Acta Psychologica 139 (2012) 486–491

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Hidden conflicts: Explanations make inconsistencies harder to detect☆

Sangeet S. Khemlani [a,*], P.N. Johnson-Laird [b]

[a] Naval Research Laboratory, United States
[b] Princeton University, United States

ARTICLE INFO

Article history:
Received 10 June 2011
Received in revised form 25 January 2012
Accepted 27 January 2012
Available online 22 February 2012

PsyINFO classification:
2340 Cognitive processes

Keywords:
Inconsistency
Explanations
Belief revision
Reasoning
Principle of resolution

ABSTRACT

A rational response to an inconsistent set of propositions is to revise it in a minimal way to restore consistency. A more important psychological goal is usually to create an explanation that resolves the inconsistency. We report five studies showing that once individuals have done so, they find inconsistencies harder to detect. Experiment 1 established the effect when participants explained inconsistencies, and Experiment 2 eliminated the possibility that the effect was a result of demand characteristics. Experiments 3a and 3b replicated the result, and showed that it did not occur in control groups that evaluated (or justified) which events in the pairs of assertions were more surprising. Experiment 4 replicated the previous findings, but the participants carried out all the conditions acting as their own controls. In all five studies, control conditions established that participants were able to detect comparable inconsistencies. Their explanations led them to reinterpret the generalizations as holding by default, and so they were less likely to treat the pairs of assertions as inconsistent. Explanations can accordingly undo the devastating consequences of logical inconsistencies, but at the cost of a subsequent failure to detect them.

1. Introduction

The word 'why' is used to elicit explanations for the mysteries of daily life. Why is my car making that noise? Why didn't the Redskins win last Sunday? Why isn't my experiment working? Indeed, a central feature of human rationality is the ability to construct explanations of observed phenomena (Harman, 1965). Recent research has explored the function and developmental trajectory of explanatory reasoning (Keil, 2006; Wellman, Hickling, & Schult, 1997). And there is consensus among researchers that explanations are related to causal inference (Johnson-Laird, Girotto, & Legrenzi, 2004; Sloman, 2005; Byrne, 2005; Walsh & Johnson-Laird, 2009), and that explanations affect reasoning, categorization, and learning (Lombrozo, 2006). Less is known about the contexts in which individuals create explanations, i.e., when and how they decide to produce explanations. One obvious context is when they are asked for an explanation. But, people also produce explanations when they are learning new information (Amsterlaw & Wellman, 2006; Chi, De Leeuw, Chiu, & Lavancher, 1994; Crowley & Siegler, 1999; Rittle-Johnson, 2006), trying to form categories (Shafto & Coley, 2003), and judging how well concepts cohere with one another (Murphy & Medin, 1985; Patalano, Chin-Parker, & Ross, 2006). Explanations also help individuals to predict future behaviors (Anderson & Ross, 1980; Einhorn & Hogarth, 1986; Lombrozo & Carey, 2006; Ross, Lepper, Strack, & Steinmetz, 1977).

Another context in which individuals spontaneously create explanations is when they detect inconsistencies (Johnson-Laird, 2006; Johnson-Laird et al., 2004; Khemlani & Johnson-Laird, 2011). Even children are likely to generate causal explanations if they observe an inconsistency with their previous experience in an experiment (Legare, 2012; Legare, Gelman, & Wellman, 2010). The relation between inconsistency and explanation is the topic of the present paper, and, in particular, how explanations can in turn make inconsistencies harder to detect.

2. Explanations resolve inconsistencies

When individuals detect an inconsistency among a set of assertions, they try to explain the origins of the inconsistency. If they know what created the inconsistency then they can make a better decision about an appropriate course of action. The explanation, of course, has a side effect of restoring consistency to the set of propositions (Johnson-Laird et al., 2004; Khemlani & Johnson-Laird, 2011). According to this *principle of resolution*, they then re-interpret the inconsistent assertions based on the consequences of their explanation.

☆ This research was supported by a National Science Foundation Graduate Research Fellowship to the first author, and by National Science Foundation Grant no. SES 0844851 to the second author to study deductive and probabilistic reasoning. For their helpful comments and criticisms, we thank Jeremy Boyd, John Darley, Sam Glucksberg, Adele Goldberg, Geoffrey Goodwin, Matt Johnson, Olivia Kang, Niklas Kunze, Max Lotstein, Joe Magliano, Gorka Navarrete, Ray Nickerson, Danny Oppenheimer, Laura Suttle, Rolf Zwaan, and the audience at the 32nd Annual Conference of the Cognitive Science Society, where an earlier version of this paper was presented.

* Corresponding author at: Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Washington, DC 20375, United States. Tel.: +1 5713399312.
E-mail address: skhemlani@gmail.com (S.S. Khemlani).

0001-6918/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2012.01.010

Consider, for instance, the following scenario:

If a person does regular aerobic exercises then the person strengthens her heart.

Someone did regular aerobic exercises, but she did not strengthen her heart.

The two assertions are inconsistent, i.e., they cannot both be true. Given such an inconsistency, it is felicitous to ask: "why not?" One explanation for the person failing to strengthen her heart is that she had a heart defect. This explanation provides an exception to the generalization about regular aerobic exercise. It also suggests that an appropriate course of action is to seek a cure for the heart defect. Individuals could abandon the generalization as false, or, more likely, construe it as an idealization that holds by default: it is true in typical cases, but it tolerates exceptions. The assertion is accordingly interpreted as akin to the *generic* assertion, i.e., aerobic exercises strengthen the heart, which also tolerates exceptions (Leslie, 2008; Leslie, Khemlani, & Glucksberg, 2011). The principle of resolution therefore predicts that individuals create explanations to resolve inconsistencies, and that the explanations can lead to the tacit re-interpretation of general assertions as generics that hold by default. This re-interpretation yields a prediction: once individuals have formulated an explanation that resolves an inconsistency, they should be less likely to detect the inconsistency. An alternative hypothesis is that explanations have no effect on the status of inconsistencies, and so the ability to detect them is not subsequently impaired. To test these contrasting predictions, we carried out five experiments in which participants had to detect inconsistencies before or after they constructed explanations of them.

3. Experiment 1

According to the principle of resolution, explanations lead to re-interpretations of inconsistent assertions, and as a result an interaction should occur: inconsistency should be harder to evaluate after individuals have explained what's going on than beforehand, and this effect should be greater than the effect of explanations on judgments of consistency. Experiment 1 tested this prediction. The participants were presented with pairs of assertions, such as:

If a person is bitten by a viper then the person dies.

Someone was bitten by a viper, but did not die.

They also answered the question, "why not?" either before or after they evaluated the consistency of the assertions. The question, of course, called for them to explain what is going on.

3.1. Method

3.1.1. Participants

40 participants volunteered through the Mechanical Turk online platform hosted through Amazon.com. Mechanical Turk is a system that distributes tasks, surveys, and experiments to thousands of people for completion. The platform is a viable alternative to laboratory experimentation (see Paolacci, Chandler, & Ipeirotis, 2010, for a comparison of different recruitment methods). Participants volunteered for the study through a listing of studies, and they completed it for monetary compensation (in the form of credit towards their accounts). They could complete the study only once, and the pool of participants was constrained to meet several demographic specifications. The experiment was made available to a) only North American Amazon.com subscribers, b) only those participants who self-reported that they were native English speakers, and c) only those participants who reported that they had no prior expertise in logic.

3.1.2. Design and procedure

On each trial, participants were given a pair of consistent or inconsistent assertions. There were two different groups: 20 participants performed an explanation task before they evaluated the consistency of pairs of assertions, and 20 performed the two tasks in the opposite order. For both groups, half of the problems contained a generalization (1) that was inconsistent with a categorical assertion (2), e.g.,

1. If someone is very kind then he or she is liked by others.
2. Someone was very kind but was not liked by others.

For the other half of the problems, the inconsistency was eliminated by omitting the first clause in the categorical assertion, e.g.,

3. If someone is very kind then he or she is liked by others.
4. Someone was not liked by others.

Participants received equal numbers of consistent and inconsistent problems, and carried out the two tasks in succession for each problem. For the consistency task, they had to answer the question, "Can both of these statements be true at the same time?" (We used this question because participants are often uncertain about the meaning of "consistent" whereas the question is unambiguous.) They responded by pressing one of two buttons marked "Yes" or "No". For the explanatory task, they answered the question, "Why not?" The question made sense for both the consistent and inconsistent pairs, because the final clause in both sorts of problem was a negative assertion. The participants typed their answers into a text box provided on the screen. They were unable to see their answer to the first task when they carried out the second task. All of the problems were similar to the two examples above (see the Appendix A for all the materials in the experiments). Each participant encountered a given pair of assertions only once, and received the pairs in a different random order. The participants were given no clues that the generalizations in the study had exceptions. Instead, they were told that the experiment was about conflicts in information, and that they would have to carry out both an evaluation task and an explanatory task. The two tasks were illustrated for them as follows:

Suppose you are told the following:

1. *If a food item is not preserved, then it rots.*
2. *This food item was not preserved, but it did not rot.*

Based on this information, you will be asked to explain what is going on. For every trial, you will also be asked if both sentences can be true at the same time. Please respond based on what you think the appropriate answer is.

3.2. Results and discussion

Table 1 presents the percentages of correct evaluations in Experiment 1. Participants were far more accurate at detecting consistencies than inconsistencies (89% vs. 45%, Wilcoxon test, $z=4.00$, $p<.0001$, Cliff's $\delta=.69$). The group that evaluated the consistency of the assertions first was more accurate than the group that provided an explanation first (79% vs. 56%, Mann–Whitney test, $z=3.07$, $p<.005$, Cliff's $\delta=.66$). Likewise, the predicted interaction was significant: the difference in accuracy between consistent and inconsistent problems

Table 1
The percentages of correct evaluations of consistency and inconsistency in Experiment 1 depending on whether this task occurred before or after the explanatory task.

	Inconsistent problems	Consistent problems
Group that carried out the consistency task first	64	93
Group that carried out the explanatory task first	27	86

was greater for the group that carried out the explanatory task first (Mann–Whitney test, $z = 2.21$, $p < .025$, Cliff's $\delta = .48$). Hence, the participants were less likely to identify inconsistencies correctly when they had first created an explanation. For inconsistent problems, 19 out of the 20 participants produced more explanations that refuted the conditional than the categorical assertion, and there was one tie (Binomial test, $p < .0001$).

The results corroborate the principle of resolution, which predicts that explanations should make it more difficult to detect inconsistencies. But, did participants' explanations change their interpretations of the premises, or is there a simpler explanation of the results? One possibility is that when the participants are asked the initial question, "why not?", they automatically infer that the assertions must be consistent and that their initial reading of one of the assertions was probably too strict. So they revise the assertions to make one of them less strict to give the experimenter the benefit of the doubt. To check whether these demand characteristics drove performance, we modified the premises in the next experiment so that they implied no possibility of an exception to the generalization. Experiment 2 also used a more direct prompt for judgments of consistency.

4. Experiment 2

The experiment introduced two changes to the previous study. First, it used strict generalizations that did not admit of exceptions: the consequent of the conditional included the word "always," e.g., "If the aperture on a camera is narrowed, then less light always falls on the film." Second, a modified prompt instructed the participants to judge the consistency of the two assertions as explicitly stated.

4.1. Method

4.1.1. Participants

53 participants from the same online platform as in the previous study carried out the experiment for monetary compensation. None reported any training in logic.

4.1.2. Design and procedure

Participants were their own controls, and received equal numbers of trials with consistent and inconsistent problems. The materials were those used in Experiment 1, except that the consequent of conditionals included the qualifier "always" in the form, *If A then always B*. The explanation task was the same as before, but the consistency task was made more explicit by using the following prompt (including the italicization): "Can both of these statements, *as explicitly stated*, be true at the same time?" 28 participants carried out the explanation task before the consistency task, and the remaining 25 participants carried out the two tasks in the opposite order. They were unable to see their answers to the initial task when they carried out the second task, and each participant received the problems in a different random order.

4.2. Results and discussion

Table 2 presents the percentages of correct evaluations of consistency and inconsistency depending on whether this task occurred before or after the explanatory task. The results replicated those of the previous study. The participants made more correct responses to consistent problems than to inconsistent problems (91% vs. 63%, Wilcoxon test, $z = 3.69$, $p < .0001$, Cliff's $\delta = .50$). The group that evaluated consistency first was more accurate than the group that created an explanation first (82% vs. 72%, Mann–Whitney test, $z = 1.99$, $p < .05$, Cliff's $\delta = .32$). And the predicted interaction was significant: the difference in accuracy between consistent and inconsistent problems was greater for the group that carried out the explanatory task first (Mann–Whitney test, $z = 1.70$, $p < .05$, Cliff's $\delta = .27$). Hence, the creation of an explanation impaired accuracy in detecting inconsistencies.

Table 2
The percentages of correct evaluations of consistency and inconsistency in Experiment 2 depending on whether this task occurred before or after the explanatory task.

	Inconsistent problems	Consistent problems
Group that carried out the consistency task first	73	92
Group that carried out the explanation task first	54	90

A possible explanation of the main result is that the explanatory task is difficult, and so subsequently the participants had less cognitive resources for detecting inconsistencies. If so, any cognitively demanding task should have the same effect. This conjecture could explain the results of both the previous and the present experiment, because the participants who began with the explanatory task went on to evaluate both consistent and inconsistent assertions less accurately than the participants who began with the evaluation task. Experiments 3a and 3b accordingly tested whether any cognitively demanding task dulled reasoners' sensitivity to inconsistencies, or whether explanations are more likely to decrease accuracy.

5. Experiments 3a and 3b

In a preliminary study, the participants had to draw an inference from a pair of assertions, i.e., they had to answer the question, "What, if anything, follows?" This task made it harder for them to detect inconsistencies. However, the inferential task was not an adequate control, because the majority of inferences from inconsistent pairs of assertions were explanations that resolved the inconsistencies. The participants in the present two experiments therefore evaluated the consistency of pairs of assertions after carrying out one of two tasks: one group provided an explanation of the assertions and the other group carried out a cognitively demanding task that did not elicit an explanation. In Experiment 3a, the task was to assess which of the two events in each pair of assertions was more surprising. This task is akin to the explanatory task, because it calls for the participants to take the pairs of assertions into account, and to make an assessment of them. A limitation of the surprise task is that participants merely choose between two given options, whereas the explanatory task calls for them to generate a new proposition (an explanation). Hence, in Experiment 3b, the participants had to create a written justification for their judgment of whichever event in each pair of assertions was more surprising.

In Experiment 3a, the participants who evaluated surprise received trials such as:

If the aperture on a camera is narrowed, then less light falls on the film

The aperture on this camera was narrowed but less light did not fall on the film

In light of these statements, which of the following is more surprising?

1. *It's more surprising that the aperture on this camera was narrowed.*
2. *It's more surprising that less light did not fall on the film.*

Once they had chosen one of the two options, they carried out the consistency task. The other group of participants typed out their answer to the question "Why not?" before carrying out the consistency task. The participants in Experiment 3b received the same problems as Experiment 3a, but they received the following alternative instructions:

> *In light of these statements, please justify which statement is more surprising.*

They typed their response in a text box below the prompt. In all other respects, Experiments 3a and 3b were identical.

5.1. Method

5.1.1. Participants

40 participants volunteered for Experiment 3a and 38 volunteered for Experiment 3b. They came from the same online platform as before, and they participated for monetary compensation.

5.1.2. Design and procedure

Participants received an equal number of consistent and inconsistent problems, and received the same set of problems used in the previous study. Half of the participants carried out the explanation task before the consistency task, and the other half of the participants evaluated surprise, or justified their judgment of surprise, before the consistency task. They were unable to see their answers to the initial task when they carried out the second task. Each participant received the same problems as before in a different random order.

5.2. Results and discussion

Table 3 presents the percentages of correct evaluations of consistency and inconsistency in Experiments 3a and 3b. The results again corroborated the principle of resolution. Participants were less accurate with inconsistent than with consistent problems after the explanatory task than after either surprise task (Experiment 3a: Mann–Whitney test, $z = 1.64$, $p = .05$, Cliff's $\delta = .30$; Experiment 3b: Mann–Whitney test, $z = 3.10$, $p < .005$, Cliff's $\delta = .59$). No reliable difference between the two groups occurred for accuracy in evaluating consistent problems (86% vs. 84% in Experiment 3a, Mann–Whitney test, $z = .63$, $p = .53$, Cliff's $\delta = .10$; 77% vs. 70% in Experiment 3b, Mann–Whitney test, $z = 1.38$, $p = .17$, Cliff's $\delta = .26$). Likewise, participants who carried out either surprise task first were not reliably worse at evaluating inconsistent pairs than at evaluating consistent pairs (75% vs. 86% in Experiment 3a, Wilcoxon test, $z = 1.54$, $p = .12$, Cliff's $\delta = .22$; 66% vs. 70% in Experiment 3b, Wilcoxon test, $z = .45$, $p = .65$, Cliff's $\delta = .02$).

The two experiments replicated the previous studies: participants who created explanations tended then to judge an inconsistent set of assertions as consistent, but the evaluation or justification of surprise had no such effect. The results count against the possibility that any cognitively demanding task concerning the assertions makes it harder to detect inconsistencies. And both groups in Experiments 3a and 3b went on to evaluate consistent problems with no reliable impairment in accuracy. The one remaining possibility is that the results were a fortuitous consequence of differences between the experimental and control groups. In the next experiment, the participants therefore acted as their own controls.

Table 3
The percentages of correct evaluations of consistency and inconsistency in Experiments 3a and 3b depending on whether participants first assessed which event was more surprising or carried out the explanatory task. In Experiment 3a, the surprise task called for participants to select which assertion was more surprising; in Experiment 3b, the task called for participants to write out what they found surprising as well as a justification for their response.

	Experiment 3a		Experiment 3b	
	Inconsistent problems	Consistent problems	Inconsistent problems	Consistent problems
Group that carried out the surprisingness task first	75	86	66	70
Group that carried out the explanation task first	47	84	39	77

6. Experiment 4

Experiment 4 tested whether explanations impair evaluations of inconsistency more than judgments of surprise when participants carried out one or other of these tasks or neither of them, before they evaluated the consistency of the assertions.

6.1. Method

6.1.1. Participants

25 participants from the same online platform as in the previous studies carried out the experiment for monetary compensation. None had received any training in logic.

6.1.2. Design and procedure

Participants served as their own controls, and received equal numbers of trials with consistent and with inconsistent problems. The materials were those in the previous studies. For a third of the trials, participants carried out only the consistency task; on another third, they carried out the task of evaluating surprise (as in Experiment 3a) before the consistency task; and on the remaining third, they carried out the explanation task before the consistency task. The three conditions were intermingled with the problems presented in a different random order to each participant. They received each set of contents only once, and the contents were rotated so that each content occurred equally often in the three conditions in the experiment as a whole.

6.2. Results and discussion

Table 4 presents the percentages of correct evaluations of consistency and inconsistency. Participants were more accurate on consistent problems than inconsistent problems (71% vs. 52%, Wilcoxon test, $z = 2.38$, $p < .01$, Cliff's $\delta = .26$), and accuracy varied over the three sorts of trial (Friedman analysis of variance, $\chi^2 = 6.20$, $p < .05$). This main effect is attributable to the drop in accuracy on inconsistent problems when participants had provided explanations. The interaction was reliable, i.e., participants were less accurate on inconsistent problems when they had carried out the explanatory task than when they had carried out either the evaluation of surprise or no prior task, whereas their accuracy with consistent problems was comparable over the different tasks (Wilcoxon test, $z = 2.08$, $p < .025$, Cliff's $\delta = .23$).

As in the previous studies, the experiment showed that explanations increased the likelihood that participants evaluated inconsistent assertions as consistent. The present design is more sensitive than the previous one in which the tasks were allocated to different groups of participants. And the present results show that the effect cannot be explained on the grounds that the creation of explanations is more

Table 4
The percentages of correct evaluations of consistency and inconsistency in Experiment 4 depending on whether participants only evaluated consistency, first assessed which event was more surprising, or first created an explanation.

	Inconsistent problems	Consistent problems
Consistency task only	60	70
Surprisingness task, then consistency task	56	76
Explanation task, then consistency task	40	68

demanding than the evaluation of surprise, because the two different tasks had no reliable effect on the evaluation of consistent problems.

7. General discussion

Across five experiments, participants were more likely to evaluate inconsistent assertions as consistent after they had created an explanation for the inconsistency than otherwise. Experiment 1 showed that the effect occurred when participants were asked to construct explanations. Experiment 2 eliminated the possibility that the effect was a result of the demand characteristics of the procedure, i.e., that participants assumed that their task was to focus on the exceptions to the generalization, and therefore judged the assertions as consistent. Experiments 3a and 3b replicated the effect by comparing those who formulated explanations with those who evaluated or justified which event in the pairs of assertions was more surprising. Experiment 4 extended the effect to a design in which participants carried out all the experimental conditions. The control conditions in all of the studies showed that participants took the evaluation of consistency seriously. If they had focused solely on this task and made no re-interpretation of the assertions, then the creation of an explanation in itself would have had no effect on their performance. Instead, they failed to detect inconsistencies as a result of creating explanations. According to the principle of resolution, their explanations led them to re-interpret generalizations as idealizations that hold by default, and so they were less likely to treat the pairs of assertions as inconsistent. Consider, for example, whether the following generalization is true or false:

If a person is bitten by a viper then the person dies.

Given the further premise, say, that someone was bitten by a viper, many studies have shown that individuals tend to infer that the person died (see, e.g., Goldvarg & Johnson-Laird, 2001). Yet, their evaluations of an inconsistent case, such as:

Someone was bitten by a viper, but did not die.

could lead them to an interpretation that tolerates exceptions. That is, the generalization expresses a truth that holds by default, but disabling conditions can lead the participants to tolerate exceptions, e.g., the person was immediately given an antidote to the poison (Cummins, 1995). They have reasoned from inconsistency to consistency (see Johnson-Laird et al., 2004), and this newfound consistency makes it harder to detect the original inconsistency of the assertions.

An alternative possibility is that an explanation leads individuals to reject outright the generalization as false. In this case, however, they would make a correct negative response to the question: "Can both of these statements be true at the same time?" for inconsistent pairs of assertions, because they have already decided that one of the pair is false. For this reason, we suspect that the rejection of an assertion is unlikely to occur often.

Three gaps in the present account remain. The first concerns the quality and content of the participants' explanations. In all five of our experiments, the explanations showed a reliable bias to overrule strict interpretations of the generalizations rather than the categorical assertions. This pattern is consistent with earlier results (Johnson-Laird et al., 2004) and with the hypothesis that explanations lead to a re-interpretation of the generalization. Participants across all the experiments created similar explanations, though we consider here the data from Experiment 3a. The vast majority of participants' explanations (86%) introduced disabling conditions for the conditional generalizations. That is, they re-interpreted the conditionals so that they tolerated exceptions, e.g.:

S83: The person who was bitten by a viper got vaccine for the poison and lived.
S87: The gun could not be loaded or the gun misfired.
S106: The person had a heart condition.

Each of these explanations tacitly refuted the generalization (see Khemlani & Johnson-Laird, 2011). Individuals allowed for an exception to the general rule, but appeared to maintain their belief in a weaker version of the generalization, e.g., "If a person is bitten by a viper and does not have a vaccine (sic), then that person will die". A small minority of responses (12%) tacitly refuted the categorical statement. To do so, participants often introduced an ambiguity about the relevant entities, e.g.:

S104: It might be a different camera.
S83: The car's engine was tuned in a different special way than the one that lowers fuel consumption.
S89: It was a guy dressed in a viper costume.

The remaining 2% of responses were explicit denials of the original categorical assertion:

S35: The engine was not tuned correctly.
S67: It was not tuned right.
S98: Rods were not inserted correctly.

All the participants who explicitly revised the categorical statement evaluated the inconsistent assertions as inconsistent. Nevertheless, further research needs to interrelate the specific explanations to the evaluations of the inconsistent assertions.

Second, the mechanism underlying the phenomenon has yet to be pinned down (but see Melhorn, Taatgen, Lebiere, & Krems, 2011, for an account based on memory activation processes). When individuals explain an apparent inconsistency among a set of assertions, their explanations could sometimes rule out one of the assertions as false. For instance, individuals are likely to judge that the following conditional generalization is true universally and unequivocally:

If a person's brain is deprived of oxygen for 1 hour then the person dies.

And so they might not be prepared to believe a description of an apparent counterexample.

Finally, the experiments corroborated the principle of resolution as it applies to explicit conflicts. However, the principle also applies to implicit conflicts. For instance, individuals often fail to notice inconsistencies in discourse (Epstein, Glenberg, & Bradley, 1984). The interpretative system is geared toward making sense of discourse, and it appears to resolve anomalies on the fly in constructing coherent models of the discourse (see, e.g., Garnham, 1987; Glenberg, Meyer, & Lindem, 1987). Moreover, anomalous information can prompt reasoners to ask questions, construct explanations, and engage in overt causal reasoning (Graesser & McMahen, 1993; Khemlani & Johnson-Laird, 2011), and the resolution principle is relevant in each of these situations.

The present studies demonstrate the power of explanatory reasoning. Humans have a natural tendency to reason from inconsistency to consistency, and to rely on an abductive procedure that generates explanations from their knowledge of causal relations (Graesser, Singer, & Trabasso, 1994; Johnson-Laird et al., 2004). The consequence is that they subsequently overlook the inconsistency. In some situations, this behavior is practical, because it enables individuals to make sensible interpretations of generalizations and to understand discourse (Graesser, Bertus, & Magliano, 1995). In other situations, however, it may lead to striking lapses in reasoning. When a plausible explanation is available, regardless of whether it is true, reasoners may overlook inconsistencies and evaluate them in accordance with the explanation. In sum, to explain an inconsistency is often to explain it away.

Appendix A. Materials used in Experiments 1–4

In Experiment 2, the generalizations were modified slightly so that they were of the form, *If A then always B,* where A and B stood for the various propositions used, e.g., "a person is bitten by a viper" and "that person dies."

Domain	Generalization	Consistent categorical	Inconsistent categorical
Biology/ physiology	If a person is bitten by a viper then that person dies	Someone did not die	Someone was bitten by a viper but did not die
Biology/ physiology	If a person does regular aerobic exercises then that person strengthens his or her heart	Someone did not strengthen his heart	Someone did regular aerobic exercises but did not strengthen his or her heart
Mechanical	If a car's engine is tuned in the special way then its fuel consumption goes down	This car's fuel consumption did not go down	This car's engine was tuned in the special way but its fuel consumption did not go down
Mechanical	If graphite rods are inserted into a nuclear reactor, then its activity slows down	The nuclear reactor's activity did not slow down	Graphite rods were inserted into this nuclear reactor but its activity did not slow down
Mechanical	If the aperture on a camera is narrowed, then less light falls on the film	Less light did not fall on the film	The aperture on this camera was narrowed but less light did not fall on the film
Mechanical	If a person pulls the trigger then the pistol fires	The pistol did not fire	Someone pulled the trigger but the pistol did not fire
Natural	If a substance such as butter is heated then it melts	This piece of butter did not melt	This piece of butter was heated but it did not melt
Natural	If these two substances come into contact with one another then there is an explosion	There was no explosion	These two substances came into contact with one another but there was no explosion
Psychological	If someone is very kind then he or she is liked by others	Someone was not liked by others	Someone was very kind but was not liked by others
Psychological	If a person receives a heavy blow to the head then that person forgets some preceding events	Pat did not forget any preceding events	Pat received a heavy blow to the head but did not forget any preceding events
Social/ economical	If people make too much noise at a party then the neighbors complain	The neighbors did not complain	People made too much noise at a party but the neighbors did not complain
Social/ economical	If the banks cut interest rates then the economy increases	The economy did not increase	The banks cut interest rates but the economy did not increase

References

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognitive Development*, *7*, 139–172.

Anderson, C. A., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*, 1037–1049.

Byrne, R. M. J. (2005). *The rational imagination.* Cambridge, MA: MIT Press.

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

Crowley, K., & Siegler, R. S. (1999). Explanation and generalization in young children's strategy learning. *Child Development*, *70*, 304–316.

Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, *23*, 646–658.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3–19.

Epstein, W., Glenberg, A. M., & Bradley, M. M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. *Memory & Cognition*, *12*, 355–360.

Garnham, A. (1987). *Mental models as representations of discourse and text.* New York, NY: Halsted.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*, 69–83.

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

Graesser, A. C., Bertus, E. L., & Magliano, J. P. (1995). Inference generation during the comprehension of narrative text. In R. Lorch, & E. O'Brien (Eds.), *Sources of coherence in reading* (pp. 295–320). Hillsdale, NJ: Lawrence Erlbaum.

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Education and Psychology*, *85*, 136–151.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.

Harman, G. H. (1965). The inference to the best explanation. *Philosophical Review*, *74*, 88–95.

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

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, *111*, 640–661.

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 225–254.

Khemlani, S., & Johnson-Laird, P. N. (2011). The need to explain. *Quarterly Journal of Experimental Psychology*, *64*, 2276–2288.

Legare, C. H. (2012). Exploring explanation: Explaining inconsistent information guides hypothesis-testing behavior in young children. *Child Development*, *83*, 173–185.

Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, *81*, 929–944.

Leslie, S. J. (2008). Generics: Cognition and acquisition. *Philosophical Review*, *117*, 1–47.

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, *65*, 15–31.

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

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

Melhorn, K., Taatgen, N., Lebiere, C., & Krems, J. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1391–1411.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411–419.

Patalano, A. L., Chin-Parker, S., & Ross, B. H. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language*, *54*, 407–424.

Rittle-Johnson, B. (2006). Promoting transfer: The effects of direct instruction and self-explanation. *Child Development*, *77*, 1–15.

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.

Shafto, P., & Coley, J. D. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naïve similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 641–649.

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

Walsh, C., & Johnson-Laird, P. N. (2009). Changing your mind. *Memory & Cognition*, *37*, 624–631.

Wellman, H. M., Hickling, A. K., & Schult, C. A. (1997). Young children's psychological, physical, and biological explanations. *New Directions for Child Development*, *75*, 7–25.