Individual Differences in Working Memory as Strategy Selection: Implications for Computational Models of Scientific Discovery[1]

Eric G. Freedman
Psychology Department
University of Michigan-Flint
Flint, MI 48502
email: freedman@umich.edu

Abstract

In contrast to storage-based models, the present paper attempts to explain the role of working memory in scientific reasoning in terms of the strategic functions of the executive control system. According to this approach, strategies are selected to achieve the individual's goals. A strategic approach provides accounts for the findings of the research on scientific discovery. Furthermore, a strategic approach has implications for a variety of areas of cognitive science.

Introduction

Although it is generally accepted that working memory influences performance in a variety of domains including reasoning and problem solving, most researchers (Engle, Cantor &, Carullo, 1992; Cowan, 1988) believe that individual performance differences are determined by the storage capacity of working memory. However, several researchers (Ericsson & Kintsch, 1995; Meyer & Kieras, 1997) have recently challenged the storage-based approaches to working memory. Instead, they argue that individual differences can be explained by use of strategies directed by the executive control system. Unfortunately, the role of the executive control system has not been explored in the context of scientific discovery.

This paper will outline a strategy-based approach to individual differences in working memory. In order to demonstrate the utility of this approach, the implications of a strategy-based model of individual differences in working memory will be applied to the problem of generating and evaluating hypotheses. As reviewed below, when multiple hypotheses are made explicit, they are an effective strategy for evaluating hypotheses.

Individual Differences in Working Memory

Traditionally, individual differences in working memory has been conceived in terms of storage limitations. Working memory consists of the temporary activation of long-term memory (Baddeley, 1986; Engle, Cantor & Carullo, 1993; Cowan, 1988). Unless refreshed, the activation of an item in working memory rapidly decays. Engle, Cantor, Carullo have suggested that individual differences in performance are a result of the amount of available activation. Individuals with high working memory can activate more information than individuals with low working memory. However, Jonides (1995) has argued that there is a need to distinguish between the storage and control functions of working memory during thinking.

A Strategy-Based Approach

A strategy-based approach stresses that the executive control system plays a central role in human performance. The executive control system has several major functions in thinking. First, according to Baddeley (1986), the executive control system allocates resources to the storage components of working memory: the visuo-

[1] This research was supported in part by a grant/fellowship from the Faculty Research Initiative Program from the University of Michigan-Flint.

spatial scratchpad and the articulatory rehearsal loop. Second, the executive control system is responsible for the generation, maintenance, and manipulation of mental representations towards the achievement of goals. Third, the executive control system selects and executes various strategies. Additionally, several cognitive scientists (Kimberg & Farah, 1993; Robin & Holyoak, 1995) have proposed that the executive control system binds various mental elements into a relational structure in service of a goal. Activating new associations permits the individual to switch goals and attend to new information. Finally, the executive control system serves to inhibit inappropriate activities or mental representations. For instance, when generating a hypothesis, experiment-generation strategies must be inhibited.

A strategy-based approach accepts the idea that working memory has a limited storage capacity. However, given the limited capacity of working memory, performance on complex real-world reasoning tasks depends on how working memory resources are allocated to the various activities needed to achieve the reasoner's goals. In this paper, strategy is defined as a plan, heuristic, algorithm, or procedure used to achieve goals. The executive control system is activated whenever a goal is satisfied or when a new goal is elicited. In complex real-world tasks, such as scientific discovery, subordinate goals are often embedded within superordinate goals. Jonides (1995) and Ward and Allport (1997) have suggested that the executive control system contributes to problem solving through subgoal management. For example, before a strategy can be selected to test a hypothesis, strategies for generating hypotheses must be executed.

Because the executive control system seeks to minimize the cognitive costs of strategy selection, the default strategy is to select the strategy which can be executed automatically. Consistent with Reder and Schunn (1996), strategy selection and execution typically occurs without conscious awareness. Selection of a strategy is determined by a strategy's current strength and level of activation. Each strategy is associated with a baseline and a current level of activation. The strategy with the highest current level of activation will be selected. The prior success of a particular strategy will determine the future likelihood of its selection. Strategy selection is determined by intrinsic factors such as the features of the information or the strength of a particular strategy as well as extrinsic factors such as task instructions.

<u>Implications for Scientific Discovery</u>

Wason (1960) originally developed the 2-4-6 task to investigate how people test hypotheses. In the 2-4-6 task, subjects are told that this sequence is one of many instances of a target hypothesis (e.g., increasing numbers). Subjects attempt to determine the target hypothesis by generating other sequences (e.g., 6-8-10). Subjects are told whether their sequences are instances of the target hypothesis. When testing hypotheses, people typically employ a strategy that involves generating instances of their current hypothesis (i.e., a positive-test). Subjects fail to appreciate the importance of falsification, that is, people do not propose tests intended to disprove their current hypotheses. As Klayman and Ha (1987) review, most explanations stress that subjects have a positive-test bias. However, recent models of scientific discovery (Freedman, 1995; Klahr & Dunbar, 1988; Mynatt, Doherty, & Dragen, 1993) have attempted to explain the cognitive processes that underlie the reliance on the positive-test strategy. In addition, several researchers (Farris & Revlin, 1989; Freedman, 1995; Klayman & Ha, 1987; Mynatt, Doherty, & Dragen, 1993) have recognized that humans typically do not consider alternative hypotheses. Specifically, Freedman (1995) and Mynatt et al. argue that working memory constrains the ability to consider multiple hypotheses. Yet, three possibilities exist concerning the relationship between working memory and scientific discovery. First, reasoners may lack working memory capacity. Indeed, Mynatt, Doherty, and

Dragen argue that reasoners can only store one hypothesis in working memory. This explanation implies a storage limitation. Second, the executive control system may lack the resources to evaluate multiple hypotheses. Because the executive control plays a central role in planning, reasoning and problem solving, the executive control system likely influences the generation and evaluation of hypotheses. Interference with the executive control system impairs syllogistic reasoning (Gilhooly, Logie, Wetherick & Wynn, 1993). A third possibility is that, as the storage capacity of working memory is reached and exceeded, resources from the executive control system are drained and, as the executive processes are needed to complete some task, less space for storage is available.

The primary goal driving hypothesis-testing is to determine the best hypothesis, conclusion, or explanation. The executive control system directs the testing of hypotheses during three phases: a generation phase, an evaluation phase and a revision phase. Furthermore, during these phases, the search through three problem spaces is constrained by the tester's working memory (see Figure 1). Extending the dual-space model of Klahr and Dunbar (1988), Freedman (1995) assumed scientific discovery reflects a search through three problem spaces: an experiment space, hypothesis space, and a strategy space. According to Klahr and Dunbar, the hypothesis space is the universe of possible hypotheses that can be applied to a particular problem and the experiment space is the possible tests of these hypotheses. In addition to the experiment and hypothesis space, reasoners also search through a strategy space. The strategy space contains the set of heuristics which direct the search through the experiment and hypothesis spaces. However, because of working memory limitations, only either the current goals, hypotheses, strategy, or experiment is active at any moment. The executive control system activates appropriate information and inhibits other information.

The Generation Process

The generation of a hypothesis or set of hypotheses is initiated by a situation where a decision is required, a problem needs to be solved, or evidence needs to be explained. Activation automatically spreads from the available information to the relevant hypotheses. Upon activating a hypothesis, a degree of confidence is assigned to that hypothesis. If the degree of confidence is sufficiently high, then the individual will terminate the hypothesis generation process. McDonald and Stenger (1993) have suggested that the plausibility of each hypothesis is determined by three factors: (1) the ease of access; (2) the consistency of the available information; and (3) the total number of hypotheses accessed. Because the most salient hypothesis is initially generated, alternative hypotheses are not generally generated. In fact, when a plausible hypothesis has been generated, the search for additional hypotheses would require controlled information processing. Given the preference of the executive control

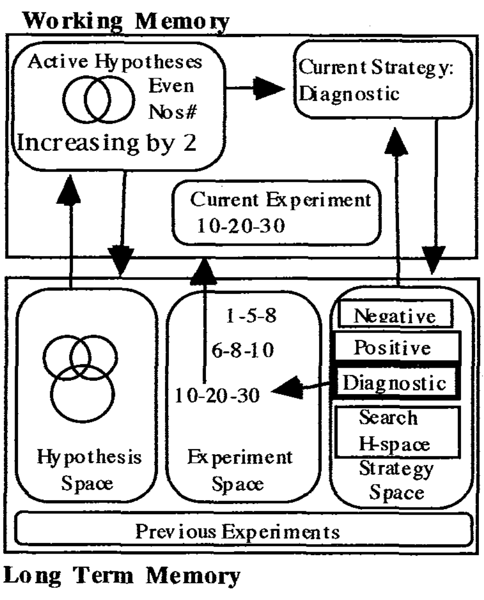

Figure 1. Hypothesis testing as a search through an experiment, hypothesis, and strategy space.

system for strategies that can be executed automatically, multiple hypotheses are typically not generated. When the perceived plausibility of the initial hypothesis is low, a search for additional hypotheses may be conducted. Mynatt et al. (1993) found that selection of information relevant to an alternative hypothesis increased when the probability of the focal hypothesis was less than .50. Because the executive control system relates various mental representations together in working memory (Robin & Holyoak, 1995), the ability to evaluate alternative hypotheses depends on the complexity of the hypotheses. When the hypotheses are simple, people should generate alternative hypotheses.

Unlike Mynatt, Doherty and Dragan (1993), who believed working memory can only store one hypothesis at a time, it is assumed that one hypothesis is the default strategy for testing hypotheses. This situation does not mean that people are not capable of evaluating more than one hypothesis at a time. However, because the generation and evaluation of multiple hypotheses requires controlled information processing, the executive control system will not test multiple hypotheses unless it has an explicit goal to do so. Additionally, because most studies on scientific discovery do not require subjects to write down their current hypotheses (cf. Klayman & Ha, 1987), the working memory resources required to maintain the current hypotheses in memory may prevent subjects from considering alternative hypotheses. Freedman (1991, 1992; Freedman & Endicott, 1997) found that the effective testing of alternatives is initiated only when this explicit goal is externally presented (i.e., explicit instructions to generate alternative hypotheses). This externally-defined goal forces the executive control system to generate alternative hypotheses. Freedman (1991, 1992; Freedman & Endicott, 1997) has made multiple hypotheses explicit by requiring subjects to write down their hypotheses. In this situation, subjects often can discover the target hypothesis with less experimentation than subjects who test a single hypothesis. As Freedman (1995) suggested, making alternative hypotheses explicit may lead subjects to be more efficient, because they reduce the working memory load. Freedman (1992, 1995; Freedman & Endicott, 1997) found that multiple hypotheses are effective because they lead to an extensive search of the hypothesis space. Similarly, Klahr, Dunbar, and Fay (1990) found that multiple hypotheses were more effective when they came from different parts of the hypothesis space. Thus, a explicit goal of generating explicitly-stated alternative hypotheses leads to an extensive and productive exploration of the hypothesis space.

The Evaluation Process

Once a hypothesis or set of hypotheses is generated, the next goal is selection of a hypothesis-evaluation strategy. The way in which subjects search through the experiment space can affect the likelihood of receiving confirmation or disconfirmation. For instance, a number sequence (e.g., 6-8-10), which is very similar to a previous sequence (e.g., 2-4-6), is likely to receive similar feedback. Therefore, it may be advantageous to select tests that are less similar to previous tests. Unfortunately, subjects typically select positive tests that are similar to previous tests. This strategy is the least informative since it typically produces ambiguous verification of subjects' hypotheses. A positive-test strategy may be selected for several reasons. First, the positive-test strategy may be relied upon because its effectiveness has made it the strongest strategy. Second, extensive experience with a positive test strategy may also lead to the automatic execution of this strategy. Thus, if one goal is to minimize the working memory load, then alternative strategies will not be solicited. Gilhooly (1996) has suggested that a positive-test strategy may be selected because it minimizes the working memory load. Oaksford et al. (1996) found that draining the executive control system's resources increases the likelihood of choosing confirmatory tests on the selection task. In other words, when the executive control system's resources are drained, reasoners fall back on their default strategies. Third, the positive-test strategy may be selected because the executive control system may activate a goal of obtaining evidence to support a current hypothesis. Negative are typically not selected for several reasons. First, negative tests require controlled information processing. Furthermore use of negative tests increases only when subjects are encouraged to try to disprove their hypotheses. Encouraging use of negative tests forces the executive control system to switch from its default strategy. Even when subjects generate negative tests, subjects are frequently unable to use this information (see Klayman & Ha, 1987, for review). Again, interpreting the evidence obtained from negative tests requires controlled information processing. Because negative tests often result in disconfirmation, the executive control system will be required to engage in the demanding task of hypothesis revision. Thus, the extra resources of executive control system needed to generate and interpret negative tests may work against its selection.

Multiple hypotheses shift the strategies selected to generate and evaluate hypotheses. Freedman (1995) argued that use of multiple hypotheses increase the effectiveness of the experiment-space search by making strategies leading to disconfirmation more salient. Specifically, alternative hypotheses suggest where in the experiment space to search for disconfirmation. When testing multiple hypotheses, use of negative tests are frequently generated and negative tests facilitates performance. Testing multiple hypotheses increases the number of negative tests produced (Freedman, 1992; Freedman & Endicott, 1997). Multiple hypotheses may increase the use of negative tests because multiple hypotheses may engage the executive control system and produce a shift in a person's goals. Specifically, the hypothesis tester may shift from providing evidence to support the current hypothesis to seeking evidence to determine which of the current hypotheses is incorrect. Both Klahr and Dunbar and Freedman (1992; Freedman & Endicott, 1997) found that testing multiple hypotheses increased the amount of disconfirmation received. When evaluating multiple hypotheses, the only way to ensure that one of your hypotheses is disconfirmed is to conduct a diagnostic test (i.e., a test that is an instance of one hypothesis and is not an instance of the other hypothesis). Freedman (1992 Freedman & Endicott, 1997) found that, when subjects tested multiple hypotheses, successful subjects evidenced a greater reliance on diagnostic tests. The presence of multiple hypotheses may make the sections of the experiment space, where diagnostic experiments can be generated, apparent. The presence of multiple hypotheses may allow subjects to integrate disconfirmatory information because the presence of an alternative to a disproven hypotheses reduces the working memory load. Finally, because multiple hypotheses may increase the use of strategies leading to disconfirmation, the presence of alternative hypotheses permit subjects to eliminate incorrect hypotheses more quickly. (Freedman, 1991, 1992; Klahr et al. 1990).

Hypothesis Revision

When a hypothesis has been eliminated, the executive control system activates a new goal to generate a new hypothesis. Human reasoners have a variety of strategies for revising hypotheses and for generating new hypotheses. The disconfirmatory information specify where in the hypothesis space subjects should search to replace a disconfirmed hypothesis. For example, an experiment (e.g., 1-3-5), which is an instance of the target hypothesis (e.g., increasing numbers), but not an instance of the subject's current hypothesis (e.g., even increasing numbers), suggests that the subject's hypothesis is too narrow. Thus, a crucial factor determining whether diagnostic or disconfirmatory tests will be effective is the way in which subjects modify their hypotheses. Furthermore, the revision of a hypothesis tends to occur for the hypothesis which is currently active in working memory. Freedman and Myers (1996) demonstrated that when evidence impacted a specific hypothesis subjects make an appropriate change in the probability of that hypothesis, but, subjects fail to make appropriate changes in the other hypotheses. Thus, people have difficulty evaluating multiple competing hypotheses. To make an appropriate revision of a current hypothesis, the executive control system must analyze the current disconfirmatory finding plus the previous findings in order to determine what hypothesis accounts of the extant data. Therefore, an inability to make adequate revisions in one's hypotheses may be due to an inability to integrate the current and previous evidence. Additionally, Farris and Revlin (1989) argued that people do not benefit from disconfirmation because subject can not conceive of an alternative to their current hypothesis. Thus, multiple hypotheses provide a means of overcoming people's difficulty with hypothesis revision.

Some Preliminary Evidence

Some preliminary evidence supports the individual differences in scientific discovery is due to strategic differences. Freedman and Endicott (1997) investigated how individual differences in working memory influenced performance on Wason's 2-4-6 task. Subjects tested either a single hypothesis, multiple hypotheses, or multiple hypotheses with a diagnostic strategy. Based on performance on a progressive matrices test, subjects were divided into high and low working-memory groups. Individuals with high working memory were more likely to discover the target hypotheses because they generated tests that facilitated the elimination of incorrect hypotheses. Specifically, compared with low working memory individuals, high working memory individuals generated more negative tests, received more negative, received more disconfirmation and generated more diagnostic tests. This pattern of findings suggests that the working-memory differences reflect the strategies employed to evaluate their hypotheses.

Conclusions

The present paper demonstrates that interpreting hypothesis-testing in terms of the strategic functions of the executive control system can provide a fruitful alternative explanation. According to this approach, strategies are selected to achieve the individual's goals. People test a single hypothesis by relying on positive tests because the executive control system selects these default strategies in order to minimize the necessary resources needed to engage in scientific discovery. While considering multiple hypotheses have some clear advantages including greater efficiency, increased use of negative tests, and more disconfirmation, multiple hypotheses require considerable resources of the executive control system. Additionally, people will not consider alternatives unless they are expressly instructed to do so and unless the alternative hypotheses are explicitly stated. In this situation, the executive control system shifts its goals and the explicitly stated hypotheses do not drain resources from the executive processes. Given the costs and benefits of multiple hypotheses, further research needs to determine the ways to minimize the costs while maximizing the benefits.

Consistent with Freedman (1995), the psychological processes involved in the generation and evaluation of hypotheses have clear implications for computational models of scientific discovery. Although computational models of scientific discovery (Cheng, 1990; Kulkarni & Simon, 1988) have included the development of strategies for the generation and evaluation of multiple hypotheses, these approaches have not focused on how the development and execution of these strategies are constrained by the executive control system. Therefore, as Freedman (1995) argued, further computational modeling of scientific discovery should include these constraints. Finally, a strategy-based approach can account for individual differences in memory. From this perspective, individuals with high working memory may not have greater storage capacity, as suggested Engle, Cantor, and Carullo (1992), but, instead, high working memory may be associated with superior encoding and retrieval strategies. In the future, a strategy-based approach should be applied to each of these areas.

References

Baddeley, A. D (1986). Working Memory. Oxford: Oxford University Press.

Cheng, P. C-H. (1990). Modeling scientific discovery. Technical Report No. 65, Human Cognition Research Laboratory, The Open University, Milton Keynes, England.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychological Bulletin, 104, 163-191.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, & Cognition, 18, 972-992.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102, 211-245.

Farris, H., & Revlin, R. (1989). Sensible reasoning in two tasks: Rule discovery and hypothesis evaluation. Memory and Cognition, 17, 221-232.

Freedman, E. G. (1991). Role of multiple hypotheses during collective induction. Paper presented at the 61st Annual Meeting of the Eastern Psychological.

Freedman, E. G. (1992). Scientific induction: Multiple hypotheses and individual and group processes. Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, 183-188.

Freedman, E. G. (1995). Making multiple hypotheses explicit: An explicit strategy for computational models of scientific discovery. Working Notes of The AAAI Spring Symposium Series: Systematic Methods of Scientific Discovery, 119-124. Menlo Park: AAAI Press.

Freedman, E. G., & Endicott, S. A. (1997). Individual differences in working memory and the testing of single versus multiple hypotheses. Paper presented at the 38th Annual Meeting of the Psychonomic Society.

Freedman, E. G., & Myers, A. M. (1996). Effects of evidence type, number of hypotheses and alternative strength on hypothesis evaluation. Paper presented at the 37th Annual Meeting of the Psychonomic Society.

Gilhooly, K. J. (1996). Thinking: Directed, undirected and creative (3rd Edition). New York: Academic Press.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. Memory and Cognition, 21, 115-124.

Jonides, J. (1995). Working memory and thinking. In Smith, E. E., & Osherson, D. N. (Eds.), An invitation to cognitive science: Volume 3. Thinking (2nd ed., pp. 215-265). Cambridge, MA: MIT Press.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. Journal of Experimental Psychology: General, 122, 411-428.

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

Klahr, D., Dunbar, K., & Fay, A. L. (1990). Designing good experiments to test bad hypotheses. In J. Shrager & P. Langley (Eds.), Computational models of scientific discovery and theory formation (pp. 356-402). Palo Alto: Morgan Kaufmann.

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

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

McDonald, J., & Stenger, J. (1993). Effects of hypothesis salience and the use of positive and diagnostic test strategies. Organizational Behavior and Human Decision Processes, 56, 213-232.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic Mechanisms. Psychological Review, 104, 3-65.

Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information Relevance, working memory and the consideration of alternatives. Quarterly Journal of Experimental Psychology, 46A, 759-778.

Oaksford, M., Morris, F., Grainger, B., & Williams, J. M. G. (1996). Mood, reasoning, and central executive processes. Journal of Experimental Psychology: Learning, Memory and Cognition, 22, 476-492.

Reder, L. M., & Schunn, C. D. (1996). Metacognitive does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), Implicit memory and metacognition (pp. 45-77). Manwah, NJ: Lawrence Erlbaum.

Robin, N., & Holyoak, K. J. (1995). Relational complexity and the functions of the prefrontal cortex. In M. Gazzaniga (Ed.), The cognitive neurosciences (pp. 987-997). Cambridge, MA: MIT Press.

Ward, G., & Allport, A. (1997). Planning and problem-solving using five-disc tower of London task. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A, 49-78.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 12, 129-140.