

Evaluating Theories in the Context of a Web of Information

Amy M. Masnick (amm13@cornell.edu)

Susan M. Barnett (smb11@cornell.edu)

Stephanie L. Thompson (slt4@cornell.edu)

Barbara Koslowski (bmk2@cornell.edu)

Department of Human Development

MVR Hall, Cornell University

Ithaca, NY 14853

Abstract

Philosophers of science have long argued that when evaluating explanations, we do not consider ideas in isolation. Instead, we possess an integrated web of information that comprises the context we consider when weighing evidence about any component of this web. In this paper, we provide empirical evidence that theories are considered in context by demonstrating that non-scientists change the strength of their belief in both of two alternative theories, even when only given information about one of these hypotheses. In addition, we seek to identify and describe some of the types of information people use in evaluating theories. Information about mechanism, inferences that discriminate between two explanations, and information about closely related situations in which the target factor operates as a mechanism can all significantly affect ratings of two rival explanations.

When evaluating theories, what types of information do we use? What changes take place as theories become increasingly refined and elaborated? In this paper, we identify some of the types of information that are often used as theories or explanations are elaborated, and we assess the extent to which such information makes rudimentary explanations increasingly or decreasingly convincing.

Quine (1969) suggests that our beliefs are not independent; instead, they are embedded in a "web of belief." Adopting Quine's notion, we suggest that knowledge forms an integrated web, and that even non-scientists evaluate theories within the context of this interconnected web of information. In effect, theories and information are interdependent, allowing people to use a bootstrapping process, iterating between theory and evidence to reach a conclusion.

Elaborated theories include several types of information.

As an example, consider Darwin's theory of evolution. It includes a mechanism by which evolution takes place (natural selection), evidence that distinguishes it from a competing theory (fossils do not fit with creationism), a way to account for anomalies (why some mammals lay eggs), and a description of related situations in which the

analogue of natural selection plays a role (e.g., animal breeding). It is the combination of all of this information that makes this theory so compelling to many.

In the literature on scientific reasoning, the question of what makes a theory convincing has typically been framed as a question about causation: What makes us decide that one factor rather than another is the cause of a phenomenon? Past studies are limited in that they deal with explanations in a restricted way. In most of the studies, participants are asked to choose which factor is the sole cause of the effect, in the absence of any potentially related information, such as information about mechanism (e.g., Kuhn, Amsel, & O'Loughlin, 1988). Indeed, many tasks in the causal reasoning literature are deliberately designed to minimize the role of other sorts of information.

There is no question that identifying causal factors constitutes one level of explanation. However, the level is a rudimentary one. Often, when people need to evaluate an explanation, the explanation consists not only of a possible causal factor, but also of some description of the process by which the factor operates. Accordingly, some researchers have also noted the importance of information about causal mechanism, anomalies and alternative accounts in deciding whether a particular factor has caused an effect (Koslowski, 1996).

However, researchers typically examine the role of alternative causes by arranging some factors to covary with the effect, and other factors not to covary (Kuhn, Amsel, & O'Loughlin, 1988). Analogously, when researchers examine the effect that anomalous or problematic data have on an explanation, the anomalies in question are situations in which the covariation is imperfect. That is, the studies of mechanism, alternative accounts, and anomalous data are couched in terms of a covariation framework. And, as noted above, information about only one aspect of the framework, such as covariation, has limited utility.

To address these issues, we set out to explore the extent to which theories are evaluated with respect to other information in a web of knowledge, and to examine what effects various types of information have on making fairly

rudimentary explanations increasingly or decreasingly convincing. We identified six types of information that might affect how convincing participants found two rival explanations. The six types of information are as follows (detailed descriptions and examples can be found in the Procedure section): 1) a discriminating inference, 2) a mechanism, 3) the target mechanism in a closely related situation, 4) the target mechanism in a distantly related situation, 5) an anomalous event and a possible resolution of the anomaly 6) and an anomalous event with no suggestion of a resolution. We hypothesize that each type of information will have some effect on the ratings of plausibility of each of two explanations offered for an observed phenomenon, even when the information is only directed at one explanation.

Experiment 1

Method and Procedure

Participants Thirty-two undergraduate students (16 women, 16 men) participated in exchange for course credit.

Procedure Participants were interviewed individually, for 20–45 minutes. Each participant was shown six story problems, consisting of an observation to be explained and two possible explanations for the observation (a target and an alternative account). The Target and Alternative were not labeled for the participants, and their order was counterbalanced across conditions. For each story, the interviewer read the information with the participant, and then asked the participant to rate how convincing he/she found each explanation on a seven-point scale. Participants were also asked to explain why the explanations were or were not convincing. Qualitative data will be summarized in another paper (Koslowski, Thompson, Barnett, & Masnick, in preparation).

After rating each explanation for a story, participants were shown another page with one of six types of additional information, as listed in the Introduction. Participants were then asked to rate the hypotheses again.

Because it was not possible to have plausible information of every type for every story, two examples are provided to demonstrate all six types of information.

The first example concerns why conquered cultures typically adopt the religion of their conquerors. The observation to be explained and the two possible explanations follow:

Observation

Historically, when one culture has been conquered by another culture, the conquered culture has come to

follow the religion of the conquering culture. For example, when Viking raiders conquered ancient Britain, the Britons began to have public celebrations of Viking religious festivals and to build shrines to Viking gods.

Possible Explanations

[Target] One hypothesis is that this happens because there are often *similarities* between the gods of the conquering culture and the gods of the conquered culture. For example, the Vikings had Freyja, the goddess of fertility, while the Britons had Brigit, who was also a fertility goddess. Therefore, calling the British Brigit by a different name was not a difficult change to make.

[Alternative] Another hypothesis is that the conquered culture does not really adopt the religion of the conquering culture. To be sure, in public, they celebrate the feast days of the new religion and they dedicate buildings to the new gods. However, this is just for political reasons, *to avoid retaliation* by the conquerors. In private, they continue to worship their own gods and to practice their own religion.

The second example concerns height differences between mountain and plains dwellers.

Observation

Medical doctors and anthropologists have found that the people who live in the mountainous regions of Mexico are smaller, in stature, than are people who live in the non-mountainous plains.

Possible explanations

[Alternative] One hypothesis is that historically, the people who migrated to the mountains came from a *different gene pool* than did people who stayed on the plains. On this hypothesis, people who moved to the mountains came from a gene pool with a smaller stature.

[Target] Another hypothesis is that mountain dwellers have access to *less oxygen* in the air and, so, their bodies adapt by growing to a smaller size.

The story problem about conquering cultures will be used to illustrate four of the types of additional information, namely, information about discriminating inferences, mechanism, anomalous observations, and a way of resolving the anomalies.

Discriminating Inference – an event that is consistent with, and can be explained by, the target theory but that is not what one would expect if the alternative theory were true. For example:

If conquered cultures do adopt the new religion because of perceived similarities, and not just for public show, then we should find statues and shrines to

the new gods even in the sections of private homes that were open only to family members. Those rooms would not be under public scrutiny so that fear of retaliation would not be an issue.

And, this is what we do find.

Mechanism Information -- additional details about a mechanism or process by which the target explanation might operate. For example:

Every culture faces certain life issues. Thus, most cultures are likely to have gods to help them with these issues. For example, most cultures have developed some sort of being to turn to for help with healing.

If conquered cultures adopt the gods of their conquerors because of similarities between the two sets of gods, then the common life issues could explain why it is that similarities are likely to occur.

Resolved Anomaly -- an anomaly and information that suggests a way of reconciling the anomaly to the target theory. In this study, resolving the anomaly consisted of modifying the theory to account for the anomaly rather than, for example, giving a reason why it might be sensible to dismiss the anomaly. (The resolved anomaly did not consist of simply adding the resolution to the anomalous information from the Anomaly condition because we did not want the resolved anomaly to be substantially longer than the anomaly itself.) For example:

Cultures (such as the Mayans) who had many gods were sometimes conquered by cultures (such as Catholic Spain) who were monotheistic, that is, who worshipped only one god.

Although monotheistic cultures had one god, they had many saints. If conquered cultures do adopt the new religion because they detect similarities, it could be because of similarities between the many gods of the conquered culture and the saints of the conquering culture.

Anomalous Information -- an observation that is not, at first glance, congruent with the target explanation. In keeping with T.S. Kuhn's treatment of anomalies, we write "at first glance" because, as Kuhn has pointed out, anomalies to an explanation can often be resolved, either by questioning the anomalous data or else by modifying or elaborating the explanation (Kuhn, 1970). For example:

Cultures (such as the Mayans) who had many gods were sometimes conquered by cultures (such as Catholic Spain) who were monotheistic, that is, who worshipped only one god.

If conquered cultures do adopt the new religion because they detect similarities, that would leave the question of how cultures who worshipped many gods

could see similarities between their many gods and the single god of a conquering culture.

The story-problem about height differences in different altitudes will be used to demonstrate the final two types of additional information, closely and distantly related analogous situations.

Close Analogous Situation -- information that the target factor is known to operate as a causal mechanism in fairly similar situations. For example:

The fact that oxygen levels decrease at higher altitudes is known to cause people who live at low altitudes to become fatigued when they visit the mountains.

So, regardless of whether decreased oxygen is the explanation for shorter stature in some groups of people, decreased oxygen has certainly been found to affect other things.

Distant Analogous Situation -- information that the target factor is known to operate as a causal mechanism in fairly distantly related situations. For example:

The fact that oxygen levels decrease at higher altitudes is known to affect baking methods at high altitudes.

So, regardless of whether decreased oxygen is the explanation for shorter stature in some groups of people, decreased oxygen has certainly been found to affect other things.

For each type of information, control information was presented in place of the experimental information for three of the six stories each participant read. This information was designed to match the experimental information in both structure and length and to give further information about the content area, without directly bearing on the plausibility of either hypothesis. (See the appendix for examples.)

Results and Discussion

A mixed model ANOVA was used to examine the effects of different types of information upon ratings of both the Target and Alternative explanations. For each type of information, the within subjects variables were experimental or control status and story content. Subject was treated as a random variable. For each type of information, the change in ratings after reading the experimental or the control information was compared. The results are summarized in Table 1.

The results of this experiment demonstrate that there are several different types of information that make people more or less convinced of an explanation. Inferences that discriminate between the two explanations significantly increased the Target rating while decreasing the rating of the Alternative. Information about mechanism and examples of similar situations in which the target factor

Table 1: Effects of Additional Information

Type of information	Mean change in ratings for Target	Mean change in ratings for Alternative
Discriminating	1.75***	-0.81***
Inference	-0.13	0.13
Mechanism	0.81***	-0.75***
	-0.13	0.13
Close Analog.	0.59***	-0.29**
Situation	-0.13	0.20
Distant Analog.	-0.13	-0.25
Situation	-0.19	-0.19
Resolved	0.50	-0.31**
Anomaly	0.19	0.19
Anomaly	-0.75	-0.06
	-0.06	0.19

Comparison of experimental (shaded) vs. control
 *p<.10, **p<.05, ***p<.01

operated each caused a significantly greater increase in ratings of the Target explanation than the control information. In addition, ratings of the Alternative explanation decreased significantly, *even though the information was directed only towards the Target explanation*. This finding provides evidence that explanations are not considered in isolation. When evaluating hypotheses, people consider several options in a set simultaneously, and use information about one explanation to evaluate other explanations. This is consistent with Quine’s suggestion of an interconnected web of information.

While anomalous information appears to decrease confidence in the Target explanation, this difference did not reach significance. This may be because participants have already resolved potential anomalies for themselves, and considered these challenges when rating both the Target and Alternative hypotheses. Several participants mentioned reasons to dismiss the anomaly in explaining their ratings.

Examples of the target factor’s application in a distantly related situation did not significantly influence participants’ views of the explanations, suggesting they did not find them relevant enough to use. This observation was supported by many participants’ comments that they did not see the relationship between the pieces of information given.

In addition to establishing that people rely on multiple types of information to evaluate even rudimentary theories, we also wanted to explore the circumstances under which this effect holds. Specifically, we wanted to investigate how evidence addressing the plausibility of different explanations influenced the effects of later information such as that presented in Experiment 1. Even after reading information that undermines one hypothesis, do people still use additional information in the same way?

Experiment 2

Method and Procedure

Participants Thirty undergraduate students (14 men, 16 women) participated in exchange for course credit.

Procedure The procedure was similar to that of Experiment 1, with one additional phase in the middle. In this study, participants were asked to rate the explanations for each of six story problems three separate times. The story problem scenarios were identical to those used in Experiment 1. For the first ratings, they were shown an observation and the two possible explanations.

For the second ratings, they were given information that either undermined the Alternative explanation or left it viable. (Each participant had the Alternative undermined for three of the six stories presented, and left viable for the other three stories.) The Alternative was undermined by presenting evidence that events that would be expected to occur if the Alternative were true, in fact did not occur. For the example of conquering cultures and their religions, the following information was given to undermine the Alternative explanation:

Conquered cultures follow the religion of the conquerors, and this is true regardless of whether the conquerors have been brutal and retaliated against non-followers.

To leave both alternatives still viable, it was suggested that it is not possible to obtain the critical information needed to prove either hypothesis. For example, for the example of conquering cultures, the following information was presented:

Historical records are often fragmentary. Therefore, it is not always clear whether there actually were similarities between the religions of the conquered and the conquerors, nor is it always clear whether the conquerors retaliated against people who did not adopt the new religion.

For the third rating, participants were shown a piece of additional information as in the first experiment (i.e., an

anomaly, an inference, etc.) and asked again to rate the two explanations. As in Experiment 1, in three of the six stories, participants were shown the experimental information in three of the six stories, and control information in the other three stories.

Results and Discussion

We first examined the effect of empirical evidence (whether the Alternative explanation was undermined or left viable) on ratings of the two explanations. A mixed model ANOVA was used to assess the overall significance of undermining the Alternative or leaving it viable on the rating of each explanation. This condition (undermined/still viable) and story were both within subjects variables. Subject was treated as a random variable. When information undermining the Alternative was presented, participants were less convinced of the viability of the Alternative (mean change = -1.46) while becoming more convinced of the target (mean change = 0.47). When the Alternative was left viable, the mean Target rating decreased by 0.47, while the mean Alternative rating decreased by 0.16. The difference in changes between the two conditions were highly significant, $F(1, 148) = 96.72, p < 0.001$, and $F(1, 148) = 81.73, p < 0.001$, for the Target and the Alternative, respectively.

After reading information suggesting that the Alternative may not be correct (i.e., that events it predicts do not actually occur) participants found this explanation far less convincing than they had before. However, they also then found the Target explanation more convincing, even though the information they read was not related to this second explanation. This finding suggests that the mere presence of an alternative alters one's belief in the accuracy of a theory. When one alternative becomes less likely, any remaining alternatives suddenly become more convincing, because there are a finite set of plausible explanations, and the evaluation of each theory occurs in the context of an interconnected web of information

Next we explored the effect of the additional types of information on ratings of both explanations. We used a general linear model to consider the effects of whether the Alternative had been undermined, for each of the six types of information, controlling for story. Table 2 shows the mean changes from the second to the third ratings of both the experimental and the control information. The table separately examines the Target (T) and the Alternative (A), considering whether the Alternative was undermined or not in the second phase.

However, even when one theory is undermined, additional information can still have a significant effect on the ratings of both explanations. Compatible with

Table 2: Effects of additional information when Alternative undermined or still viable

Type of information	Mean Δ in T ratings		Mean Δ in A ratings	
	A undermined	A still viable	A undermined	A still viable
Disc. Inference	1.57**	2.00**	-0.86*	-1.00**
	-0.13	0.14	0.25	0.14
Mechanism	0.50	1.00**	-0.25	-0.14
	0.14	-0.13	0.00	0.00
Close Analog. Sitn	0.43**	0.25	-0.29	0.00
	0.00	0.00	0.00	0.00
Distant Analog. Sitn	-0.50	0.14	0.00	-0.14
	0.00	0.00	0.00	0.00
Resolved Anom.	0.14	0.63	0.29	-0.25
	0.13	0.00	0.13	0.00
Anomaly	-1.00	-0.86	0.25	0.00
	-0.14	0.00	0.29	0.38

Comparison of experimental (shaded) vs. control
* $p < .10$, ** $p < .05$, *** $p < .01$

Experiment 1, we found that discriminating inferences, mechanism information, and closely analogous situations significantly influenced ratings of the Target explanation, at least under some conditions. Again, anomalous information had a non-significant impact, and distant analogies were viewed as too different to affect views on the Target hypothesis. No consistent pattern emerged in the changes in ratings of the Alternative explanation.

General Discussion

Taken together, the results of these two studies indicate that people consider evidence in context, and rely on several different types of information when evaluating theories. A complete "web of information" consists of many different types of knowledge, collateral information that buttresses a central concept in multiple ways. It also includes information about multiple explanations, so that the plausibility of one hypothesis is dependent on information about the other.

Previous research about how people evaluate theories has looked at the theories in isolation, and has only examined the effects of new information on the target account.

Clearly, this is only one piece of the picture. When considering two hypotheses together, information about only one hypothesis can change the ratings of both. This finding argues for a more comprehensive approach in future research, addressing the impact of information on more than simply one theory in isolation. Instead, it is critical to explore the effects of new information on all relevant pieces of an interconnected web of information used to evaluate evidence.

After confirming that even rudimentary explanations are evaluated in context, the obvious next step is to describe what sorts of information affect such evaluations.

Consistent with past research (e.g., Koslowski, 1996), we found that information about mechanism makes an explanation increasingly convincing. When no other information is given, this information also makes an alternative explanation less convincing.

Information that differentiates between two competing explanations by supporting one and not the other is most compelling -- it makes one explanation much more convincing, and the other much less so.

Anomalous information changes ratings but not significantly. Participants' comments suggest that they may be considering such information when rating the explanations. Chinn and Brewer (1993) have suggested that there are seven different ways people respond to anomalous data, indicating a potential direction for exploring the role of anomalous information. Further exploration of these seven responses may be necessary to determine why the anomalous information had a nonsignificant effect in these studies.

While knowledge of a closely related analogous situation in which the target factor operates as a causal mechanism increases ratings of the Target explanation, knowledge about a distantly related situation does not. This finding is consistent with the observation that "long-distance analogies" are far less likely yield new insights than much more closely related analogies (Dunbar, 1995).

In summary, evidence related to one explanation also affects the ratings of the plausibility of another explanation, suggesting the existence of a web of information. These exploratory studies suggest that several types of evidence comprise this web, including knowledge about mechanism, inferences that differentiate between hypotheses, and closely related analogous situations. All yield information that significantly influences ratings of two alternative explanations. Accordingly, we believe that future studies of theory evaluation must take into account the interrelated web of information in order to capture accurately how people reason.

Acknowledgments

We thank Geoff Moran for assistance with data collection, and Yong Liao and Melanie Swiderek for comments on an earlier draft of this paper.

References

- Chinn, C.A., & Brewer, W.F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1-49.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J. Sternberg, & J. Davidson (Eds.). *Mechanisms of Insight*. Cambridge, MA: MIT Press.
- Koslowski, B., Thompson, S., Barnett, S., & Masnick, A. (in preparation). Qualitative responses to scientific reasoning problems.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.
- Kuhn, D., Amsel, E. and O'Loughlin. (1988). *The development of scientific thinking skills*. Orlando, FL: Academic Press.
- Kuhn, T.S. (1970). *The structure of scientific revolutions*. 2nd ed., Chicago: University of Chicago Press.
- Quine, W.V. (1969). Natural kinds. In W.V. Quine, *Ontological relativity and other essays*. New York: Columbia University Press.

Appendix

An example of control information for discriminating inferences:

If conquered cultures do adopt the new religion because of perceived similarities, and not just for public show, then we should find statues and shrines to the new gods even in the sections of private homes that were open only to family members. Those rooms would not be under public scrutiny so that fear of retaliation would not be an issue.

However, after so long a time, it is not always possible to tell which rooms were and were not under public scrutiny.

An example of control information for the analogous situations:

People who live in the mountains of Mexico have fascinating culture that has been much studied by anthropologists.

They have complicated kinship structures and a strong tradition of oral storytelling, which enables them to pass on stories and legends to new generations.