Philosophica 61 (1998, 1) pp. 39-50

# THE INFLUENCE OF ANOMALOUS DATA ON SOLVING HUMAN ABDUCTIVE TASKS

*Andreas Keinath & Josef F. Krems*

ABSTRACT

This paper describes an abductive process model of anomalous data integration. The model makes use of the entrenchment of the current explanation (amount of data explained) and the probability of alternative explanations. It is hypothesised that increasing confirmation of the anomaly itself increases the probability of alternative explanations. In an experimental study we found that both the entrenchment of an existing explanation and confirmation of the anomaly clearly influence how people resolve anomalous data. These results are in agreement with the predictions of the model.

## 1. Introduction

The integration of anomalous data in a current explanation or theory is an essential part of scientific discovery, trouble shooting and other everyday tasks like language comprehension (Kuhn, 1962; Kintsch, 1988, Thagard, 1989). In this paper we focus on the integration of anomalous evidence into an existing multicausal explanation and on abductive inference. Abduction refers to an inference to the best explanation by which from a given rule (A causes B) and a given observation (B) a hypothetical explanation (A) has to be inferred (Josephson and Josephson, 1994). Therefore, multicausal abductive explanations are explanations that are composed of multiple single causal hypothesis that together explain a set of observations. An anomaly occurs in this context if a new observation contradicts the existing explanation. This implies that data are presented in a sequential way, which is also usual in real life abductive tasks like medical diagnosis. The task for a problem solver results then in the resolution of the anomaly by changing parts of the existing explanation or by changing to a different explanation after the occurence of an anomaly.

The study was undertaken in the context of the abductive process model by Johnson and Krems (in press). We focused on the influence of the time of the anomaly occurrence and the amount of evidence confirming the anomaly itself as an indicator of the likelihood of alternative hypothesis. We hypothesised that anomaly resolution should be easier when an anomaly occurs early in the problem solving stage and when there are more observations that confirm the anomaly itself. On the contrary, the later an anomaly occurs and the less observations confirm the anomaly itself the harder it will be to resolve the anomaly. This paper describes the model framework in which the study was conducted and will also present empirical results.

## 2. Anomaly Resolution and Multicausal Explanations

In the case of abductive reasoning an anomaly occurs whenever new evidence contradicts one or more hypotheses in the current multicausal explanation. This is usually the way in many real life tasks in which the problem solver encounters new evidence in a sequential way. Because most explanations can be viewed as explanations which consist of compound hypothesis, the reasoner either has to modify his existing explanation so it is valid for the new and older evidence or select a new hypothesis that can explain the new evidence without contradicting the existing explanation. In our study we wanted to clarify the factors that influence the processes of hypotheses change or persistence.

It is widely accepted that anomalies play an important role in the case of hypothesis change. In philosophy of science Kuhn (1962) is the most prominent researcher who stressed the importance of anomalies as triggers for theory change. Other authors like Kulkarni and Simon (1988) or Darden (1992) use anomalies in their models of scientific discovery as triggers for hypothesis changes. An important feature of Darden's model is that he emphasises the confirmation of the anomaly itself as an important step in anomaly resolution. However, there was no implementation of this strategy in Darden's model.

Although there exist several models of scientific discovery and abductive reasoning, and the role of anomalies is viewed as an important feature of this process, most of them do not provide detailed processing of anomalous data. For example, Dunbar and Klahr's (1989) model (SDDS, Scientific Discovery as Dual Search) only provides a model of how explanations are formed and modified by searching in hypothesis and experiment spaces. However, it does not provide a description of what happens when new evidence or data contradicts the current explanation. On the other hand, Thagard's (1989) theory of explanatory coherence (TEC) offers an account on how anomalous data affects the strength (or coherence) of hypothesis but it ignores the sequential nature of abduction. Another problem of TEC and the corresponding process model implementation (ECHO) is the assumption that people can determine the coherence between all propositions in parallel (Cheng and Keane, 1989). This assumption seems unlikely for complex problem solving tasks, such as medical diagnosis or scientific discovery.

Studies on the interpretation of anomalous data provide evidence on more detailed factors but provide no model framework for these. Chinn and Brewer (1992, 1993) argue that the entrenchment of a current hypothesis is a crucial factor of how people respond to anomalous data. This means the more entrenched by data an existing explanation is, the more preserving behaviour favouring the initial explanation a person will show. However, their study does not reveal anything about how anomalies may lead to hypothesis change. The only important factor is the entrenchment of the existing explanation. These findings are close to the literature on the confirmation bias (e.g. Klayman and Ha, 1987) that indicate that people tend to focus on theory confirming evidence. Likewise, research on the primacy effect shows that with sequential processing of evidence early-presented information is overweighed in contrast to later-presented, probably inconsistent information (Nisbett and Ross, 1980). Other studies on contradictory evidence stress the importance of how the availability and likelihood of alternative hypotheses can influence the response to anomalous data (Burbules and Linn, 1988; Krems, 1994).

## 3. The Computational Model

Johnson and Krems (in press) developed a mental model based theory of abduction. Abduction is considered as the sequential comprehension and

integration of data into a single situation model that represents the current best explanation of the data.
The basic procedure for integrating new evidence in this model is as follows: Suppose that a new datum is available. First, the situation model is updated to include the new datum. Next, the new datum is comprehended, that is, knowledge is brought to bear to determine what the new datum implies about the situation. Comprehension results in one or more explanations for the datum, where each explanation consists of one or more hypotheses together with the data they explain. If the generated explanation is inconsistent with any hypothesis or data in the existing situation model, an anomaly has occurred and the model must be updated by either finding an alternative explanation for the new datum or by altering an explanation for the old data.

## 4. Processing Anomalous Data

The model responds to anomalous data by rejecting all but one of the anomalous explanations and then constructing alternative explanations for the data left unexplained by the rejection(s). It does this using a limited lookahead search to determine which explanation is the best to keep. In the lookahead search the model first selects one of the explanations, say e1, and rejects it. Then it searches for an alternative explanation for the data and evaluates the resulting situation model. Next it returns to the original anomalous situation, rejects e2, searches for an alternative to explain d2 and evaluates the results. The model then rejects the explanation whose rejection resulted in the best situation model (where best is defined as the model that explains the most data with the fewest number of explanatory components). If an alternative explanation for one of the data items cannot be found, then the explanation for that datum will be retained and the explanation for the other datum will be modified.

If rejecting the alternative explanations results in equally good situation models, then the difference between the probabilities (if known) to explain the evidence is used to break the tie. Thus, when deciding which explanation to reject, the model makes use of three factors: entrenchment of the current theory (the amount of data explained), the relative probability of the contradictory explanations, and the availability of alternative explanations. The more often a situation is faced in which the existing explanation is replaced by an alternative, the more likely it is that the person has generated an appropriate alternative for that explanation. This means that availability of potential explanations should increase with problem solving experience. It also means that subjects' confidence in their explanations should correlate with the relative frequencies that the explanations were correct for a set of data. These are the general predictions of our theory.

## 5. The Task

To study the human abductive problem solving we used a task called Black Box (BBX). In this tasks several atoms are hidden in a box (8x8 matrix), and the subjects goal is to discover their locations by shooting light rays into the box. The BBX device is shown in Figure 1(a). Each atom has a field of influence that is also invisible to the subjects. Our subjects are trained on the rules of how the light rays interact with the fields of influence. The light rays can be absorbed, reflected, or deflected by the fields of influence. The interactions of the light rays with the fields of influence according to the rules of the task can be seen in Figure 1(b). Because the atoms and the fields of influence are invisible, the only information a subject has to locate the atoms are markers surrounding the Box. The path of the rays through the black box has to be inferred by the different markers that indicate where a ray entered and left the box or that the ray was reflected or absorbed.

There are primarily three reasons for selecting the BBX task in studying human abductive inference. First, it is similar in many ways to real-world abductive tasks. These similarities include: 1) Sequential data processing, that means that new data must be collected and evaluated based on a current hypothesis. 2) The task allows the problem solver to form multicausal hypotheses that explain the data together. Second, the task is easy to understand. Our subjects learn the rules within one hour. Other studies indicate that even children are able to understand the task (Simon et al., 1987). Third, we are able to control the background knowledge of our subjects. By using such a simple domain we can ensure that all our subjects need to know about the task they will learn in our training phases. This is normally much more difficult in other domains like trouble shooting or scientific discovery (Chinn and Brewer, 1992).

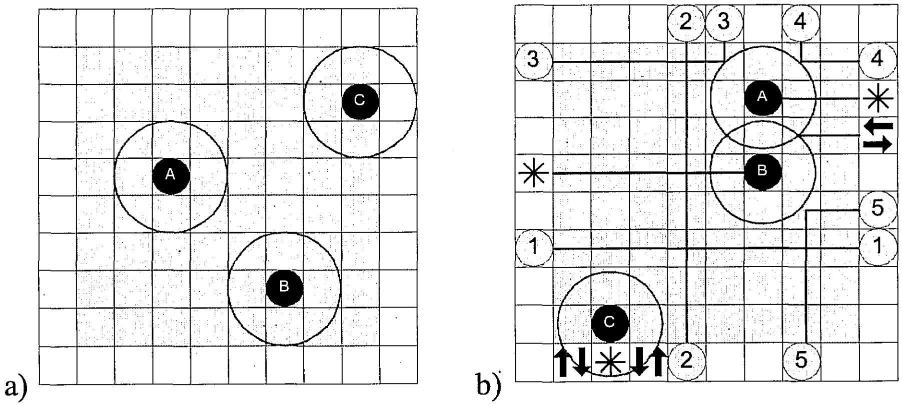


*Figure 1:* (a) The BBX device with 3 atoms and fields of influence; (b) The BBX devise with all possible interactions between light rays and fields of influence.

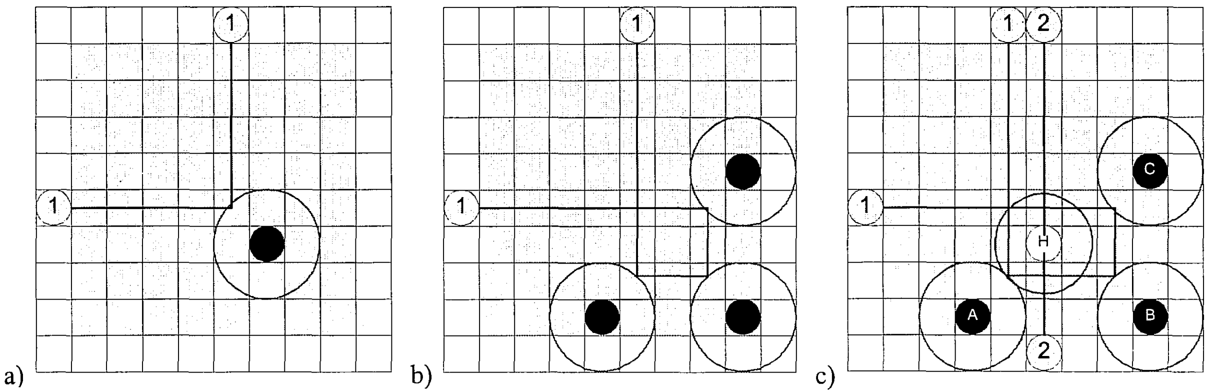


*Figure 2:* (a) simple pattern; (b) complex pattern; (c) anomalous data situation. Pattern (a) and (b) produce the same visible output.

## 6. Interpretation of Anomalous Data in the Black Box Task

A typical example of anomalous data in the BBX is illustrated in Figure 2. The situation in Figures 2(a) and 2(b) produce the same visible output for the subjects. Therefore, we are able to induce pattern 3a in our subjects as an explanation where pattern 2(b) represents the correct solution. An anomalous situation can be seen in Figure 2(c). If a subject chooses the simple pattern as the explanation for the light ray 1 by placing the atom H. the light ray with the number 2 leads to an anomalous situation. The second light ray can not go through the box if there is an atom at location H. This anomaly must be resolved by changing from the simple explanation of Figure 2(a) to the more complex explanation of Figure 2(b) that explains all the data.

The use of entrenchment in this task refers to the amount of rays that confirm the initial hypotheses before the anomaly occurred. Therefore, we constructed tasks in which all the evidence seemed to confirm the simple pattern as an explanation for the data until the anomaly occurred. On the other hand, the use of the probability of alternative hypotheses refers to the amount of evidence that confirms the anomaly itself. This is empirically evaluated by data that support the anomaly itself after its occurrence.

The model predicts that the later an anomaly occurs, the more entrenched the initial hypothesis is, and, therefore, hypothesis change is more difficult. On the other hand, the more evidence that confirms the anomaly, the more likely alternative explanations get, and, therefore, hypotheses change should be facilitated.

## 7. Experimental Evaluations: Design and Procedure

To test the predictions of the model we conducted an experimental study. 48 undergraduate students of the University of Regensburg solved a total of 192 tasks (without training).

A 2 (entrenchment of the existing explanation) $\times$ 3 (amount of evidence confirming the anomaly) between-subjects design was used. The factor amount of evidence had three levels, ranging from one to three confirming pieces of evidence for the anomaly itself. The factor entrenchment of the existing explanation was measured by the time of the anomaly's occurrence. This could be very early in the problem solving stage (anomaly as the second datum) or later in the problem solving stage (anomaly as the fourth datum).

Every session consisted of a training phase followed by a test phase to ensure that all subjects had the same level of knowledge and practice with the task. In the training phase subjects were trained on 25 tasks that could be solved by applying all previously learned rules. In the test phase subjects were presented experimental tasks followed by distracter tasks. In the critical tasks we tried to induce an initial but wrong explanation that would explain all data until the anomaly occurred. After the anomalous data was presented, the subjects had to resolve the anomaly by changing their explanation to solve the task correctly. The distracter tasks should prevent our subjects from learning the underlying patterns of the initial wrong hypothesis (simple pattern, see Figure 3a) and the true pattern that produced the anomalous situation (complex pattern, see Figure 3b).

The subjects were able to place and remove atoms in the black box and ask for new data. New data was requested by clicking on a "More Data" Button, which highlighted one of the perimeter cells. This indicated where the next light ray would enter the black box and by clicking on that cell the subjects were shown what had happened to this ray by markers. Thus the data were presented sequentially as if the subjects had actually shot the rays themselves. By still using predefined data-sets we were better able to control the situations presented to the subjects. Subjects had to place as many atoms as were hidden in the black box to finish a task. This means that they still could finish the task without solving it but should prevent them from hurrying through the tasks without thinking about the problems. All atom placements, removals, and data requests were time-stamped and recorded.

## 8. Results

### 8.1. Inducing an Initial Explanation

Inducing an explanation that can be contradicted by anomalous data is crucial for the study of anomalies. This initial wrong hypothesis can then be contradicted by anomalous data so subjects have to change their existing explanation. In our case we looked at every experimental task that was performed by our subjects and checked if they had placed the atom indicating the use of the simple pattern as an explanation before the anomaly occurred (H-atom, see figure 3c). This initial explanation could not be found in only one task, so we eliminated it from our further analysis while all other tasks could still be used. This is a strong indication that we were able to induce a wrong initial explanation that could be contradicted by our anomaly.

### 8.2 Amount of Evidence and Time of Anomaly Occurrence

The model predicts that subjects should find it harder to solve tasks that are more entrenched before the anomaly occurs. On the other hand, anomaly resolution should be facilitated the more evidence is presented for the anomaly itself. This means our subjects should solve more tasks correctly if the anomaly is presented early in the problem solving stage, and the amount of solved tasks should increase for both conditions the more evidence confirms the anomaly (see Figure 3). In a Hierarchical Log-linear analysis we revealed both significant effects of our main factors but no interactions. This means that we have a significant effect of *anomaly occurrence* × *amount of solved tasks*, $\chi^2$ (1)= 19.26, $p < 0.01$ and a significant effect of *amount of evidence confirming the anomaly* × *amount of solved tasks*, $\chi^2$ (2)= 15.69, $p < 0.01$.

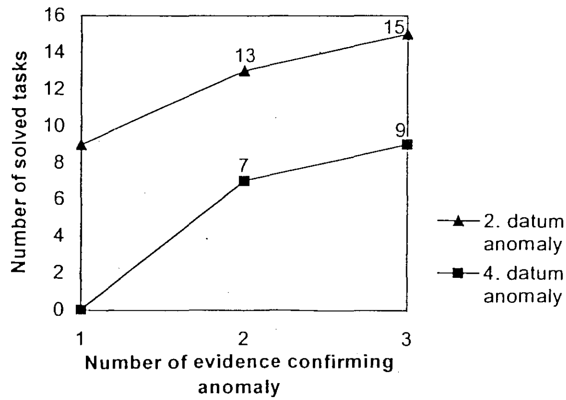


*Figure 3:* Influence of confirming evidence for anomaly and time of anomaly occurrence on the amount of correctly solved tasks.

Furthermore we could identify in a Wilcoxon Test that there are significant different amounts of tasks solved between both levels of anomaly occurrence over all levels of anomaly confirmation (see Table 1).

Table 1: *Differences between second datum and fourth datum of anomaly occurrence over all levels of anomaly confirmation.*

| | Amount of anomaly confirmation | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Wilcoxon *W* | 192 | 216 | -216 |
| *z* | -3.48 | -2.15 | -2.41 |
| *p* | .000 | .031 | .016 |

## 9. Discussion and Conclusions

The analysis of the data showed clearly that in nearly all cases we induced successfully an initial explanation that could later be contradicted by an anomaly. Therefore, we had an experimental controlled situation of anomaly occurrence.

With respect to our hypothesis we obtained data that confirmed the model predictions. The amount of evidence and the time of anomaly occurrence had a clear influence on how people responded to anomalous data. The sooner the anomaly occurred or the less entrenched an existing explanation was, the easier our subjects found it to change their hypothesis. In contrast, it was much harder to change the existing hypothesis the later the anomaly occurred and the more entrenched this initial hypothesis was. This is consistent with the literature on cognitive biases (Klayman and Ha, 1987) and the findings of Chinn and Brewer (1992, 1993). On the other hand, hypothesis change was facilitated the more data confirmed the anomaly itself, or the more entrenched the anomaly itself was. We assume that confirmation of the anomaly itself increases the likelihood of alternative hypothesis and this helps the subjects to change their explanation. This explanation is consistent with the assumptions of the model in processing anomalous data. We therefore believe that confirmation of the anomaly dissipates the cognitive bias effect of entrenchment by the initial explanation. In other words, the more entrenched the anomaly gets the weaker the influence of the entrenchment of the initial explanation.

In order to contradict a well entrenched explanation and force subjects to a hypothesis change it seems to be important to increase the likelihood of alternative explanations by increasing the entrenchment of the anomaly. Finally, we can say that the use of the entrenchment of an existing explanation and the likelihood of alternative explanations are reasonable and empirical confirmed components of the abductive process model by Johnson and Krems (in press).

University of Technology Chemnitz

**REFERENCES**

Burbules N.C. & Linn M.C. (1988), Response to contradiction: Scientific reasoning during adolescence, *Journal of Educational Psychology,* 80:67-75.

Cheng P.C-H. & Keane M. (1989) Explanatory coherence as a psychological theory, *Behavioral and Brain Sciences*, 12(3):469-470.

Chinn C.A. & Brewer W.F. (1992), Psychological responses to anomalous data, in: *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum Associates, Hillsdale.

Chinn C.A. & Brewer W.F. (1993), Factors that influence how people respond to anomalous data, in: *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society,* Lawrence Erlbaum Associates, Hillsdale.

Darden L. (199)2, Strategies for anomaly resolution, in: *Minnesota Studies in the Philosophy of Science: Vol. 15. Cognitive Models of Science,* R.N. Giere, ed., University of Minnesota Press, Minneapolis.

Dunbar K. & Klahr D. (1988), Dual space search during scientific reasoning, *Cognitive Science,* 12(1):1-48.

Johnson T.R. & Krems J.F. (in press). Context use in multicausal abductive reasoning. *Cognitive Science*

Josephson J.R. & Josephson S.G. (1994), *Abductive Inference*. Cambridge University Press, Cambridge.

Kintsch W. (1988), The role of knowledge in discourse comprehension: A construction-integration model, *Psychological Review*, 95 (2):163-182.

Klayman J. & Ha Y.-W. (1987), Confirmation, disconfirmation and information in hypothesis testing, *Psychological Review*, 94 (2):211-228.

Krems J. & Johnson T.R. (1995), Integration of anomalous data in multi-causal explanations, in: *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society,* J. D. Moore and J. F. Lehmann, eds., Lawrence Erlbaum Associates, Mahwah.

Kuhn T.S. (1962), *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago.

Kulkarni D. & Simon H.A. (1988), The process of scientific discovery: The strategy of experimentation, *Cognitive Science*, 12:139-175.

Nisbett R. & Ross L. (1980), *Human Inference: Strategies and Shortcomings of Social Judgement*, Prentice-Hall, Englewood Cliffs.

Simon T., McShane J. & Radley S. (1987), Learning with microcomputers: Training primary school children on a problem-solving program, *Applied Cognitive Psychology*, 1:35-44.

Thagard P. (1989), Explanatory coherence, *Behavioral and Brain Sciences*, 12 (3):435-502.