Title
Temporal explanations help resolve temporal conflicts

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Kelly, Laura J
Khemlani, Sangeet

Publication Date
2021

Peer reviewed

Temporal explanations help resolve temporal conflicts

Laura Kelly and Sangeet Khemlani
{laura.kelly.ctr, sangeet.khemlani}@nrl.navy.mil
Navy Center for Applied Research in Artificial Intelligence
US Naval Research Laboratory, Washington, DC 20375 USA

Abstract

People can explain phenomena by appealing to temporal relations, e.g., you might explain a colleague's absence at a meeting by inferring that their prior meeting did not end on time. Cognitive scientists have yet to investigate temporal explanations, and explanatory reasoning research tends to focus on how people assess causal explanations; it shows that reasoners often generate causal explanations to resolve conflicts. We posit that temporal explanations help reasoners resolve temporal conflicts, and describe three experiments that test the hypothesis. Experiment 1 provided participants with temporal information that was consistent or inconsistent and elicited their inferences about what followed. Participants spontaneously provided temporal explanations to resolve inconsistencies, and many of them also provided more conservative refutations. In Experiments 2 and 3, participants evaluated explanations and refutations in light of conflicting information. The studies showed that participants spontaneously generate temporal explanations, and in certain cases, they prefer temporal explanations when a more conservative refutation was available. The research is the first to examine patterns in temporal explanatory reasoning.

Keywords: temporal explanations; conflicts; inconsistencies; duration; event cognition

Introduction

These three sentences are inconsistent:

The football game went from 1pm to 5pm.
Ria arrived at the game at 6pm.
She attended the game.

The situation is impossible: how can Ria attend the game if she arrived after it occurred? One of the three premises must be false, i.e., they cannot all be true at the same time. Psychologists since William James have argued that people cope with conflicts by minimally revising their information, i.e., they reject as few of the premises as possible. But recent work shows that reasoners generate explanations to resolve conflicts (Khemlani & Johnson-Laird, 2011, 2012). Here are some plausible explanations that might suffice:

Ria attended the game virtually.	[spatial]
The posted schedule was wrong.	[epistemic]
Ria is the Queen; she shifted the schedule.	[causal]
Ria attended a different game.	[temporal]
The game was delayed.	[temporal]

Explanations help reasoners understand the past and predict the future (Anderson & Ross, 1980; Craik, 1943; Einhorn & Hogarth, 1986; Gopnik, 2000; Lombrozo & Carey, 2006; Ross, Lepper, Strack, & Steinmetz, 1977), and many cognitive scientists argue that they are a hallmark of human rationality (Harman, 1965; Horne, Muradoglu, & Cimpian, 2019; Johnson-Laird, 1983; Lombrozo, 2007), though they also serve as the basis of magical thinking, conspiracy theory, and pseudoscience (Gronchi, Zemla, & Brondi, 2017; Weisberg, Keil, Goodstein, & Rawson, 2008). They add new relations or entities into the discourse that were not present in the given information (e.g., a virtual game, the posted schedule, Ria's royal status). The consequences of a particular explanation can help guide reasoners to reject information, e.g., if a reasoner infers that Ria attended the game virtually, then they may tacitly reject the second premise that she arrived at the game at 6pm. Many types of explanation can help reasoners resolve the conflict, e.g., a spatial explanation appeals to spatial locations, and an epistemic explanation appeals to knowledge and belief.

The majority of empirical research into explanatory reasoning has focused on how people assess causal explanations (e.g., Ahn & Kalish, 2000; Alicke, Mandel, Hilton, Gerstenberg, & Lagnado, 2015; Fernbach, Macris, & Sobel, 2012; Johnson-Laird, Girotto, & Legrenzi, 2004; Lombrozo, 2016; Sloman, 2005). One reason for the focus on causality may be because people tend to spontaneously generate causal explanations when given the opportunity in the laboratory (Khemlani & Johnson-Laird, 2011) and in more natural contexts (Zemla, Sloman, & Bechlivanidis, 2017). In particular, reasoners are more likely to generate causal explanations to resolve conflicting, inconsistent information rather than to elaborate on a consistent description (Khemlani & Johnson-Laird, 2011, 2012, 2013; Legare, 2012). Consider this description from Khemlani & Johnson-Laird (2013):

If a person does regular aerobic exercises then the person strengthens her heart.
Someone did regular aerobic exercises, but she *[did / did not]* strengthen her heart.
What, if anything, follows?

If the woman strengthened her heart, there is nothing to explain, and reasoners often respond that "nothing follows". If she did not strengthen her heart, however, the two premises are inconsistent with one another, i.e., reasoners can draw contradictory conclusions from them. They often infer explanations to eliminate the conflict, e.g.,

Perhaps she has a health condition that prevents her heart from getting stronger.

Since *prevention* is a causal relation, the explanation is causal in nature, and reasoners rely on causal knowledge to resolve other kinds of inconsistencies, too (Khemlani & Johnson-Laird, 2020). They need not have constructed an explanation: a more conservative response would have been to directly refute the premises, e.g., they could have inferred that the first premise is strictly false (it describes a generalization that has exceptions). But, as studies show, people prefer causal explanations to refutations and generate them more often.

In the example above, the first premise describes a causal relation: regular aerobic exercises cause a person's heart to strengthen. So, in retrospect, it may not be surprising that reasoners infer causal explanations to resolve causal conflicts. But conflicts can arise in other sorts of information, too, e.g., temporal descriptions can contain conflicts. Reasoners can make sophisticated inferences about time and duration (Hoerl & McCormack, 2019), and they can detect conflicts in temporal descriptions. Consider this description (from Kelly, Khemlani, & Johnson-Laird, 2020):

The meeting happened during the conference.
The sale happened before the conference.
The meeting happened before the sale.

The three sentences cannot all be true at the same time, and reasoners have little difficulty detecting the inconsistency – indeed, they often inaccurately assess descriptions to be inconsistent, because they fail to consider all the possibilities consistent with the sentences.

When people detect an inconsistency in temporal information, they may try to explain it by introducing new events and temporal relations to the discourse, i.e., they may try to construct temporal explanations. Cognitive scientists have yet to examine temporal explanations in humans, and no studies have assessed whether people make them in response to temporal conflicts. This paper accordingly examines how reasoners construct and evaluate explanations of time as a way to cope with conflicts. Three studies test the hypothesis that conflicts should prompt reasoners to generate temporal explanations and consider them as sensible more often than more conservative refutations. Experiment 1 shows that people produce temporal explanations to resolve conflicting information; Experiment 2 reveals cases in which people prefer temporal explanations to refutations, and Experiment 3 controls for a confound in Experiment 2 and further shows that participants often prefer explanations. We conclude by describing how temporal explanations differ from other kinds of explanation and why they are a particularly helpful strategy for resolving conflicts in information.

Experiment 1

Experiment 1 sought to test whether reasoners can generate temporal explanations in a systematic way. It gathered reasoners' natural responses to conflicting temporal information. Participants typed out their responses to problems such as:

Suppose that you are told the following:
The blood drive was open from 9am to 4pm on Monday.
Trisha arrived at the blood drive at 5pm on Monday.
You discover the following fact:
Trisha gave blood at the blood drive.
What, if anything, follows?

The set of premises is inconsistent, because they describe a scenario in which Trisha gave blood at a blood drive after it had closed. The study varied whether the premises described consistent or inconsistent scenarios; previous work on causal explanations suggests that reasoners should generate explanations more often for inconsistent scenarios.

Method

Participants. 51 participants completed the experiment for monetary compensation ($2) through Amazon's Mechanical Turk. 6 participants produced a majority of nonsensical responses, so we dropped their data. The analyses reported are based on the remaining 45 participants (18 female, mean age = 37.2). The participants were native English speakers, and 6 had taken one or more courses in logic.

Preregistration and data-availability. The experimental designs, predictions, and analyses for Experiments 1-3 were pre-registered through the Open Science Framework platform (https://osf.io/v23ah/). The same link provides the corresponding experimental code, materials, and data.

Task, materials, and design. Participants completed 8 problems which each presented a participant with information concerning the duration of an event, information about when an individual arrived to the event, and information about whether or not that individual took part in the event. Provided that an individual can take part in an event only if they arrive sometime between when the event started and when it ended, the premises in each problem could conflict with one another. For instance, the set of premises provided in the example above is inconsistent, because Trisha arrived after the drive ended – but a change to the first premise makes the set of premises consistent:

The blood drive was open from 9am to **8pm** on Monday.

In this description, Trisha arrived during the event hours and therefore it is consistent to say she gave blood at the event. Half of the problems were consistent and half were inconsistent; the experiment randomized the consistency of each problem. Participants' typed their response to the question, "What, if anything, follows?" into a response box. The experiment required participants to type a response at least 1 character long for each problem.

The premises of the problems in Experiment 1 came from 8 separate scenarios that concerned everyday events (e.g., attending a class, speaking at a meeting, picking up a prescription), and each set of premises could be made consistent or inconsistent by manipulating the interval of the event described. There was an error in one scenario such that

Table 1. The types of responses produced by participants in Experiment 1; the percentages of those responses; examples of each response type; and the percentages of responses that unambiguously refuted or explained one of the three premises, along with relevant examples.

Type of response	%	Example
Direct refutation of premises:	38%	"Kiana did not pick up her medication."
Temporal explanations:	37%	"The staff meeting was postponed."
Other explanations:	5%	
Causal:	<1%	"Kiana broke into the pharmacy after it closed."
Epistemic:	<2%	"The hours given for the doctor's office were inaccurate."
Spatial:	<2%	"Ria attended the meeting remotely."
Miscellaneous:	<1%	*[omitted for brevity]*
Premise that was either refuted or explained:		
Premise 1:	32%	"The blood drive location decided to stay open later."
Premise 2:	5%	"Ria got to work early before … the meeting."
Premise 3:	10%	"Ria did not make the meeting, she was too late."

Note: Percentages describing responses were orthogonal to one another and do not add to 100%.

the timeline was incoherent if taken literally, i.e., "The party was scheduled to occur from 7pm to 12am on Friday." The analysis below focuses on the remaining 7; excluding the erroneous scenario had no qualitative effect on the results.

Rejection and coding criteria. The first author coded participants' typed responses. Responses that appeared nonsensical, copied from the premises, or otherwise inappropriate to the task were dropped from further analysis (14% of the data from the 51 original participants; 2% of the data from the 45 participants included in the analyses). The subsequent analyses concerned the remaining 309 responses (see Table 1 for examples). If participants generated more than one plausible response (this occurred for 6% of the trials), we coded only their first response.

We coded responses on the following four criteria:

1. *Did the response directly refute one of the premises?* Each response was coded on whether it explicitly denied the truth of one of the premises (e.g., "The blood drive didn't close at 4pm, it was still open at 5pm.").
2. *Did the response explain the premises by reference to some temporal concept?* Temporal explanations are responses that introduce a new temporal relation, e.g., "The class *was pushed back* that day", or a new event, e.g., "She *left* after the meeting *and then came back*."
3. *Did the response explain the premises in some other way?* Other explanations concerned responses that introduce non-temporal entities or relations, such as spatial, epistemic, or causal relations, or else relations that were ambiguous in nature.
4. *Which premise did a response refute or explain*? Refutations or explanations could concern: the event's time interval (premise 1); the time an individual arrived at the event (premise 2); or whether the individual attended the event (premise 3). Responses that were equivocal were not considered for further analyses.

Results and Discussion

Participants produced temporal explanations on 37% of the trials and direct refutations on 38% of the trials; Table 1 provides a breakdown of the different types of responses. Participants produced more explanations for inconsistent problems than for consistent problems (75% vs. 4%; Wilcoxon test, $z = 5.78$, $p < .0001$, Cliff's $\delta = .87$), and 40 out of 45 participants exhibited this pattern (binomial test, $p < .0001$). The pattern is analogous to how individuals cope with conflicts in causal sets of premises (Khemlani & Johnson-Laird, 2011). Likewise, participants produced refutations more often for inconsistent versus consistent problems (67% vs. 4%; Wilcoxon test, $z = 5.79$, $p < .0001$, Cliff's $\delta = .89$), a pattern that 41 out of 45 participants displayed (binomial test, $p < .0001$). The pattern was robust to the different materials, as well; when aggregated by the 7 different scenarios, all 7 revealed more temporal explanations and more refutations for inconsistent problems than consistent problems (binomial tests, *p*s < .01).

Experiment 1 revealed that individuals could spontaneously construct temporal explanations. They did so on roughly 3 out of 4 inconsistent trials. Many of the participants' responses (29%) to inconsistent trials directly refuted the premises, e.g., "The meeting started later than 10am." We had not anticipated such responses, given that they seldom occur in the case of causal conflicts (Khemlani & Johnson-Laird, 2011, Experiment 1). We suspect that the ability to type open-ended responses, as well as ambiguities in the materials themselves, allowed participants to immediately infer the consequences of the explanations they generated. But, the data obscured the participants' preferred strategies in coping with temporal conflicts. Experiment 2 accordingly tested participants' relative preferences between explanations and refutations. It used a forced choice task to directly compare participants' preferences between the two.

Experiment 2

Experiment 2 tested whether participants prefer temporal explanations to direct refutations when coping with premises that describe a temporal inconsistency. On half of the trials, participants received problems and response options such as:

> Suppose that you are told the following:
> The concert was occurring from 9pm to 11pm.
> Ruthie arrived at the concert hall at 11:30pm.
> You discover the following fact:
> Ruthie attended the concert.
>
> What, if anything, follows?
> Ruthie did not attend the concert. [refutation]
> The concert was delayed by two hours. [explanation]
> Nothing follows from the given information.

On the remaining trials, participants compared the same response options for consistent problems, e.g., a problem akin to the one above except where the concert ended at midnight, after Ruthie arrived. Refutations are simpler than explanations because explanations introduce concepts not present in the premises, e.g., the explanation above introduces the concept of a delay while the refutation does not introduce anything. Hence, a conservative response may be to prefer refutations over explanations. But, as previous research on causal explanations shows, people often prefer explanations to refutations because explanations provide a more complete narrative of what gave rise to the inconsistency.

Method

Participants. 55 participants completed the experiment for monetary compensation ($2.50) through Amazon's Mechanical Turk, commensurate with minimum-wage standards. We dropped data from 5 participants who took less than 2 minutes to complete the task, gave nonsensical responses to the debriefing questions, or were non-native English speakers. Of the remaining 50 participants, all but 12 had taken one or fewer courses in introductory logic. Their mean age was 39.1; 17 participants were female, 32 were male, and one preferred not to say.

Task and design. As in Experiment 1, participants were presented with 8 problems, half of which were consistent and half of which were inconsistent. Each problem presented a forced choice task between 3 possible responses to the prompt, "What, if anything, follows?"

Materials. The 8 scenarios were based on participants' natural responses from Experiment 1. Some of the problems in Experiment 1 contained various ambiguities that permitted participants to construe the problems as consistent when they were designed to be inconsistent. The materials in Experiment 2 were modified so that they described unambiguous event intervals. For each scenario, the experiment provided three response options: a temporal explanation, a refutation, and "Nothing follows from the given information." The temporal explanation implied a change to the interval described in the first premise, e.g., an extension or a postponement. The refutation focused on the third premise by denying that the agent attended the event or that they carried out the action that required attending the event, e.g., "Yasmine did not speak at the meeting." The explanations and refutations were constructed to have the same number of syllables (see OSF for materials). There was an issue with the same scenario as in Experiment 1. The analysis we report was conducted on the other 7 scenarios; excluding the erroneous scenario did not qualitatively affect the results.

Table 2. The percentages of participants' selections of the three different response options in Experiment 2 as a function of whether the problem was consistent or inconsistent.

Type	Consistent	Inconsistent	All
Explanation	22%	82%	51%
Refutation	5%	10%	7%
Nothing follows	73%	8%	41%

Procedure. Each problem began by displaying the event information and the question. After a 3 second delay, the three response options appeared in a randomized order. Participants selected a response to move to the next problem.

Results and Discussion

Table 2 provides the percentages of participants' choices for explanations, refutations or 'nothing follows' responses in Experiment 2. Participants preferred temporal explanations over refutations (51% vs. 7%; Wilcoxon test, z = 5.76, $p < .0001$, Cliff's $\delta = .89$) and 'nothing follows' responses (51% vs. 41%; Wilcoxon test, $z = 2.29$, $p = .022$, Cliff's $\delta = .33$). Their pattern of responses depended on the consistency of the scenario, $X^2(2, N = 350) = 155.07$, $p < .0001$. The results validate the prediction that participants should prefer explanations over refutations in the inconsistent condition (82% vs. 10%; Wilcoxon test, $z = 5.49$, $p < .0001$, Cliff's $\delta = .85$). Participants chose explanations more frequently in the inconsistent condition than the consistent condition (82% vs. 23%; Wilcoxon test, $z = 8.89$, $p < .0001$, Cliff's $\delta = .79$). In the consistent condition, participants preferred 'nothing follows' responses over the other response options (73% vs. 27%; Wilcoxon test, $z = 4.45$, $p < .0001$, Cliff's $\delta = .71$).

Experiment 2 directly tested participants' preferences for refutations, explanations, and 'nothing follows' responses as answers to consistent and inconsistent problems. One limitation of the study is that it confounded the type of response with the premise under consideration. That is, explanations explained the first premise (the premise describing the time interval of the relevant event) and refutations refuted the third premise (the premise describing the agent's participation in the event). Experiment 3 addressed the confound by presenting explanations and refutations that both concerned the first premise.

Experiment 3

Experiment 3 was similar to Experiment 2: it provided participants with three response options to consider, i.e., an explanation, a refutation, and a 'nothing follows' response. The explanations and refutations both concerned the first premise. For example, one problem in the study included the following premises:

> Suppose that you are told the following:
> The pottery class was Thursday from 6:30pm to 8pm.
> Matteo arrived at the pottery studio at [7pm/9pm] on Thursday.
> You discover the following fact:
> Matteo attended the pottery class.

The response options were:

> The class did not end at 8pm on Thursday.
> The teacher was late delaying the start of the class.
> Nothing follows from the given information.

Method

Participants. 50 participants completed the experiment for monetary compensation ($2.50) through Amazon's Mechanical Turk. We dropped the data from 1 participant who took less than 2 minutes to complete the experiment. The analyses reported below are based on the remaining 49 participants (24 female, mean age = 36.4). All of the participants were native English speakers and 32 had taken one or fewer courses in introductory logic.

Task, design, and procedure. Same as in Experiment 2.

Materials. The refutations for this experiment stated that the event interval did not end at the time provided, e.g., "The blood drive did not end at 4pm on Monday" (see OSF for materials). Unlike in the two previous experiments, the first premise did not vary based on consistency. The experiment manipulated the arrival time described in the second premise to create consistent and inconsistent problems, e.g., "Viv arrived at the blood drive at [2pm / 5pm] on Monday." The length (i.e., number of syllables) between the refutations and explanations were roughly matched.

Results and Discussion

Table 3 provides the percentages of participants' choices for explanations, refutations and 'nothing follows' responses in Experiment 3. Participants did not prefer explanations over refutations or over 'nothing follows' responses; indeed, all three responses types did not differ reliably from chance (Wilcoxon tests, *z*s < 1.78, *p*s > .07). However, their response preferences varied as a function of the consistency of the scenario, $X^2(2, N = 392) = 120.06$, $p < .0001$. Participants preferred 'nothing follows' responses over other response types in the consistent problems (65% vs. 35%; Wilcoxon test, $z = 2.56$, $p = .01$, Cliff's $\delta = .41$), but they preferred

Table 3. The percentages of participants' selections of the three different response options in Experiment 3 as a function of whether the problem was consistent or inconsistent.

Type	Consistent	Inconsistent	All
Explanation	24%	37%	30%
Refutation	11%	50%	31%
Nothing follows	65%	13%	39%

explanations over 'nothing follows' responses in the inconsistent condition (37% vs. 13%, Wilcoxon test, $z = 3.44$, $p < .001$, Cliff's $\delta = .48$), and they likewise preferred refutations over 'nothing follows' responses in the inconsistent condition (50% vs. 13%; Wilcoxon test, $z = 4.71$, $p < .0001$, Cliff's $\delta = .68$).

Participants chose refutations more often than explanations in the inconsistent problems, but the difference was not reliable (50% vs. 37%; Wilcoxon test, $z = 1.73$, $p = .083$, Cliff's $\delta = .25$). Participants chose explanations frequently more often when there was an inconsistency to resolve (37% vs. 24%; Wilcoxon test, $z = 8.63$, $p < .0001$, Cliff's $\delta = .26$).

Experiment 3 tested participants' preferences among explanations, refutations, and 'nothing follows' responses for consistent and inconsistent scenarios when the information being explained and refuted were the same, i.e., the event interval. Participants generally chose either a refutation or an explanation over 'nothing follows' for inconsistent scenarios but they did not show a preference between the two.

When compared with Experiment 2, the present results suggest that the preference for explanations over refutations in Experiment 2 could be due to differences in the contents of the information being considered. It may also be a result of our problem format which presents the third premise, i.e., the agent's participation, as new information that could be have called into question the veracity of the previous information. Regardless, in the present experiment participants chose to resolve inconsistencies with explanations just slightly less often than they chose refutations. The refutation is simpler and doesn't add in new information yet it is not grossly preferred over the temporal explanations.

All three experiments, reveal new patterns of reasoning about temporal explanations. They show that: participants generate temporal explanations (Experiment 1); they can prefer temporal explanations over refutations to resolve inconsistencies (Experiment 2); and that they endorse explanations at a similar rate to refutations when possible confounds are eliminated from the study design.

General Discussion

We describe evidence that reasoners can spontaneously generate temporal explanations – i.e., explanations that introduce novel events and temporal relations – particularly when those explanations resolve conflicts in premises that describe temporal relations. A series of experiments presented participants with problems of the following form:

The party occurred from 7pm to 10pm on Friday.
Maryam arrived at the party at 10:30pm on Friday.
Maryam attended the party.
What, if anything, follows?

The premises explicitly concern temporal information, i.e., the durations of the events and a particular individual's arrival time, and people have no difficulty assessing the consistency of such descriptions (though they have difficulty when durational relations yield ambiguous mental simulations; see Kelly, Khemlani, & Johnson-Laird, 2020). For instance, the premises above are clearly inconsistent: they cannot all be true at the same time.

Participants in Experiment 1 typed out their natural responses to such problems. In theory, their responses needn't have appealed to temporal explanations: the premises imply other relations. For instance, if Maryam arrived at the party, it may be reasonable to induce that she *knew* that the party was happening (an epistemic relation). Maryam *arrived* at the party, which means that she hadn't been in the same spatial location as the party (a spatial relation). Arrival at the party demands some means of moving from one location to another (a causal relation). And so, in theory, participants could have appealed to any number of explanations to resolve the conflict, such as this epistemic explanation:

Maryam *mistakenly thought* the party was happening until midnight.

Instead, reasoners preferred to resolve the conflict by introducing novel temporal relations, as in this explanation:

Participant 43: "The party ran later than scheduled."

The relation *later* is temporal, and it helps to resolve the conflict by implicitly refuting the first premise in the description. Other temporal explanations are possible, e.g.,

Maryam went to the after-party, not the main party.

This explanation introduces a novel event (the after-party), which presumably occurs directly after the main party, and so it does not refute the first premise; it refutes the second.

Perhaps participants' tendency to generate temporal explanations was an artifact of the generative task in Experiment 1, i.e., there may be a cognitive burden associated with constructing an explanation from scratch (Horne et al., 2019). Experiments 2 and 3 accordingly provided participants with several options to resolve inconsistencies, including an explanation and a direct refutation. In Experiment 2, participants chose the explanation far more often than they chose the refutation, though the experiment was confounded such that the two options implicitly refuted different premises. In Experiment 3, participants chose the refutations more often to resolve conflicting information, but they nevertheless chose explanations far more often than the most conservative response option, i.e., that nothing follows from the premises. All three experiments show that temporal explanations help reasoners resolve inconsistencies in temporal information, and they serve as the first datasets on explanatory reasoning about time.

Yet all three experiments are limited, particularly the latter two: they presented participants with multiple options to consider, which demands that participants compare and weigh the options in tandem. Future studies should ask participants to assess the merits of each option in isolation to eliminate any noise that is a consequence of comparisons.

In sum, we report the discovery of a novel class of non-causal explanation: temporal explanations. Reasoners generate them and evaluate them systematically, and temporal explanations may help explain a trade-off in explanatory reasoning: explanations demand cognitive resources to produce and evaluate, but they yield a better and more coherent understanding of the world.

Acknowledgments

This research was performed while the first author held an NRC Research Associateship award at the U.S. Naval Research Laboratory. It was also supported by a grant from the Office of Naval Research to the second author. We are grateful to Kalyan Gupta and Danielle Paterno at the Knexus Research Corporation for their help in conducting the experiments. And we thank Bill Adams, Gordon Briggs, Monica Bucciarelli, Hillary Harner, Tony Harrison, Phil Johnson-Laird, Joanna Korman, Markus Knauff, Andrew Lovett, Greg Trafton, and Ben Wright for their advice and comments.

References

Ahn, W., & Kalish, C. W. (2000). The role of mechanism beliefs in causal reasoning. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition*. Cambridge, MA: MIT Press.

Alicke, M. D., Mandel, D. R., Hilton, D. J., Gerstenberg, T., & Lagnado, D. A. (2015). Causal conceptions in social explanation and moral evaluation: A historical tour. *Perspectives on Psychological Science*, *10*.

Anderson, C.A., & Ross, L. (1980). Perseverence of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*.

Craik, K. (1943). *The nature of explanation*. Cam- bridge, United Kingdom: Cambridge University Press.

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*.

Fernbach, P. M., Macris, D. M., & Sobel, D. M. (2012). Which one made it go? The emergence of diagnostic reasoning in preschoolers. *Cognitive Development*, *27*.

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory-formation system. In F. Keil, & R. A. Wilson (Eds.), *Explanation and cognition.* Cambridge, MA: MIT Press.

Gronchi, G., Zemla, J., & Brondi, M. (2017). Cognitive style predicts magical beliefs. *Proceedings Of The 39th Annual Conference of the Cognitive Science Society*.

Harman, G. H. (1965). The inference to the best explanation. *Philosophical Review*, *74*.

Hoerl, C., & McCormack, T. (2019). Thinking in and about time:

A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, *42*.

Horne, Z., Muradoglu, M., & Cimpian, A., (2019). Explanation as a cognitive process. *Trends in Cognitive Sciences*, *23*.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, *111*.

Kelly, L., Khemlani, S., & Johnson-Laird, P.N. (2020). Reasoning about durations. *Journal of Cognitive Neuroscience, 32*.

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

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

Khemlani, S., & Johnson-Laird, P.N. (2020). Causal conflicts produce domino effects. *Quarterly Journal of Experimental Psychology,* 73.

Legare, C. H. (2012). Exploring explanation: Explaining inconsistent information guides hypothesis-testing behavior in young children. *Child Development*, *83*.

Legrenzi, P., & Johnson-Laird, P. N. (2005). The evaluation of diagnostic explanations for inconsistencies. *Psychologica Belgica*, *45*, 19–28.

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

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*.

Lombrozo, T. & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*, 167–204.

Ross, L., Lepper, M. R., Strack, F., Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, *35*, 817–829.

Sloman, S. A. (2005). *Causal models: How we think about the world and its alternatives*. New York, NY: Oxford University Press.

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, *20*, 470-477.

Zemla, J. C., Sloman, S., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review*, *24*, 1488–1500.