Title
Goal Specificity in Hypothesis Testing and Problem Solving

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 16(0)

Authors
Vollmeyer, Regina
Holyoak, Keith J.
Burns, Bruce D.

Publication Date
1994

Peer reviewed

Goal Specificity in Hypothesis Testing and Problem Solving

Regina Vollmeyer
Department of Psychology
University of California, LA
Los Angeles, CA 90024-1563
regina@psych.ucla.edu

Keith J. Holyoak
Department of Psychology
University of California, LA
Los Angeles, CA 90024-1563
holyoak@psych.ucla.edu

Bruce D. Burns
Department of Psychology
University of California, LA
Los Angeles, CA 90024-1563
burns@psych.ucla.edu

Abstract

Theories of skill acquisition have made radically different predictions about the role of means-ends analysis in acquiring general rules that promote effective transfer to new problems. Under one view, means-ends analysis is assumed to provide the basis for efficient knowledge compilation (Anderson, 1987), whereas under the alternative view means-ends analysis is believed to disrupt rule induction (Sweller, 1988). We suggest that in the absence of a specific goal people are more likely to use a rule-induction learning strategy, whereas providing a specific goal fosters use of means-ends analysis, which is a non-rule-induction strategy. We performed an experiment to investigate the impact of goal specificity and systematicity of rule-induction strategies in learning and transfer within a complex dynamic system. Subjects who were provided with a specific goal were able to solve the initial problem, but were impaired on a transfer test using a similar problem with a different goal, relative to subjects who were encouraged to use a systematic rule-induction strategy to freely explore the problem space. Our results support Sweller's proposal that means-ends analysis leads to specific knowledge of an isolated solution path, but does not provide an effective method for learning the overall structure of a problem space.

Introduction

A central problem in cognitive science is to identify the relationship between problem solving and learning. People can learn from solving problems, but it is unclear exactly how learning takes place or what is learned. People sometimes seem to learn little from a problem-solving episode except a specific solution to a particular problem; yet on other occasions people acquire more general knowledge that can be applied to a wide range of related problems. What is the difference?

Goal Specificity and Rule Induction

A particularly intriguing possibility is that some solution methods may be effective for finding solutions to specific problems, but relatively ineffective in promoting abstraction of more general knowledge that would support transfer to novel but related problems. A case of particular theoretical interest concerns the role of means-ends analysis in learning. Some theories of learning have claimed that means-ends analysis, while itself a weak problem-solving method used primarily by novices, is nonetheless a valuable stepping stone toward expertise. According to this view, solutions first generated by means-ends analysis are subsequently compiled into rules that allow more efficient solutions to be found for problems similar to the original one (e.g., Anderson, 1987; Larkin, 1981).

Other theorists, however, have argued that means-ends analysis and similar problem-solving methods can actually impede the acquisition of general rules (e.g., Mawer & Sweller, 1982; see Holyoak, 1991, for a brief review). Means-ends analysis can be applied to well-defined problems with a specific goal, and its immediate product is not a general rule, but simply a solution path that achieves the immediate goal. We will term a strategy that achieves a specific goal without necessarily yielding general rules a *non-rule-induction* (NRI) strategy. In contrast, other learning strategies can operate on ill-defined problem situations that lack a specific goal. In the absence of a specific goal, free exploration of a problem space may yield general rules about state transitions, which can later be used to achieve a relatively wide variety of goals, thus promoting transfer to a family of similar problems. We will term a strategy that focuses directly on rule acquisition, rather than on achievement of a specific goal, a *rule-induction* (RI) strategy.

Sweller and coworkers found evidence that people with a non-specific goal gained more knowledge about a task than did people with a specific goal (e.g., Sweller, 1988; Sweller, Mawer, & Ward, 1982). They interpreted their results as evidence that different learning strategies were applied depending on goal specificity. For example, subjects in one set of experiments involving solving geometry problems were provided with partial information about the angles and sides of a triangle, and were asked either to calculate a particular angle (specific goal), or else to calculate all possible angles and sides (non-specific goal). However, both groups had to calculate the same sides and angles. Subjects who received the non-specific goal instructions were subsequently more successful in solving transfer problems.

Such evidence suggests that non-specific goals encourage use of RI strategies, which promote acquisition of more general knowledge about the structure of a problem space. However, RI strategies can vary in their effectiveness. People differ in the degree of systematicity with which they formulate and test their hypotheses, and those who formulate hypotheses in a task-appropriate and testable way generally gain more knowledge (e.g., Dunbar, 1993; Klahr & Dunbar, 1988; Klahr, Fay, & Dunbar, 1993). In general, the optimal strategy for testing hypotheses about the influence of multiple factors on one or more dependent variables is the VOTAT (*v*ary *o*ne *t*hing *a*t a *t*ime) strategy (Tschirgi, 1982), in which one factor is varied while the others are held constant. This RI strategy is central to experimental design in science.

The present study was designed to investigate the role of goal specificity and learning strategies in the acquisition and transfer of knowledge about a complex dynamic system. The learning domain was chosen because it is especially suitable for investigating the interrelationships between problem solving and hypothesis testing. By manipulating both goal specificity and the systematicity of subjects' hypothesis-testing strategies, we attempted to determine whether an NRI strategy or an RI strategy is most effective in promoting learning and transfer. Our hypothesis was that while an NRI strategy would be adequate for achieving a specific goal, a systematic RI strategy, VOTAT, would be more effective for acquiring general structural knowledge about the domain, resulting in greater transfer to similar problems with different goal states.

Biology Lab: A Dynamic Problem Environment

Since the early 1980s, researchers have used computer-simulated scenarios to study complex problem solving (for a review see Funke, 1991). These tasks are relatively complex, as multiple variables have to be manipulated in order to achieve multiple goals simultaneously. In the present study we used a computer-driven dynamic problem environment we termed *biology lab*, constructed using the shell DYNAMIS (Funke, 1991). In our cover story, subjects were told that they were in a biology lab in which there is a tank with four species of sea animal (crabs, prawns, lobsters, sea bass). These species are affected by four input variables (temperature, salt, oxygen, current). The structure of the environment, illustrated in Figure 1 (which was never shown to our subjects), was such that two of the outputs (prawns, crabs) are relatively simple to manipulate because each is influenced by only one input (as shown by relations I and II, respectively). The other two outputs are more complex, because each is influenced by two factors. One output (sea bass) is affected by two inputs, and the other (lobster) is affected by a decay factor (marked as a circle connected to the output) in addition to a single input variable. The decay factor was implemented by multiplying an output by a constant negative factor on each trial. Decay is a dynamic aspect of the system, because it yields state changes even if there is no input (i.e., all inputs are set to zero). The system is thus complex in that it involves multiple input variables that must be manipulated to control multiple output variables, and dynamic in that the state of the system changes as a joint function of external inputs and internal decay.

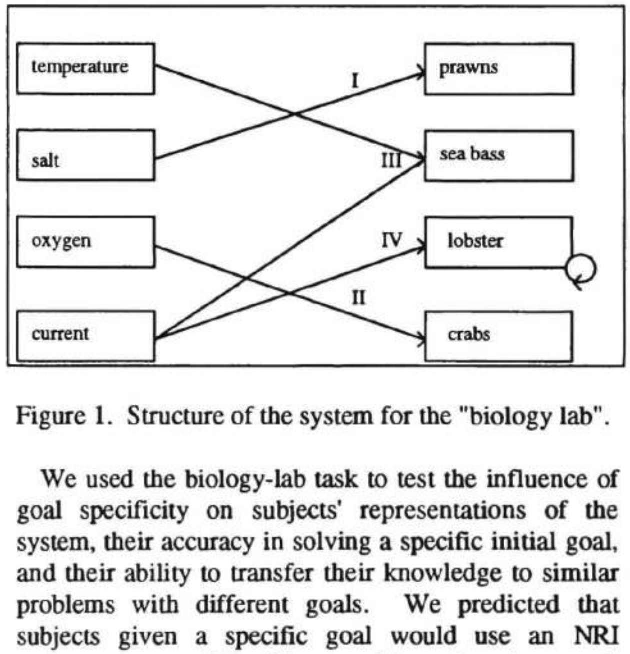

Figure 1. Structure of the system for the "biology lab".

We used the biology-lab task to test the influence of goal specificity on subjects' representations of the system, their accuracy in solving a specific initial goal, and their ability to transfer their knowledge to similar problems with different goals. We predicted that subjects given a specific goal would use an NRI strategy that would suffice to achieve the given goal, but would yield a poorer representation than would be achieved by subjects given a non-specific goal coupled with instruction in an optimal RI strategy, VOTAT.

The more complete representation of the system attained by the latter subjects was expected to lead to more effective transfer to other biology-lab problems with altered goals.

A Study of Learning in a Dynamic Environment

Method

Subjects: Sixty undergraduate students at the University of California, Los Angeles, participated for course credit.

Design: The experiment included four conditions, defined by the factorial combination of two levels of goal specificity (specific vs. non-specific) and two levels of strategy instruction (instruction to use VOTAT vs. no such instruction). Fifteen subjects served in each condition.

Procedure: The biology-lab problem required subjects on each trial to set the levels of the four input variables and observe the resulting values of the output variables (numbers of each of four species of sea animals). The underlying structure of the system was as depicted in Figure 1. Each series of six trials was defined as a "round", at the beginning of which the system was set to a specific state. All subjects received three initial exploratory rounds followed by a fourth round in which they were asked to produce a specific goal state (namely, 50 crabs, 400 prawns, 900 lobsters, and 700 sea bass). Subjects in the specific-goal condition were informed of this goal from the outset of round 1, and thus had a total of four rounds to achieve the goal. In contrast, subjects in the non-specific-goal condition were not given any specific goal until round 4. In rounds 1-3, these subjects were simply asked to set inputs and observe outputs in order to figure out how the system works.

Before starting to manipulate the system on the computer all subjects received general instructions about the task. In addition, subjects in the strategy-instructed group were given written instructions explaining the optimal strategy (VOTAT) of varying just one variable at a time, setting the remaining variables to zero. Subjects in the strategy-uninstructed group received no advice on how to explore the system. After each round of the exploratory phase (rounds 1-3), subjects completed a "structure diagram", in which subjects indicated how they believed the input variables affect the output variables. They were provided with a diagram showing the inputs and outputs as in Figure 1, but with all links omitted. The subjects' task was to draw links between variables that they believed to be dependent, and also to assign weights indicating how strong they felt each influence was.

In round 4, all subjects were presented with a specific goal state. This goal state was the same as that which those in the specific-goal group had had throughout the exploratory phase (rounds 1-3). Finally, in round 5 all subjects were asked to achieve a different goal state (namely, 250 crabs, 200 prawns, 1000 lobsters, and 350 sea bass) that was new to subjects in all conditions. Performance on this new goal provided a measure of the degree to which learning over rounds 1-4 yielded transfer to a novel problem drawn from the same problem space. The entire experiment took an hour to complete.

Results

Dependent variables: Three dependent variables were analyzed to provide evidence of learning and transfer. (1) *Structure score*. The structure diagram completed by all subjects after each of the first three rounds was used to derive a score reflecting degree of knowledge of the underlying structure of the system. This structure score was computed as the sum of the number of correct specifications of links, directions, and weights, adjusted with a correction for guessing (see Woodworth & Schlosberg, 1954, p. 700). Because the structure score after round 3 was most informative about subjects' knowledge at the end of the initial learning phase, this score was used in all analyses reported here. (2) *Solution error*. Solution error in reaching the specific goal state during round 4 was computed as the sum of the absolute differences between the target and the obtained number for each of the four output variables. As this measure produced a skewed distribution, the variance was corrected by applying a logarithmic transformation. Solution error was computed for each of the six trials that comprised round 4, in order to determine how quickly subjects were able to approach the target goal. As there was no difference in performance between trials the mean error for the six trials was used. (3) *Transfer error*. Transfer error in achieving the new goal introduced in round 5 was measured in exactly the same way as solution error in round 4.

Preliminary analyses. Preliminary analyses were performed to determine whether our measures of learning were systematically related, and whether the manipulation of subjects' learning strategy by instructions had been effective. If the structure score derived from subjects' completions of structure diagrams in round 3 provides a valid assessment of what they had learned about the system, then the structure score would be expected to correlate inversely with solution error measured on round 4 and transfer error on round 5. This was indeed the case. Subjects with higher structure scores produced lower solution error when they had to reach the goal state in round 4, $r = -.50$, $p < .001$, as well as lower transfer error in round 5, $r = -.58$, $p < .001$.

We also tested whether our manipulation of learning strategy by instructions was successful. We examined subjects' patterns of settings for the four inputs to determine their basic strategy. Subjects were classified as trying to reach the goal state (NRI strategy) when two criteria were met: 1) at least one of the four output states for the specific goal was reached; and 2) they displayed a pattern of gradually coming closer to the goal (as opposed to directly calculating the correct output value). In contrast, subjects were classified as using the RI strategy of VOTAT if on at least four out of the six trials of a round they set the pattern of varying a single input while setting the remaining three inputs to zero. Other strategies (e.g., varying multiple inputs at once) formed a heterogeneous set. These additional patterns were difficult to classify firmly as NRI or RI strategies. Eighty percent of all subjects in the strategy-instructed conditions followed the VOTAT strategy in the first round. Figure 2 shows how the goal conditions influenced the percentage of strategy-instructed subjects using each non-VOTAT strategy on each round. (The missing percentage reflects subjects using VOTAT.) Most strategy-instructed subjects in the non-specific-goal condition continued with the VOTAT strategy through round 3. However, the strategy-instructed subjects who had a specific goal exhibited a strong tendency to switch from the VOTAT strategy to an NRI strategy that focuses directly on reaching the stated goal. Figure 3 shows the same information as Figure 2, but for strategy-uninstructed subjects. As can be seen from Figure 3, in the absence of strategy instructions most subjects did not spontaneously use the VOTAT strategy. Giving subjects a goal to reach, did have an effect on strategies, however, as many subjects with a specific goal used an NRI strategy whereas no subject in the non-specific-goal group ever used that strategy. It thus appears that our strategy instruction was indeed effective in promoting use of the RI strategy of VOTAT, but that providing a specific goal created a strong pressure to employ an NRI strategy.

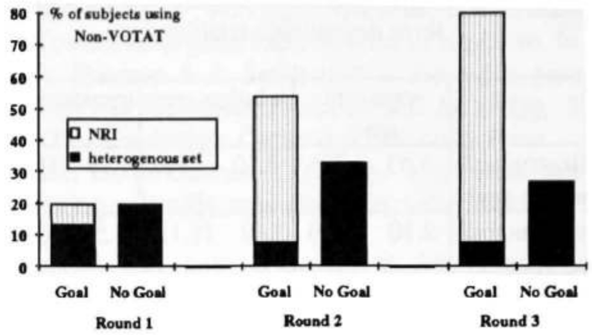

Figure 2. Percentage of strategy-instructed subjects using non-VOTAT strategies over three rounds.

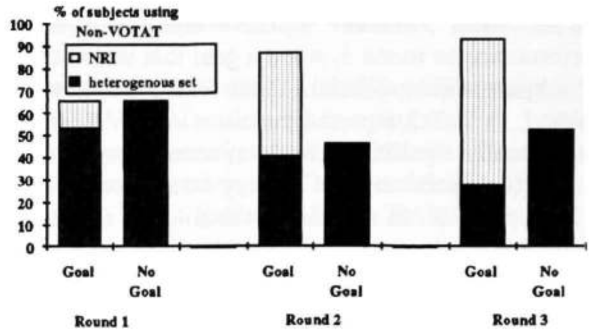

Figure 3. Percentage of strategy-uninstructed subjects using non-VOTAT strategies over three rounds.

Influence of goal specificity and strategy on learning and transfer. Subjects achieved higher structure scores when given a non-specific goal (M = 2.35) rather than a specific goal (M = 1.69), $F(1,56) = 15.5$, $p < .001$; and when given instruction in the VOTAT strategy (M = 2.22) rather than no strategy instruction (M = 1.82), $F(1,56) = 5.87$, $p < .05$. The interaction between goal specificity and strategy instruction as determinants of structure score was not significant. In addition, as depicted in Table 1, subjects instructed in use of the VOTAT strategy achieved lower solution error in round 4 over all six trials, $F(1,56) = 4.9$, $p < .05$. Solution error did not differ as a function of goal specificity, $F(1,56) = .35$. Although subjects in the non-specific-goal condition achieved greater overall knowledge of the system structure, those in the specific-goal condition had three additional rounds of practice in attaining the goal set for all subjects in round 4. These offsetting advantages may explain the groups' equal performance.

Table 1. Means and variances (in parentheses) for the three dependent variables.

	structure score	solution error	transfer error
uninstructed/ specific goal	1.53 (.73)	4.10 (.96)	3.99 (1.08)
uninstructed/ non-sp. goal	2.10 (.69)	3.69 (1.18)	3.58 (.87)
instructed/ specific goal	1.84 (.69)	2.77 (2.89)	3.22 (2.98)
instructed/ non-sp. goal	2.60 (.45)	2.52 (2.85)	2.22 (2.90)

The most crucial results concern transfer performance on round 5, when a goal that was novel to all subjects was introduced. These results are shown in Table 1. A 2x2x2 repeated-measures ANOVA yielded a marginally significant three-way interaction between round, goal specificity and strategy instruction, $F(1,56) = 2.80$, $p < .10$. In contrast to the solution round, the group that received the non-specific goal in rounds 1-3, coupled with instruction in the VOTAT strategy, achieved lower error scores on the transfer problem than did the other three groups combined, $F(1,58) = 4.54$, $p < .05$. Thus although subjects given a specific goal, who predominantly used an NRI learning strategy, were able to effectively achieve that specific goal, they were relatively poor in transferring their knowledge to a similar problem with a new goal. The combination of low goal specificity and instruction in the VOTAT strategy maximized knowledge of the overall structure of the dynamic system, and thereby maximized transfer to the new problem.

Discussion

The aim of the present study was to test alternative theories of the relationship between problem solving and acquisition of generalized rules. The biology-lab domain, a complex dynamic system involving multiple input variables that must be manipulated to control multiple output variables, provided a rich environment in which to explore the influence of goal specificity and hypothesis-testing strategies on learning and transfer. We found that providing subjects with a specific goal from the outset of learning produced a strong tendency to use a non-rule-induction strategy. The predominant strategy for such subjects was a variant of means-ends analysis that focused on incrementally reducing the difference between obtained outputs and the specific goal. This strategy was adequate for eventually solving the particular goal, but was suboptimal as a vehicle for discovering the overall structure of the system. As a result, provision of a specific goal impaired eventual transfer to a new problem drawn from the same problem space but involving a different goal state.

Acquisition of the structure of the system was fostered both by using a non-specific goal and by providing explicit instruction in an optimal rule-induction strategy, VOTAT, which involves varying a single factor at a time while holding other factors constant at zero. However, subjects who were given a specific goal tended to abandon the VOTAT strategy over the course of the learning session, shifting to a non-rule-induction strategy. Subjects who were not taught the VOTAT strategy tended to use either a non-rule-induction strategy (if a specific goal was provided) or some other suboptimal strategy (if no specific goal was provided). Thus optimal transfer performance required a combination of a non-specific goal coupled with instruction in use of an effective rule-induction strategy.

Our results run counter to theories of skill acquisition that stress the importance of learning from weak problem-solving methods as a means of inducing general rules (e.g., Anderson, 1987; Larkin, 1981). It is certainly possible that people sometimes learn general rules in the aftermath of solving problems by variants of means-ends analysis; however, at least in the absence of prior knowledge of the domain, this approach does not appear to provide an optimal path toward either general knowledge of the structure of a complex system or successful transfer to problems with an altered goal. Rather, acquisition of system structure is fostered to a greater extent by free exploration of the problem space.

It should be noted that the NRI strategy used by our subjects, although goal-directed, did not meet the technical definition of means-ends analysis (i.e., removing the largest difference between the current state and goal state, in the process recursively solving the subproblem of getting from the current state to that which satisfies the preconditions of required operators). Thus our results do not directly show that the full means-end strategy would fail to promote learning of overall problem structure. Nonetheless, the present NRI strategy did involve difference reduction (i.e., search in which each step progresses closer to the specified goal), which is a major component of means-ends analysis. It is possible that the key factor limiting acquisition of overall structure is focus on a specific goal, in which case full means-ends analysis, like the NRI strategy used by our subjects, would also prove relatively ineffective in promoting learning. However, further research will be required to test this possibility.

Another caveat concerning the present findings relates to the fact that our study used a problem domain in which our subjects were complete novices. A different pattern of results might emerge in a problem domain for which subjects have a prior theory of the domain. In a more knowledge-rich domain, mechanisms of explanation-based learning (e.g., Mitchell, Keller, & Kedar-Cabelli, 1986) might allow people to form generalizations of solutions initially obtained by weak methods, such as means-ends analysis. One direction for future work would involve manipulating domain knowledge together with subjects' learning strategies, and examining transfer performance in the aftermath of initial problem solving.

The present results are broadly in agreement with the findings of Sweller and his colleagues (Sweller, 1988; Mawer & Sweller, 1982), who also found that reduced goal specificity yields greater transfer. The present study increases the generality of this conclusion by demonstrating similar results in the domain of a complex dynamic system, as opposed to the static mathematical domains primarily used in earlier studies. In addition, the present study goes beyond previous work in identifying the interactive relationship between hypothesis-testing strategy and the impact of reduced goal specificity. In a complex task environment such as the biology lab, college students are not generally prepared to make spontaneous use of an effective rule-induction strategy, even when they are given a non-specific goal. It is therefore important to provide instruction in the use of such a strategy in order to allow maximum benefit from free exploration of the problem space. It is not enough to simply "wander" through a haphazard series of input-output relations; rather, effective learning depends on systematic investigation of controlled variations in the inputs. Our results thus have important educational implications for designing effective techniques for encouraging problem-based learning in complex domains.

References

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397-434.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R.J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Erlbaum.

Holyoak, K. J. (1991). Symbolic connectionism: Toward third-generation theories of expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 301-335). Cambridge: Cambridge University Press.

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

Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*, 111-146.

Larkin, J. H. (1981). Enriching formal knowledge: A model for learning to solve textbook physics problems. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp 311-334). Hillsdale, NJ: Erlbaum.

Mawer, R. F., & Sweller, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 252-259.

Mitchell, T. M., Keller, R. M., & Kedar-Cabelli, S. T. (1986). Explanation-based generalization: A unifying view. *Machine Learning, 1*, 48-80.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 112*, 639-661.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*, 1-10.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt, Rinehart, and Winston.

Acknowledgements

This research was financially supported by NSF Grant SBR-9310614 and Vo 514/2-1 of the Deutsche Forschungsgemeinschaft