## Title
Does Meta-space Theory Explain Insight?



## Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

## Authors
Wolf, David F.
Beskin, Jonathan
Dietrich, Eric

## Publication Date
1995

Peer reviewed

# Does Meta-space Theory Explain Insight?

**David F. Wolf II**
Philosophy Department
SUNY at Albany
Albany, NY 12222
dw4570@cnsunix.albany.edu

**Jonathan Beskin**
**Eric Dietrich**
PACCS Program in Philosophy
Binghamton University
Binghamton, NY 13902-6000
noj@turing.pacss.binghamton.edu
dietrich@crow.pacss.binghamton.edu

**Abstract**

Previous computational theories of problem solving have not accounted for the occasional display of accelerated problem solving by humans working on conceptually hard problems. Researchers refer to this behavior as *insight*. Kaplan and Simon describe insight as the selection of a good representation of the problem by the problem solver. They propose a dual-state space theory, meta-space theory, to explain insight (Kaplan and Simon, 1990). We show that meta-space theory is unfalsifiable. We then show that the nature of meta-space theory makes it superfluous for the study of human problem solving.

## Introduction

Traditional heuristic search space theories cannot account for *insight* in creative problem solving (Boden, 1988). 'Insight', is the accelerated rate of problem solving that occurs once the problem solver selects a good representation (Kaplan and Simon 1990). To explain how humans select these representations, Kaplan and Simon employ a dual search space schema. Their dual-space architecture consists of a meta-level space whose nodes are themselves possible problem spaces. The meta-level space is also a type of problem space, so we begin with a brief review of heuristic search space theory. We then breifly discuss the importance of representations in creative problem solving. Finally, we argue that search space theory is unfalsifiable and hence unnecessary as a theory of human problem solving.

## Heuristic Search Space Theory

Researchers in the paradigm of heuristic search space theory think of problems as *problem spaces*. The 3-tuple **<O, G, I>** describes a problem space where:

- **O** = a set of operations. These specify the rules that enable a state to be transformed into another state.
- **G** = a set of goal states. These states are acceptable as a solution to the problem.
- **I** = a set of initial states. These states describe all the possible starting situations for the problem solving process.

A problem space consists of all the possible states that the operations can generate from the initial states. A problem space is like a mathematical tree in which the root node is the initial state, the nodes are possible states, and the leaves are states on which no further operation apply. The set of goal states, **G**, is a subset of those leaves.

Heuristic search space theory's basic assumption is that problem solvers search problem spaces. Starting at an initial state, problem solvers search for a goal state. The 3-tuple **<D, O, C>** defines a problem solver where:

- **D** = a knowledge base that consists of one or more data bases with information appropriate to a particular task and about the desired goal states. The primitives to create spaces for a search are also in the knowledge base.
- **O** = a set of operations. These rules are the options the problem solver may perform on any given state. These are the same operations as in the definition of the problem space.
- **C** = a control strategy. This selects what operator to apply to the current state. It does this by assigning values to the possible states and selects the state that has the best value assigned to it. The control strategy also resolves conflicts that may occur from the selection.

The problem solver creates a *search tree* to find a path from the initial states to one of the goal states (see Figure 1). It accomplishes this by applying operators to the current states. This determines all the possible states that could result from the current one. The control strategy, **C**, rates the desirability of these possibilities. The state with the most favorable rating will determine which operator, and thereby which state, to select next.

An important aspect of heuristic search space theory is the notion of *heuristics*. Heuristics are imperfect problem solving processes that often work. Although heuristics do not guarentee a solution, they can solve problems that are otherwise intractable. For example, the average game of chess consists of approximately forty moves by each player. Its problem space consists of $10^{118}$ states. A million computers searching a million nodes a second would still require $10^{100}$ years to search the problem space chess creates. Problem spaces like these must be pruned to make

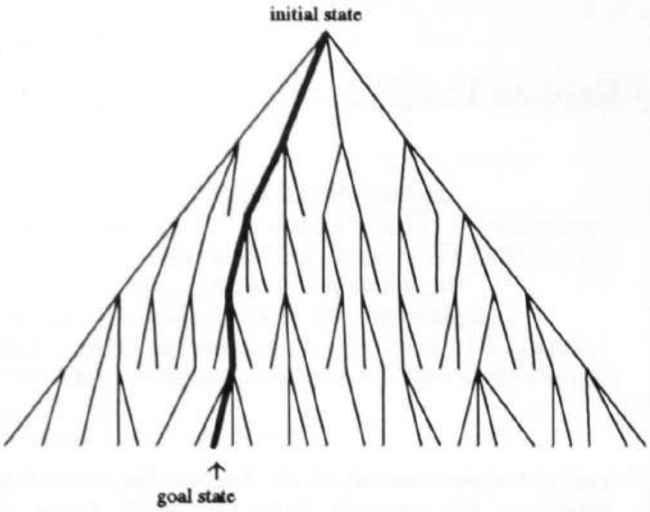


Figure 1: Above is a small problem space. The darkened lines identify the search tree, which traces a portion of the problem solver's path. In this figure the search tree was created by a depth-first search procedure.

them tractable. Heuristics exist in two places within the problem solver. They are explicit within the control strategy, C, and are implicit within the set of operations, O. These 'rules of thumb' help the problem solver streamline its search by eliminating undesirable options from the search.

Traditionally, each problem space designated one way that a problem solver could characterize a problem. Real world problem solvers often use multiple representations to solve problems. For example, when faced with problems that would require an unreasonable amount of search time, people first attempt to change the problem's representation (Sinnott, 1989). The reason for this is that an appropriate representation emphasizes relevant features of the problem, and can drastically reduce the time to solve the problem (Kaplan and Simon, 1990). However, this is a secondary feature of representations. For many problems, only a small number of representations provide solutions. Furthermore, some very hard problems cannot be solved without the use of several representations (Minsky, 1991). Any flexible problem solving system must be capable of generating multiple representations.

Heuristic search space theory does not account for representations in problem solving. Instead, Newell and Simon state that the theory explains problem solving after a representation is selected (1972). Later they suggest that problem solving occurs in one large problem space that incorporates all the representations of the problem (Newell and Simon, 1976). However, this explanation of representations cannot account for the ability to rapidly change representations. Previous research suggests that complex problems require dual search spaces (Simon and Lea, 1974). In dual space frameworks, one space contains possible hypotheses, and each hypothesis is itself a search space for investigating the hypothesis. Kaplan and Simon (1990) rely on this dual-space technique to account for the use of multiple representations in human problem solving.

## Meta-space Theory

Kaplan and Simon's dual-space architecture consists of a meta-level space and a set of representation problem spaces. They argue that insight occurs in the meta-space.[1] The meta-space is a problem space where each node is a unique representation of the problem in the form of a problem space. The problem solver in meta-space theory behaves similarly to its counterpart in heuristic search space theory. It traverses the meta-space searching for a representation that will solve the problem. Meta-heuristics improve search by pruning the meta-space. The four currently hypothesized meta-heuristics are: Try a Switch, Notice Invariants, Form Hypothesis, and Compare Alternatives (Kaplan and Simon, 1990, p. 381, 402).

Once the problem solver selects a node in the meta-space it enters the search space that that representation creates. The problem solver now searches this problem space for a solution to the problem. If no operators within the problem space seem to lead to a successful solution, the problem solver abandons the representation.[2] It then continues its search within the meta-space (see Figure 2).

A meta-space that could account for any possible representation would be infinite and thereby not computable by finite beings in a finite amount of time. According to Kaplan and Simon (1990), humans "are not equipped with generators for searching the space of 'all possible representations'" (p. 403). Rather, humans have a subset of these generators. Even so, this smaller set of generators must still be vast to accomodate all the representations that humans might use. At the moment, the boundaries of this set are unknown, although some representations and meta-heuristics are known to be used for some problems (Kaplan and Simon, 1990).

## Problems with Meta-space Theory

We claim that meta-space theory is unfalisifiable. This means that no possible data about how humans solve problems can falsify it. Given time to analyze a subject's solution to a problem, meta-space theory can always provide an explanation of that solution. Therefore, meta-space theory is compatible with any data, no matter how surprizing, about the way humans solve problems. Hence, meta-space theory makes no predictions, and rules out no possibilities. It is therefore vacuous; it says nothing about human problem solving. We argue that to make meta-space theory useful to psychology, Kaplan and Simon require a theory of problem solving independent of the meta-space framework.

[1] Kaplan and Simon refer to the second space as the meta-level space. For brevity, we refer to it as the *meta-space*.

[2] Kaplan and Simon have not fully defined this procedure. They do suggest that the amount of time spent in the problem space is often a relevant factor (Kaplan and Simon, 1990: p. 377).

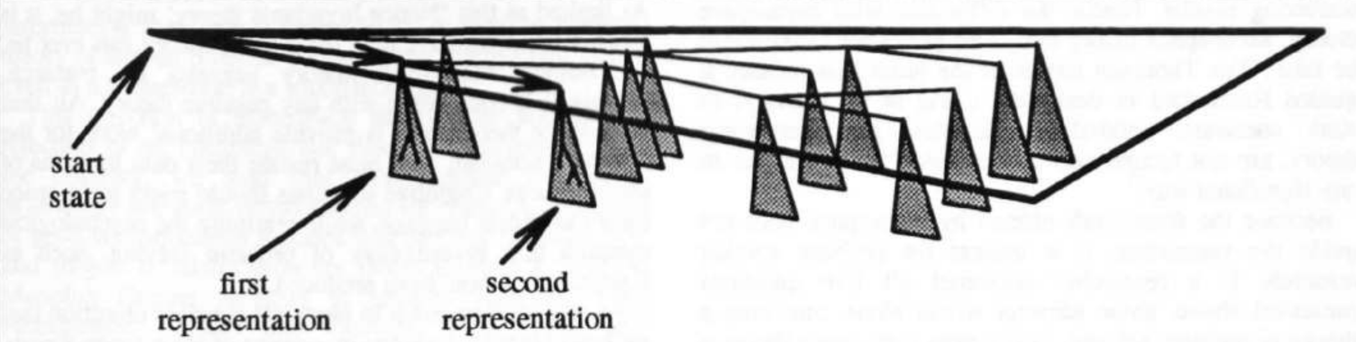


Figure 2: The meta-space is a problem space that contains all possible representations of a given problem. It is a search space with each node containing a search space that represents the problem differently. The problem solver searches the meta-space for a representation. When it selects a node, it enters the search space that that representaion creates and searches for a solution. If the operators of this search space do not seem to solve the problem, then the problem solver returns to the meta-space to search for another representation.

A good theory of problem solving should answer five questions:

1) How do subjects think about problems, *i.e.*, what mental representations do subjects use?
2) How do subjects produce and select representations?
3) How do subjects use selected representations to solve problems?
4) How do subjects decide when a representation is not working, and what do they then do?
5) What information is retained from representation to representation?

Meta-space theory fails to answer any of these questions. At best, meta-space theory must leave these questions open for future research. In short, meta-space theory is best described as a *framework* for studying problem solving, not a theory of problem solving. This framework is reducible to a single simple tenet; *viz.* problem solving is the process of searching for a representation in a meta-space, and then searching selected representations for a solution.

Kaplan and Simon (1990) implicitly agree that meta-space theory is a framework for studying problem solving. Rather than claiming that meta-space theory is a completed theory, they advocate a research program of exploring the representations and meta-heuristics that subjects use to solve problems.

> "[We] must discover how subjects constrain their search for a new representation when the initial one does not suffice, and how the new representation constrains their search for the solution. We must also discover what triggers subjects to shift from searching for a problem solution to searching for a better problem space in which to conduct the solution search." (Kaplan and Simon, 1990: p. 377)

The information that Kaplan and Simon say researchers still need to discover is the same information that constitutes a theory of problem solving. Thus, a researcher can rewrite any theory of problem solving in terms of meta-space theory, including a theory that does not allow for representational change at all. To see this, note that one could characterize such a theory as a meta-space with only one node. This universal compatibility makes meta-space theory unfalsifiable, and therefore vacuous.

Some might think meta-space theory is not unfalsifiable, but merely flexible and that this flexibility is a strength. One could argue that Kaplan and Simon left the framework of meta-space open to accommodate any possible future research. However, this is not a strength, but a weakness. In an attempt to make meta-space theory general enough to explain all problem solving, it fails to explain any problem solving. A researcher can account for any experimental results, no matter how surprising, within the meta-space framework by inserting the appropriate *ad hoc* representation or meta-heuristic. There is nothing in the theory to stop a researcher from adding a new meta-heuristic for each observed anomaly in the data as we unravel how humans solve problems.

A theory that does not constrain research also does not provide any guidelines for future exploration. A fecund, productive, theory suggests some avenues of research, and denies others. A theory's success depends on how well these avenues predict observable behavior, in contrast to how well other avenues predict the same behavior. As an example of a productive theory, consider the case of Rutherford and his "mini solar system" model of the atom. Rutherford was a student of J. J. Thomson. At the turn of the century, Thomson had the best model of the atom, called simply the Thomson model. This model was also referred to as the plum-pudding model, since it viewed the atom as a large positively charged ball with negatively charged electrons stuck in it, like plums in a pudding. The Thomson model predicted that alpha particles shot at an atom would not be scattered or deflected from their original path, but would instead pass right through. Rutherford, Geiger, and Marsden performed this experiment in 1909 and discovered, to their amazement, that the alpha particles were scattered every which way. The Thomson model had to be false, some other model had to be correct. This was when Rutherford started developing his solar system model, or nuclear model, which concentrated most of the mass of the atom in the nucleus. Such a model explains the scattering results. Notice the difference with meta-space theory. Meta-space theory makes no predictions that might be false. The Thomson model of the atom was fecund; it guided Rutherford in designing a test of its veracity. In stark constrast, unfalsifiable theories, like meta-space theory, are not fecund and do not guide the researcher in any significant way.[3]

Because the framework offered by meta-space does not guide the researcher, it is useless for problem solving research. If a researcher answered all five questions presented above, those answers would alone constitute a theory of problem solving. This makes meta-space theory a useless shell, *even as a framework*. This is further evidence that meta-space theory is unfalsifiable. Anything that a researcher might say about problem solvers and their strategies can be said irrespective of whether or not one adopts the meta-space architecture. The Notice Invariants heuristic is an excellent example. This heuristic informs subjects to pay attention to features of the problem that do not change between representations (Kaplan and Simon, 1990: p. 404ff). Kaplan and Simon show that the subjects that solved the mutilated checkerboard problem quickly were ones who noticed relevant perceptual invariants sooner. These subjects also noticed a wider range of invariants. More specifically, the data indicates that fast subjects usually noticed that a domino covers squares of different types (one black and one white), while the two missing squares are of the same color. Kaplan and Simon refer to this feature of the mutilated checkerboard problem as *parity*.

The results of the Notice Invariants experiments are important to the understanding of general problem solving regardless of the status of meta-space theory. In fact, the Notice Invariants heuristics could have been proposed by psychologists who never thought in terms of search spaces, but rather only in terms of constructed representations and representational change. This does not logically require any version of search space theory.[4] The hypothesis that noticing invariants is important to solving problems suggests future research. One possible implication of the data is that it appeared to be more important for subjects to notice invariants between representations than to use any specific representation. Noticing the parity feature almost invariably led to representations that used this feature to prove the problem impossible. One viable interpretation is that the solution representation was merely a vehicle that the subject used to highlight the parity feature. If this is the case, then subjects that have not noticed the importance of the parity feature should have a hard time solving the problem even if they are told which representation to use. Notice that the research on the Notice Invariants heuristic immediately implies additional experiments and hypotheses to test. This line of research constitutes a potential theory of problem solving without any need for meta-space theory. As limited as this 'Notice Invariants theory' might be, it is much more productive than meta-space theory can ever be. In contrast, meta-space theory suggests no research, because it is compatible with any possible theory. All that meta-space theory does is provide additional work for the cognitive scientist, who must restate their data in terms of search spaces. Cognitive scientists should reject meta-space theory as excess baggage, while retaining the psychological research that investigators of problem solving, such as Kaplan and Simon, have produced.

At this point we wish to block the possible objection that we have ignored evidence in support of meta-space theory. Kaplan and Simon (1990: p. 412) claim that search space theory, "is supported by a substantial body of empirical data". Their evidence for this claims comes from interpreting all occurrences of changing representation as search within a meta-space (Kaplan and Simon, p. 376-7). If, as we argued earlier, meta-space theory is unfalsifiable, then no evidence can confirm it. No amount of positive evidence can ever absolutely confirm any theory. Rather, one confirms a theory by testing it against rival theories (Aronson, 1984: p. 377ff). Rival theories cannot both be true over the same body of data, so an experiment confirms one theory in part by disconfirming an alternate theory.[5] To make this work, the theories must make opposite predictions with regard to the experiment. Depending on the outcome of the experiment, one theory is correct and the other is wrong. This means that an experiment confirms a theory only if, by failing, it could potentially disconfirm the theory. A theory that was compatible with both outcomes of an experiment would not be confirmed by either outcome. By definition, an unfalsifiable theory is compatible with all potential outcomes. Therefore, no amount of evidence could ever confirm an unfalsifiable theory, such as meta-space theory. Kaplan and Simon present data that is compatible with meta-space theory, not confirmation of the theory.[6]

## Conclusion

We showed that the meta-space framework is unfalsifiable. Fundamentally, meta-space theory is a theory of *selecting* representations, when what we need is a theory of how representations are constructed during problem solving. Cognitive science should therefore jettison it from research on human problem solving. Such theories do not inform psychologists, and usually only serve to cloud the issues. The only way to make the theory useful for psychology would be to develop a new theory of problem solving independent of the meta-space framework. The

[3] For more on the importance of fecundity, see Kitcher (1992: pp. 47 - 49).

[4] A deeper point here is that the researcher could also have produced the Notice Invariants heuristic by analogy to heuristic search space theory without ever using meta-space theory.

[5] Already we note a problem as meta-space theory cannot have any rival theories. It is compatible with all potential theories of problem solving.

[6] Note that while there is evidence in support of the various meta-heuristics that Kaplan and Simon (1990) investigate, this does not constitute evidence in favor of meta-space theory. Individual meta-heuristics are falsifiable, and therefore can be confirmed by experiment. Meta-space theory is not falsifiable, and cannot be confirmed.

information meta-space theory lacks is itself a complete theory of human problem solving. Thus, meta-space theory, even as a framework, is a superfluous shell.

This rejection of meta-space theory does not entail a rejection of all computational theories of psychology. On the contrary, we hold that computational theories of psychology are good methodologies for research. Examples of good computational theories that are highly informative and subject to falsification are Dedre Gentner's Structure Mapping Engine program in analogy (Gentner 1983; Falenhainer, Forbus, and Gentner, 1986, 1989) and Stephen Kosslyn's work on mental imagery (Kosslyn, 1994). These approaches are falsifiable, informative, and consequently important to the field of psychology. We reject meta-space theory, not because it is computational, but because it is useless for the study of problem solving.

## References


Barr, A. and Feigenbaum, E. (1981). *The Handbook to Artificial Intelligence*. vol. 1. Los Altos: William Kaufmann Inc.

Boden, M. (1988). *Computer Models of Mind*. Cambridge. Cambridge University Press.

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1986). "The structure-mapping engine". *Proceedings of the American Association for Artificial Intelligence* (pp. 272-277), Philadelphia.

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1989). "The Structure-Mapping Engine: algorithm and examples". *Artifical Intelligence* 41: 1-63.

Gentner, D. (1983). "Structure-mapping: A theoretical framework for analogy". *Cognitive Science* 7:155-170.

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

Kitcher, P (1992). *Abusing Science: The Case Against Creationism*, Cambridge, MA: MIT Press.

Kosslyn, S. (1994). *Image and Brain*. Cambridge, MA: MIT Press.

Minsky, M. (1991). "Logical verses Analogical or Symbolic verses Connectionist or Neats verses Scruffy". *AI Magazine* 12: 35-51.

Newell, A. and Simon, H. (1972). *Human Problem Solving*. New Jersey: Prentice-Hall.

Newell, A. and Simon, H. (1976). "Computer Science as Empirical Enquiry: Symbols and Search". *The Tenth Turing Lecture*, first published in *Communication of the Association of Computing Machinery* 19.

Simon, H. and Lea, G. (1974). "Problem Solving and Rule Induction: A Unified View" in *Knowledge and Cognition*. Gregg, L (Ed.). Hillsdale: Lawrence Erlbaum Associates.

Sinnott, J. (1989). "A Model for Solutions of Ill-structured Problems: Implications for Everyday and Abstract Problem Solving" in *Everyday Problem Solving*. Sinnott (Ed.). New York: Preager Publishers.