Schunn, C. D., & Klahr, D. (1996). *The problem of problem spaces: When and how to go beyond a 2-space model of scientific discovery*. In the Proceedings of the 18th Annual Conference of the Cognitive Science Society.

The problem of problem spaces: When and how to go beyond a 2-space model of scientific discovery

Christian D. Schunn and **David Klahr**
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
{schunn, klahr}@cmu.edu

Adding Problem Spaces — The Issues

The process of scientific discovery has been characterized as a search in two problem spaces: the space of possible experiments and the space of possible hypotheses (Klahr & Dunbar, 1988). More recently, there have been several proposals to include additional problem spaces. In particular, we have proposed the addition of a space of experimental paradigms and a space of data representations (Schunn & Klahr, 1995). These proposals raise meta-theoretical issues: (a) Why these spaces? (b) Why *only* these spaces? Moreover, these issues are not specific to models of scientific discovery—they are general to all problem solving activities.

In our talk, we propose a general set of logical, empirical and implementational criteria for resolving these issues. We illustrate this distinction in the context of distinguishing between the data representation and hypothesis spaces in scientific discovery, focusing on results of psychological labs studies of discovery activities. Before presenting the general criteria, we describe a simple, concrete example which we believe provides an intuitive feel for when distinct problem spaces do and do not occur.

Consider subjects trying to solve simple letter series completion tasks (cf., Simon & Kotovsky, 1963), such as ABMCDM_. Subjects will initially select a representation that involves relations on the English alphabet, and they will seek patterns of sames, differents, nexts, double nexts, priors, etc. In this case, the representation is so obviously immediate and evoked by the stimulus that it makes no sense to view it as involving any search.

Now consider what happens when subjects get a trick problem, such as: OTTFFSS_. Here too, they start with the obvious representation in which the letters are simply letters. But leads subjects to formulate increasingly complex and ad hoc explanations for the O at the beginning of the series. Ultimately, subjects using this straightforward representation hit an impasse. Then they start to consider whether the letters stand for something else (like the first letters of the names of the integers).

At this point, we argue, they are searching a data representation space. That is, they must consider other ways to characterize the items of the list. In the present instance, they might reason that, since the earlier lists are based on the alphabet, and the alphabet is a (the most) familiar ordered list of symbols, then there might be other ordered lists that are relevant: days of the week, names of the Kings of England, Presidents of the US, number names, months, etc. Perhaps the letters in the sequence are related to some feature of these other lists (in this case the first letter of their English names of the integers.)

General Criteria

How does one decide whether to add a new problem space? Below we list three kinds of criteria that 1) we found very useful in making the distinction, and 2) we believe can be applied to understand problem-solving behavior in any task.

Logical Criteria

First, one needs to be able to define the spaces such that they are unambiguously different. Furthermore, the distinction between spaces should be categorical rather than a continuum. There are two important components to the definition of a problem space: the goals used for searching the space, and the entities that are searched. Distinct problem spaces should involve distinct entities and distinct goals. Of course, the problem spaces can be coordinated such that information from one space is used to search the other space—this feature is ubiquitous to the scientific discovery tasks that we have considered.

Empirical Criteria

Second, there should be empirical differences between the spaces. In particular, they should occur (at least occasionally) at different times, they should involve different search heuristics, and there should be different factors that influence behavior in each search space.

Most importantly, there should be activity in each search space. If a subject always stays in the same state with respect to a particular problem space, then that search space is not a useful tool for describing that subject's behavior. This might occur in two ways. First, the subject could know so little about a domain that only one state is available to that subject. Second, subjects might be so knowledgeable about a domain that they are able to pick a good state immediately and need not search any further. Thus, the existence of a particular problem space for a given subject is determined, in part, by their knowledge.

Implementational Criteria

Third, one should be able to represent the problem spaces distinctly in a computational model that is capable of performing the task. In other words, one should be able to map unambiguously from observed behavior to a state in the problem space. Furthermore, search through a specific problem space should be distinct from coordination between problem spaces. This provides an exacting test of the degree to which the theoretical distinctions can actually account for the empirical behavior at a precise level.

An Example

Having described the general criteria, we will now describe how we have applied these criteria in proposing our 4-space model. However, in this short abstract, we are only able to explore the addition of the data representation space.

Logical Criteria

First, there are the entities that are searched. The hypothesis space involves propositions about the world, potentially at varying levels of abstraction and of varying levels of generality. For example, one might have the following hypotheses—varying in scope and abstraction—about the results of an experiment on reading comprehension: "there was no effect of manipulating motivation in this experiment (on comprehension of some history text)", "there is generally no effect of monetary manipulations of motivation on comprehension of history texts", or "there is generally no effect of motivation on reading skills".

By contrast, the data representation space involves the objects and object features of the data. For example, one might graph different the relationship between different variables, one might use different kinds of graphs, and one might re-code, collapse, or expand the same data.

The two spaces also differ in terms of their goals. The goal of the hypothesis space is to produce parsimonious explanations/descriptions of objects and relations in the world. By contrast, the goal of the data representation space is to find regularities. A data representation is abandoned if it doesn't lead to regularities or interpretable patterns, whereas it is maintained when it does. In other words, the search in the data representation space seeks to find the regularities, and search in the hypothesis space seeks to explain them.

Note that there are not separate search spaces for hypotheses at different levels of abstraction because 1) this distinction is a continuum; and 2) the goals of the search at different levels is the same. Similarly, there are not different search spaces for tables versus graphs, or internal versus external representations.

Empirical Criteria

The particular data that lead us to propose the data representation space came from a complex discovery microworld called MilkTruck (Schunn & Klahr, 1992, 1995). Subjects in this task were asked to discover the role of a complex, mystery function by conducting experiments in the domain. What is the empirical evidence from this domain for the data representation space?

First, the majority of the subjects radically altered their representation of the task over the course of discovering the correct function. Based on their verbal think-aloud descriptions of the experiments and experiment outcomes, we found that the subjects changed both the kinds of objects they were examining (new objects and new levels of aggregation of the same objects) as well as which features of the objects the described. Interestingly, we found that the subjects used multiple representations of the same experimental outcome for most experiments. Moreover, we found that, from one experiment to the next, subjects made at least minor changes to the data representation over 90% of the time. This can be contrasted with performance in some simpler tasks, in which people do not need to search the data representation space since there is usually only one way to represent the data.

Second, we found that search in the two spaces did occur at different times. Occasionally, subjects considered several different representations of the data without coming up with a new hypothesis. Similarly, they occasionally considered several different hypotheses all within the same representation of the data.

Third, we found several new heuristics that subjects used in producing data representation change. For example, we found that subjects used a Notice Invariants heuristic (cf., Kaplan & Simon, 1990) in which they recast their representation of the data into the parts of the data that remain constant across experiments and the parts that vary.

Implementational Criteria

While we have not yet completely implemented a running model, in coding the data, we have developed a sense of how easy/difficult it is to unambiguously distinguish between hypothesis and data representation, and it has forced us to revise some previous distinctions that we had initially made only on the basis of logical criteria.

Furthermore, in the symbolic production system schemes that we have constructed, changing data representations is always an important issue that is clearly distinct from selecting among and constructing novel hypotheses.

Conclusions

We have outlined three types of criteria (logical, empirical, and implementational) that we believe are important in determining the number of problems spaces that should be used in describing problem solving behavior in a particular task. In this brief abstract, we could only describe the application of these criteria to justifying the addition of the data representation space in characterizing scientific discovery behavior. However, we have also been able to use these criteria to propose the experimental paradigm space, and we believe these criteria can be applied to resolve these kinds of issues in any problem-solving task.

References

Kaplan, C. A., & Simon, H. A. (1990). In Search of Insight. *Cognitive Psychology*, **22,** 374-419.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science,* **12(1),** 1-55.

Schunn, C. D., & Klahr, D. (1992). Complexity Management in a Discovery Task. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Schunn, C. D., & Klahr, D. (1995). A 4-space model of scientific discovery. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review,* **70,** 534-46.