The Theory-Ladenness of Data: An Experimental Demonstration

William F. Brewer
Department of Psychology
University of Illinois at Urbana-Champaign
603 East Daniel St.
Champaign, IL 61820
wbrewer@s.psych.uiuc.edu

Clark A. Chinn
Center for the Study of Reading
University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, IL 61820
chinn@vmd.cso.uiuc.edu

Abstract

Most philosophers of science now believe that scientific data are theory laden, i.e., the evaluation of data is influenced by prior theoretical beliefs. Although there is historical and psychological evidence that is consistent with the theory-laden position, experimental evidence is needed to directly test whether prior beliefs influence the evaluation of scientific data. In a fully counterbalanced design, one group of subjects received evidence that dinosaurs were cold-blooded, and another group of subjects received evidence that dinosaurs were warm-blooded. The subjects reported a strong belief in whichever theory they had read about. Then subjects were presented with a piece of data that supported one theory and contradicted the other theory. The identical piece of data was rated as more believable when it was consistent with the subject's theory than when it was inconsistent. These results provide clear support for the position that scientific data are theory laden.

Introduction

The purpose of this paper is to provide experimental support for the position that the evaluation of scientific data can be strongly influenced by prior theoretical beliefs. This study uses the techniques of cognitive science to investigate issues in the philosophy of science. Thus, we consider this line of work to be a contribution to a naturalist approach to the philosophy of science (e.g., Giere, 1985; Maffie, 1990).

Theory-Ladenness of Observation

The influence of theory on scientific data was emphasized in the work of Hanson (1958) and Kuhn (1962). Both of these philosophers were particularly interested in the role of theory in influencing the scientist's direct perceptual observations. For example, both of them argued that when looking at a sunrise, a geocentric theorist (e.g., Tycho Brahe) would see the sun rising, while a heliocentric theorist (e.g., Copernicus) would see the earth turning toward the sun. In a recent paper, Brewer and Lambert (1993) argue that the particular psychological examples used by Hanson and Kuhn may not be adequate to support the view that scientific observation is theory laden, but that there is, in fact, convincing laboratory evidence to support their position. Brewer and Lambert take the position, however, that current work in cognitive science only supports a moderate form of the theory-laden position, in which the top-down (theory-driven) processes only have major influences when the sensory input is weak or ambiguous.

Theory-Ladenness of Data

In the present paper we want to focus on the theory-ladenness of scientific data in the broader sense that includes both personal observations and reports of experimental data. Philosophers of science initially focused on the issue of the theory-ladenness of observation because there has been a long tradition in philosophy of attempting to show that immediate perception could lead to knowledge (in the philosopher's sense of justified true belief). However, it seems to us that in trying to understand the psychology of scientific reasoning, the issue of the theory-ladenness of data in the broad sense is of even more importance. The information used by working scientists to formulate and evaluate theories is much more likely to be data in the broad sense than it is to be personal perceptual observations.

The Quine-Duhem Thesis

The particular aspect of the theory-ladenness of data that we focus on in this paper is the role of theory in the evaluation of the quality of scientific data. In the terms of current philosophy of science, our experiment can be seen as examining an aspect of the Quine-Duhem thesis. Duhem (1914/1991) and Quine (1951/1963) both argued that theory-inconsistent data cannot disprove a theory. They pointed out that when faced with data that are inconsistent with a theory, a scientist can always make adjustments to the theory or to other background assumptions that will preserve the theory.

In a series of recent papers we (Chinn & Brewer, 1992, 1993) have elaborated the Quine-Duhem thesis using evidence from the history of science and from psychology to examine the responses made by scientists and nonscientists to anomalous data. We concluded that there are seven basic responses that can be made to theory-inconsistent data. The individual can: (1) *ignore* the data; (2) *reject* the data on methodological grounds; (3) *exclude* the data from the domain of the theory; (4) hold the data in *abeyance*; (5) *reinterpret* the data to make it consistent with the theory; (6) accept the data and make *peripheral changes* to the theory; and (7) accept the data and *change the theory*. Note that of the seven responses we described, all but one allow the individual to retain the original theory in the face of the anomalous data. The present paper is an experimental study of the role of theory in the evaluation of theory-consistent and theory-inconsistent data.

Evaluation of Theory-Consistent and Theory-Inconsistent Data

Historical evidence. When one examines historical cases of the evaluation of theory-consistent and theory-inconsistent data, one finds examples that fall into all of the categories described by Chinn and Brewer. For example the proponents of rival theories may view each other's experiments as conceptually flawed and therefore easily reinterpretable. When Pasteur and his opponents were debating the issue of whether spontaneous generation is possible, scientists conducted many experiments in which they examined whether living germs would appear in sterile solutions. Pasteur discounted the experiments in which germs did appear by saying that there must have been some unknown source of contamination. Similarly, Pasteur's opponents reinterpreted experiments in which germs did not appear by arguing that by sterilizing the air the experimenters destroyed the life-giving properties that permitted air to support spontaneous generation (Collins & Pinch, 1993).

Even when proponents of rival theories can agree on the interpretation of a particular piece of data, they may reach different conclusions about the soundness of the methodology. In the debates over continental drift, there were data that suggested that Greenland was moving relative to the continents, thus providing support for continental drift. Opponents of continental drift dismissed these data as unreliable. Proponents of continental drift admitted that there was unreliability in the data but evidently considered them sound enough to support the claim that Greenland was in fact moving westward (Frankel, 1988).

Psychological evidence. A number of quite different kinds of psychological data can be seen as providing evidence relevant to the issue of the evaluation of theory-consistent and theory-inconsistent data. Pickering and Monts (1982) examined the lab reports of students in an undergraduate chemistry course. The particular experiment that the students carried out gave data inconsistent with the theories they had learned in the course. In writing up the report of the lab, over half of the students rejected their own data and most of them argued that they must have made methodological errors in the way that they carried out the experiment. In the area of social psychology, researchers (e.g., Lord, Ross, & Lepper, 1979) have investigated the impact of preexisting social attitudes on the evaluation of evidence. For example, individuals who favor capital punishment are more critical of the methodology of studies that are inconsistent with their beliefs than they are of studies that are consistent with their beliefs. Mahoney (1977) carried out a study of how reviewers evaluate manuscripts that are either consistent or inconsistent with their theoretical perspective. He found that for artificially constructed manuscripts with identical methodology, theory-inconsistent reviewers considered the methodology to be much worse than did theory-consistent reviewers.

Thus, while there is some historical and psychological evidence that is consistent with the view that theory leads to differential evaluation of data, it seems to us that what is needed is an experiment that experimentally manipulates the scientific beliefs of individuals and directly examines the hypothesis that individuals differentially evaluate scientific data on the basis of their current scientific theories.

Experiment

The basic purpose of the experiment was to examine the role of individuals' theoretical beliefs on their evaluation of data using realistic scientific theories and data.

Method

Design. There were three experimental groups in the experiment. Each group received the same piece of data. However, the experiment was designed so that for one group the data were consistent with their current theory, for another group the data were inconsistent with their current theory, and there was a control group who received the data without a strong prior theory. In order to have the same piece of data be consistent for one group of subjects and inconsistent for another we used a pair of scientific theories with the logical property that data supporting one of theories implied the rejection of the other theory. Thus, data supporting the first theory were inconsistent with the second theory and vice versa.

In order to minimize the problems that could be caused by our subjects' prior knowledge and beliefs about the theories, the particular theory received by the subjects was counterbalanced. Thus, within the consistent data group, half of the subjects received Theory A followed by data supporting Theory A; the other half of the subjects received Theory B followed by data supporting Theory B. There was a corresponding counterbalancing for the inconsistent data condition.

Materials. The theories we chose to use were current theories about the warm-bloodedness or cold-bloodedness of dinosaurs (Thomas & Olson, 1980). Pilot work showed that our undergraduate population could easily understand these theories and that each theory, when read alone, provided a convincing account of the characteristics of the dinosaurs. Also, this pair of theories had the needed property that data supporting one of the theories implied the rejection of the other theory.

Each subject received a theory passage or a control passage followed by a data passage. There were three different experimental passages (2 theory passages, 1 control passage) and two different data passages. All experimental passages began with a discussion of the differences between warm-blooded and cold-blooded animals. This section of the experimental passages pointed out that warm-blooded animals generate their own body heat internally and that cold-blooded animals do not. Thus, this initial passage exposed each subject to the broad outline of each theory, but did not supply any supporting evidence.

The *warm-blooded theory passage* began with the initial information describing the differences between warm-blooded and cold-blooded animals and then continued by giving a variety of arguments and data supporting the position that dinosaurs were warm blooded (e.g., all warm-blooded animals have an upright posture, while cold-blooded animals have limbs that project from the sides, and the fossil record shows that dinosaurs had upright posture). The total warm-blooded passage (initial information and warm-blooded theory) was 1,303 words long.

The *cold-blooded theory passage* began with the same initial passage discussing the differences between warm- and cold-blooded animals and then continued with arguments and data supporting the position that dinosaurs were cold blooded (e.g., warm-blooded animals have to eat much more than cold-blooded animals to keep their metabolism going, and a warm-blooded animal as large as a dinosaur would have required impossible amounts of food to survive). The total cold-blooded passage was 1,478 words long.

The *control passage* contained only the initial information discussing the differences between warm and cold-blooded animals with no discussion of how these differences applied to dinosaurs. The control passage was 245 words long.

There were two data passages, one containing data consistent with the warm-blooded theory (thus inconsistent with the cold-blooded theory) and one containing data consistent with the cold-blooded theory (thus inconsistent with the warm-blooded theory).

The *warm-blooded data passage* stated that fast-growing animals have characteristic patterns of dense bone tissue and that only warm-blooded animals grow fast enough to have these patterns. The passage then noted that dinosaurs show the characteristic patterns for fast growth, so they must have been warm blooded.

The *cold-blooded data passage* stated that the brains of cold-blooded animals have small, simple pineal systems, whereas warm-blooded animals have large, complex pineal systems. The passage noted that the pineal systems leave a clear impression in the skull and that dinosaur skulls show the impression of simple pineal systems, so dinosaurs must have been cold blooded.

Within each type of data passage there was an additional factor. Half of the data passages were written in an impersonal textbook format For example: "Microanalysis of dinosaurs' bones indicates that the growth rate of dinosaurs was much, much faster than that of crocodiles, and as fast as a baby ostrich." The other half of the data passages gave the same information in the form of the results of an actual experiment. For example: "Robert Stockton, a paleontologist at Oregon State University, has conducted a microscopic analysis of fossilized rib bones of two separate species, stegosaurus and ankylosaurus. He concluded that the microstructure of these dinosaurs' bones indicates that their growth rate was much, much faster than that of crocodiles, and as fast as a baby ostrich." We had thought the subjects might find it harder to reject the data written in textbook fashion, but this difference in presentation style showed little or no effect in our data and so will be ignored in the remainder of this paper.

Rating scales. The subjects carried out three different types of rating tasks. They rated their belief in the theories, they rated their belief in the data, and they indicated how consistent they thought the data were with each theory.

The subjects indicated their degree of *belief in the theory* through the use of a 0 to 10 scale which stated that they were to use "0" for a "completely false theory" and "10" for a "perfectly correct theory."

The subjects indicated their degree of *belief in the data* through the use of a 0 to 10 scale where "0" was labeled as "false" and "10" was labeled as "true."

The subjects also indicated their belief in *data/theory consistency* through the use of a 0 to 10 scale. The scale asked the subjects to "Rate your belief in this assertion: The [X] theory is *consistent* with [data Y]" where X indicates one of the two theories and Y indicates one piece of data.

Procedure. Subjects read a theory or control passage. Those subjects who had received a theory passage were then asked to rate their belief in the theory they had just read and, after completing that rating, they were asked to rate their belief in the opposing theory. Next the subjects in the theory conditions were presented with a piece of data (consistent with the theory they had read about for half the subjects, inconsistent for the other half). They were then asked to rate their belief in the data they had just read about. After this the subjects were asked to indicate if the data were consistent with the theory they had originally read about and then, if the data were consistent with the opposing theory. Finally, they were asked to re-rate their belief in the original theory they had read about and their belief in the opposing theory.

The subjects in the control condition went through essentially the same procedure. They read the control passage and were asked to rate their belief in the theories (order of theory rated was counterbalanced). They then received data (half received data consistent with the warm-blooded theory and half data consistent with the cold-blooded theory) and were asked to rate their belief in the data. This provided a baseline on how believable the data (in isolation) were for our undergraduate population. Next the control subjects rated the consistency of the data they received with each of the theories, and finally they re-rated their belief in each of the two theories.

Subjects. There were a total of 72 undergraduate subjects in the experiment; 24 in each experimental condition.

Results

The first data of interest are the results of the subjects' initial ratings of belief in the theories that they had just read. These ratings allow us to see if the theory passages we developed were effective in convincing our subjects that the theory they read gave the correct account of dinosaur metabolism. In order to develop an overall measure of belief in the two theories, each subject's rating of the cold-blooded theory was subtracted from their rating of the warm-blooded theory. The results are given in Table 1.

Table 1: Initial Ratings of Belief in the Theory

Passage type	Belief rating[a]
Warm-blooded theory	-5.5
Cold-blooded theory	5.9
Control condition	1.5

[a]Cold-blooded rating minus warm-blooded rating, so positive scores show belief in the cold-blooded theory.

The results in Table 1 show that the subjects in each condition were fairly strongly convinced about the correctness of the theory that they had read about. What they did not know, of course, is that one third of the people in the room with them had just read a completely different theory and found it just as convincing! The subjects in the control condition showed a small tendency to prefer the cold-blooded theory.

The crucial data for this experiment are the subjects' ratings of their belief in the data that had been presented in the data passage. These results are presented in Table 2.

Table 2: Ratings of Belief in the Data

Theory/data relation	Belief rating[a]
Consistent	8.0
Inconsistent	6.0
Control	7.8

[a] rating scale from 0 to 10 with higher ratings indicating stronger belief in the data.

The data show that subjects give different ratings to the quality of the exact same piece of data depending on their current theoretical beliefs. An ANOVA shows that these differences are highly reliable ($F(2,66) = 7.292$, $p < .001$). When the data were inconsistent with the subject's current theory, the data were rated roughly 2 scale units lower than if the data were consistent with the subject's current theory ($p < .05$). Comparison with the control group shows that essentially all of the effect comes from the subjects with inconsistent data showing less belief in the data.

The final data of interest are the subjects' rating of the consistency of the data with the theory they had read. This rating was designed to see if the subjects in the inconsistent condition would perceive the data as less inconsistent with their preferred theory than would the subjects in the consistent condition. The data are presented in Table 3.

Table 3: Ratings of Data/Theory Consistency

Theory/data relation	Consistency rating[a]
Consistent	8.5
Inconsistent	7.2
Control	8.4

[a] rating scale from 0 to 10 with higher ratings indicatingthat the data are consistent with the theory.

There was a modest reduction in the subjects' beliefs in the consistency of the data with the theory for subjects receiving inconsistent data, but this difference was not reliable, $p > .08$.

Conclusions

The data from this experiment provide clear evidence that individuals' beliefs in a scientific theory can strongly influence their evaluation of scientific data. In particular the data show that individuals will tend to be critical of data that is inconsistent with their current theoretical beliefs. Our data are thus consistent with work in the philosophy of science that argues for the theory-laden nature of data and are also consistent with the Quine-Duhem thesis that anomalous data do not require the rejection of a scientific theory.

References

Brewer, W. F., & Lambert, B. L. (1993). The theory-ladenness of observation: Evidence from cognitive psychology. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 254-259). Hillsdale, NJ: Erlbaum.

Chinn, C. A. & Brewer, W. F. (1992). Psychological responses to anomalous data. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 165-170). Hillsdale, NJ: Erlbaum.

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.

Collins, H., & Pinch, T. (1993). *The golem: What everyone should know about science.* Cambridge: Cambridge University Press.

Duhem, P. (1914/1991). *The aim and structure of physical theory.* Princeton: Princeton University Press.

Frankel, H. (1988). Plate tectonics and inter-theory relations. In A. Donovan, L. Laudan, & R. Laudan (Eds.), *Scrutinizing science: Empirical studies of scientific change* (pp. 269-287). Baltimore, MD: The Johns Hopkins University.

Giere, R. N. (1985). Philosophy of science naturalized. *Philosophy of Science, 52*, 331-356.

Hanson, N. R. (1958). *Patterns of discovery.* Cambridge: Cambridge University Press.

Kuhn, T. S. (1962) *The structure of scientific revolutions.* Chicago: Chicago University Press.

Lord, C. G., Ross, L. , & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.

Maffie, J. (1990). Recent work on naturalized epistemology. *American Philosophical Quarterly, 27*, 281-293.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*, 161-175.

Pickering, M., & Monts, D. L. (1982). How students reconcile discordant data. *Journal of Chemical Education, 59*, 794-796.

Quine, W. V. O. (1951/1963). *From a logical point of view*. New York: Harper Torchbooks.

Thomas, R. D. K., & Olson, E. C. (Eds.) (1980). *A cold look at the warm-blooded dinosaurs*. Washington, DC: American Association for the Advancement of Science.