne, Philipp Koralus, David Lobina, Robert Mackiewicz, Salvador Mascarenhas, Adam Moore, Alex Doumas, and Greg Trafton for advice. This work was supported by a Jerome and Isabella Karle Fellowship from the Naval Research Laboratory (to S.S.K.) and by National Science Foundation Grant SES 0844851 (to P.N.J-L.) to study deductive and probabilistic reasoning.

References

Ahn, W., & Bailenson, J. (1996). Causal attribution as a search for underlying mechanism: An explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology*, 31, 82−123.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327-18332.

Bramley, N. R., Gerstenberg T. & Lagnado, D. A. (2014). The order of things: Inferring causal structure from temporal patterns. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Bucciarelli, M., Mackiewicz, R., Khemlani, S., & Johnson-Laird, P. N. (under review). Simulations and gestures in children's formulations of algorithms. Manuscript under review.

Byrne, R.M.J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT.

Cheng, P. W. (1997). From covariation to causation: a causal power theory. *Psychological Review, 104,* 367–405.

Cheng, P. W., & Novick, L. R. (1991). Causes versus enabling conditions. *Cognition* 40, 83–120.

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

Galinsky, A., & Moskowitz, G. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology, 34,* 384-409.

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.

Hegarty, M. (2004). Mechanical reasoning as mental simulation. *Trends in Cognitive Sciences*, *8*, 280-285.

Hilton, D.J., & Erb, H-P. (1996). Mental models and causal explanation: Judgements of probable cause and explanatory relevance. *Thinking & Reasoning*, *2*, 273−308.

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

Johnson-Laird, P.N. (2006). *How we reason*. New York: Oxford University Press.

Johnson-Laird, P. N., & Byrne, R.M.J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Kahneman, D., & Tversky, A. (1998). The simulation heuristic. In (Eds.) D. Kahneman, P. Slovic, A. Tversky. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.

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

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. (2013). Cognitive changes from explanations. *Journal of Cognitive Psychology, 25,* 139–146.

Khemlani, S., Lotstein, M., Trafton, J.G., & Johnson-Laird, P. N. (2015). Immediate inferences from quantified assertions. *Quarterly Journal of Experimental Psychology.*

Khemlani, S., Mackiewicz, R., Bucciarelli, M., & Johnson-Laird, P. N. (2013). Kinematic mental simulations in abduction and deduction. *Proceedings of the National Academy of Sciences*, 110, 16766–16771.

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.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(3), 451–60.

Park, J., & Sloman, S. (2014). Causal explanation in the face of contradiction. *Memory & Cognition, 42,* 806-820.

Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed). NY: Cambridge University Press.

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.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: observations and interventions. Cognitive Psychology, 64(1), 93–125.

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.

Suppes, P. (1970). *A probabilistic theory of causality*, Amsterdam: North-Holland Publishing.

White, P.A. (2014). Singular cues to causality and their use in human causal judgment. *Cognitive Science, 38,* 38–75.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, *136*, 82−111.

Appendix. The 16 problems used in Experiments 1 and 2. Where relevant, changes to the stimuli between Experiment 1 and Experiment 2 are marked in bolded text.

# of moves	Initial conf.	Premise	Question	Causal reln.	Correct answer
0	A[BC]	Suppose B has just moved to the empty siding.	Did that move cause C to move to the siding?	Cause	Yes
0	A[BC]	Suppose B has just moved to the empty siding.	Did that move enable C to move to the siding?	Enable	No
0	A[BC]	Suppose B has just moved to the empty siding.	Did that move cause A to stay on the left track?	Cause	No
0	A[BC]	Suppose B has just moved to the empty siding.	Did that move enable A to stay on the left track?	Enable	Yes
1	ABC[]	Suppose B moved/**moves** to the empty siding.	Did/**Does** that move cause C to move to the siding?	Cause	Yes
1	ABC[]	Suppose B moved/**moves** to the empty siding.	Did/**Does** that move enable C to move to the siding?	Enable	No
1	ABC[]	Suppose B moved/**moves** to the empty siding.	Did/**Does** that move cause A to stay on the left track?	Cause	No
1	ABC[]	Suppose B moved/**moves** to the empty siding.	Did/**Does** that move enable A to stay on the left track?	Enable	Yes
2	ABC[]	Suppose A moved/**moves** to be alone on the right track.	Did/**With that move, does** B cause C to move to the siding?	Cause	Yes
2	ABC[]	Suppose A moved/**moves** to be alone on the right track.	Did/**With that move, does** B enable C to move to the siding?	Enable	No
2	ABC[]	Suppose A moved/**moves** to be alone on the right track.	Did/**Does** that move cause B to move to the left track?	Cause	No
2	ABC[]	Suppose A moved/**moves** to be alone on the right track.	Did/**Does** that move enable B to move to the left track?	Enable	Yes
3	[ABC]	Suppose C moved/**moves** to be alone on the left track.	Did/**With that move, does** A cause B to move to the right track?	Cause	Yes
3	[ABC]	Suppose C moved/**moves** to be alone on the left track.	Did/**With that move, does** A enable B to move to the right track?	Enable	No
3	[ABC]	Suppose C moved/**moves** to be alone on the left track.	Did/**Does** that move cause C to move to the right track?	Cause	No
3	[ABC]	Suppose C moved/**moves** to be alone on the left track.	Did/**Does** that move enable C to move to the right track?	Enable	Yes