UC Merced

Title
Factors Mediating the Success of Observation-based Problem Solving

Permalink
https://escholarship.org/uc/item/7r16z35x

Author
Osman, Magda

Peer reviewed

Factors Mediating the Success of Observation-based Problem Solving

Magda Osman (m.osman@ucl.ac.uk)
Department of Psychology, University College London. Gower Street
London, WC1E 6BT England

Abstract

Studies of complex dynamic control tasks (CDCTs) have revealed that when problem solvers learn about a system indirectly through observation (Berry, 1991; Lee, 1995), their procedural knowledge of the system is impaired relative to their declarative knowledge. However, when learning is through direct interactions with the system, then problem solvers declarative knowledge is impaired relative to their procedural knowledge. Osman (in press) claims that one reason that observation based learning produces such poor procedural knowledge is that observers are prevented from hypothesis testing during learning and monitoring the status of their knowledge of the system; this has been shown to be critical in the acquisition and application of relevant knowledge in CDCTs (e.g., Sanderson, 1989). The present study explored the effects of preventing and encouraging hypothesis testing during observational learning on the accuracy of declarative and procedural knowledge of a dynamic problem solving task. The findings show that when instructions promote self evaluative processes during hypothesis testing, problem solving ability is improved compared to when self evaluative processes are prevented.

Conflicting theoretical issues

CDCTs have been a popular task environment for examining many phenomena, including motivational and affective processes in complex decision making (e.g., Locke & Latham, 2002), skill learning in naturalistic decision making (e.g., Brehmer, 1992) memory and attentional processes in problem solving (Burns & Vollmeyer, 2002), and implicit learning (e.g., Berry, 1991; Lee, 1995). Their popularity and range make them ideal for studying the acquisition and transfer of skill-based knowledge in a variety of complex interactive environments.

Declarative vs. procedural skill learning: CDCTs have been described as procedural tasks because they involve perceptual-motor behaviors that fulfill a set of constraints in order to achieve a goal. Procedural learning involves controlling an environment that is dynamic: i.e., it is changing as a consequence of the learner's actions. The knowledge that is acquired is procedural, and represents "knowing how" to perform actions that are tied to specific goals. This is different from declarative knowledge, which is "knowing that" of particular facts about the underlying actions and structural knowledge concerned with the goal itself (e.g. Anderson, 1982). The popular position on procedural learning in problem solving is that procedural knowledge and declarative knowledge are dissociated (Berry, 1991; Berry & Broadbent, 1984; Dienes & Berry, 1997; Lee, 1995), and that they are supported by functionally separate cognitive mechanisms (e.g., Squire, 1986). It follows that having declarative knowledge alone will impair one's later ability to perform a procedural task. One method used to demonstrate this involves training people on a procedural task by observing another perform it first: Because the learners are explicitly monitoring what they are observing, this is claimed to generate declarative knowledge (e.g., Kelly & Burton, 2001). Berry (1991) and Lee (1995) used this method to compare the effects of procedural-based and observation-based learning. They showed that, when participants later came to problem solve, the observers' ability to perform the procedural task was poorer than that of procedural-based learners. Returning to the example, the suggestion here is that John's ability to master the device would have been more successfully achieved had he tried to learn-by-doing rather than learn-by-observing. Although popular, the claim that procedural-based learning has an advantage over observation-based learning in problem solving has attracted little empirical support (e.g., Berry, 1991; Lee, 1995).

Goal specificity: Some (e.g., Burns & Vollmeyer, 2002; Sweller, 1988; Vollmeyer et al., 1996) claim that during learning, the nature of CDCT encourages people to generate hypotheses (i.e., they select which inputs they want to manipulate) and test them (i.e., they make predictions as to which outputs the inputs changed will effect, and the outcome of the manipulation is monitored and used to update the model of the system). Evidence for this comes from studies that compare different types of goal instructions during learning. For instance, instructions like "explore the system", an example of non-specific goal, are contrasted with "learn about the system while trying to reach and maintain specific outcomes", an example of a specific goal. In the control test phase, specific goal learners perform more poorly than non-specific goal learners.

Burns and Vollmeyer's (2002) recent account of the differential effects of goal specificity on problem solving develops on Dual-Space theory (Klahr & Dunbar, 1988). The theory proposes that a skill is acquired by using the same principles that underlie scientific discovery: i.e., designing the appropriate procedure (experimental design) to evaluate a theory (hypothesis formation). Dual-Space theory deconstructs a task into spaces: the hypothesis space, which consists of the hypotheses generated, and the experimental space, which consists of instances of the problem that can potentially be tested. Search in the hypothesis space is guided by self-evaluative processes and

prior knowledge, whereas search in the experimental space is guided by the current goal. Specific goal learning promotes exclusive search of the experimental space with limited self-evaluative processing, whereas non-specific goal learning involves moving through both spaces via hypothesis testing, and therefore necessarily involves self-evaluative processes.

Osman (in press) examined the goal specificity effect (i.e., Specific Goal vs. Non-Specific Goal) under conditions in which problem solvers' mode of learning was either procedural-based or observation-based. The findings showed that differences in control performance were not the result of the format in which participants learnt (i.e., procedural, observation), but of the goal that was pursed during learning. This evidence further supports Burns and Vollmeyer's Dual-Space theory, by showing that preventing hypothesis testing and self-evaluative thinking through specific goal learning leads to decrements in problem solving ability, regardless of whether learning is observation-based or procedural-based. Osman argued that one reason that Berry (1991) and Lee (1995) reported poor control performance in observation-based learning of a control task (Berry, 1991; Lee, 1995) was that participants were given specific goal instructions that prevented hypothesis testing and self-evaluative thinking. In Berry's (1991) study, observation-based problem solvers were actively discouraged from hypothesis testing, and in Lee's study they were required to learn a rule passively without employing any evaluative thinking. This also indicates that observation-based learners are sensitive to goal specificity in the same way as procedural-based learners.

In sum, to further our understanding of the effectiveness of learning vicariously, this study examines how goals mediate learning processes that later affect problem solving ability. To achieve this, the present study contrasts the effects of specific goals and non-specific goals under observation-based conditions, in order to explore further the relationship between declarative and procedural knowledge in a complex dynamic problem solving task. To extend the work on goal specificity, the study examines the extent to which attenuating hypothesis testing and self-evaluative processes hinders problem solving ability when these behaviors are prevented under non-specific goal learning conditions, and encouraged under specific goal learning conditions.

An Example of a Complex Dynamic Control Task: Typically complex dynamic control tasks involve several input variables which are continuous (e.g., concentration levels of salt, carbon, and lime) and that are connected via a complex causal structure or rule to several output variables that are also continuous (e.g., Chlorine concentration, Oxygenation levels, temperature) (See Figure 1). The example used here is taken from Burns and Vollmeyer's (2002) task which was originally based on a water tank purification system. In their system the starting values of the inputs were set to 0, and those of the outputs are: Oxygen = 100, Chlorine concentration = 500, Temperature = 1000,

and the objective was to learn the causal structure and numerical relationship between the inputs and outputs; which is described as linear, but with constant value added to each connection.

Thus, learning about the system and then attempting to control it requires that participants accurately incorporate the continuous feedback they receive on the output variables while changing the input variables. For instance, if on the first trial a participant decides to change the level of the input variable Carbon to 100 units, the value for the output Temperature will be 1104 (i.e., 1000 (which is the starting value) + 100 (value of Carbon) + 4 (the added value)). Because the input Carbon belongs to a common effect causal structure, the output Chlorine Concentration is also affected, and its value will be 599.5 (i.e., 500 (which is the starting value) + 100 (value of Carbon) + -0.5 (the added value)).
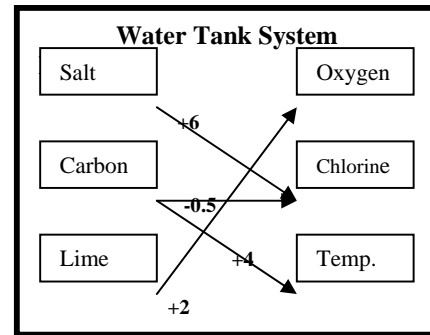

Figure 1: Water tank system

If on trial 2 the value of the input Salt is changed, the output values of Temperature and Oxygenation will remain the same as the previous trial and only Chlorine Concentration will change because it is the only output connected to the input Salt.

Present Experiment

The present experiment had two main objectives. The first was to investigate the goal specificity effect, using observation-based versions of Burns and Vollmeyer's Water Tank control task. To examine this, the present study included two conditions in which the instructions corresponded to the specific goal and non-specific goal (SG, NSG, respectively) instructions presented to participants in Burns and Vollmeyer's original study, but the learning phase was observation-based. The second objective was to examine the effects on structural knowledge and control performance when hypothesis testing behavior and self-evaluative thinking are prevented during learning. If, as has been suggested, these behaviors are critical in the acquisition and application of knowledge in control tasks, then attenuating them will induce decrements in problem solving ability. To examine this, two further conditions were

included to complement the non-specific goal and specific goal conditions. The NSG-Hypo(-) included similar instructions to the non-specific goal condition, with the exception that solvers were explicitly told to avoid hypothesis testing, consistent with Berry's (1991) original instructions and evaluative thinking. The SG-Hypo(+) condition included the same instructions as the SG condition, with the exception that solvers were told to hypothesis test explicitly and to examine their behavior. First, if problem solving ability in CDC-tasks requires monitoring and evaluation of hypothesis testing behavior, then attenuating these behaviors should lead to decrements in performance in the SG and NSG-Hypo(-) conditions compared to the NSG and SG-Hypo(+) conditions. However, if observation-based learning impairs knowledge acquisition, irrespective of the presence or absence of hypothesis testing (Berry, 1995; Lee, 1991), then there should be no difference in problem solving ability between the four conditions.

Method

Participants Sixty-four students from University College London volunteered to take part in the experiment, and were paid £4 for their participation. Participants were randomly allocated to one of four conditions [NSG, SG, NSG-Hypo(-), SG-Hypo(+)], with sixteen in each of the four conditions. Participants were tested individually and were presented with a fully automated version of Burns and Vollmeyer's (2002) control water tank system task, which was run on Dell Optiplex computers. The experimental program was written in Visual Basic 6.

Design Experiment 1 included two types of goal specificity instructions (NSG, SG), but for two conditions the instructions were designed to reverse the effects of goal specificity, by encouraging [SG-Hypo(+)] and discouraging [NSG-Hypo(-)] hypothesis testing and self-evaluation. There was thus a total of four conditions [NSG, SG, NSG-Hypo(-), SG-Hypo(+)]. Participants were presented with a learning phase that was divided into blocks, each consisting of 6 trials, and a test phase that included two control tests (Control Test 1, Control Test 2), each consisting of 6 trials. The CDCS presented in Figure 1 is taken from Burns and Vollmeyer's (2002) task, which was based on a water tank purification plant. By manipulating the input values, problem solvers can track the effects on the outputs, which enables them to reason from cause (input changes) to effect (output changes), via acquisition of the causal structure or the rule that relates inputs and outputs. In Burns and Vollmeyer's example, the input-output relations are linear, but with a constant value added to each input-output connection. This is indicated in Figure 1 as the values on the input-output links. The learning phase included two structure tests (Structure Test 1, Structure Test 2) which were designed to measure participants' declarative

knowledge of the underlying structure of the control system. In the control test phase, participants were required to manage the system in order to reach and maintain specific output criteria, and in that phase the criteria differed for each control test. This phase examined participants' ability to apply their procedural knowledge of the system to control the task.

Procedure Participants were told that they would be taking part in a problem solving task, and that they would be given an opportunity to learn about a water tank system. They were informed that their knowledge of the underlying structure of the system would be examined during this phase, and that they would be tested on their ability to control the system in two tests of control. The critical manipulations occurred during the learning phase between the four conditions included in this study.

Learning phase. The learning phase comprised 12 trials, which were divided into two blocks each with 6 trials. Each trial involved participants tracking changes in the values of an input: This was indicated by a slider that corresponded to moving automatically to a pre-specified value, which also appeared as a number above the input label. Each slider ranged on a scale from -100 to 100 units. Participants also clicked a button to reveal the effects of the changes in input values on the output values.

Participants began by clicking a button to reveal the pre-specified input values for the first trial (no time limit was imposed on the time spent studying the input values or output values on each trial). When they had studied the values of the inputs for that trial, participants clicked a second button to reveal the corresponding output values for that trial. As soon as they were ready, they clicked a button to indicate that they were proceeding to the next trial: The button hid the output values from view. Participants then repeated the process of seeing the input values, and then making the corresponding changes to the output values. After Trial 6, and after Trial 12, participants were presented with a structure task. This task was designed to index knowledge of the causal structure of the control system. A diagram of the water system was shown on screen, and participants were asked to indicate which inputs were connected to which outputs.

NSG condition. The NSG condition was given general instructions as to which features of the system to attend to when pressing particular buttons. Participants were also informed that they were observing a worker from the water purification plant inputting values into the system, in order to run some checks. They were told to pay close attention to the changes to inputs and outputs.

SG condition. In addition to the instructions presented to the NSG condition, the SG condition was told that, from the outset, they had to assess how effective the worker was at achieving and then maintaining specific output values (i.e., Oxygenation = 50, Chlorine CL Concentration = 700, Temperature = 900) throughout the 12 trials.

NSG-Hypo(-). The critical difference between the NSG and NSG-Hypo(-) was the inclusion of extra instructions that were based on those that Berry (1991) presented to participants: These were designed to discourage hypothesis testing behavior. "Subjects should be encouraged to pay attention to the observed interaction but not be induced to use a more deliberate hypothesis testing mode of performance" (Berry, 1991, p. 885). To achieve this, participants were told that they should avoid making any explicit predictions about the effects on the outputs, based on the changes to the inputs that they were observing. Instead, they should just focus on the changing inputs and the values of the outputs on each trial.

SG-Hypo(+). The critical difference between the SG and SG-Hypo(+) was the inclusion of extra instructions that were designed to encourage hypothesis testing behavior. Participants were told that, on each trial, they should examine their own knowledge about the underlying structure of the system, by using the changes in inputs and outputs as a way of testing and updating their own knowledge. They should make predictions based on their assumptions about the relationship between inputs and outputs, and use the values presented to them as a way of updating their knowledge of the system.

Control test 1: In this phase all participants were required to change the input values to achieve the following output values: Oxygen = 50, Chlorine concentration = 700, Temperature = 900. Participants were allocated 6 trials in which they were to reach and then maintain the output values given. *Control test 2*: As with Control Test 1, all participants were now required to change the input values to achieve given output values. However, the required output values were now: Oxygen = 250, Chlorine concentration = 350, Temperature = 1100.

Scoring

Structure Test scores. The scoring scheme used to score performance on Structure Tests 1 and 2 involved computing the proportion of input-output links correctly identified for each test. A correction for guessing was incorporated, and was based on the procedure used by Vollmeyer et al. (1996), which was simply correct responses (i.e., the number of correct links included, and incorrect links avoided) – incorrect responses (i.e., the number of incorrect links included, and correct links avoided)/ N (the total number of links that can be made). The maximum value for each structure score was 1. This scoring scheme was applied to score performance on both structure tests presented during the learning phase.

Control Test 1 and 2 scores. The scoring procedure used was based on Burns and Vollmeyer's scoring system. Control performance was measured as error scores. Error scores were obtained by calculating the difference between each target output value (i.e., the criterion according to the solution phase) and the actual output values produced by the participant, for each trial of the of the control test. To minimize the skewedness of the distribution of scores, a log

transformation (base 10) was applied to the error scores of each individual participant for each trial.

All analyses of error scores for each control test were based on participants' mean error score averaged over all 6 trials across all three output variables. Success in control performance in both control tests is indexed by the difference between the achieved and target output values. Therefore, the lower the error score, the better is the performance.

Results

Structure Test scores. Figure 2 shows that overall the mean structure test scores of the NSG and SG-Hypo(+) conditions were higher than those of the SG and NSG-Hypo(-) conditions, indicating that the presence of hypothesis testing behavior and self-evaluative thinking influenced the accuracy of participants' structural knowledge of the system.
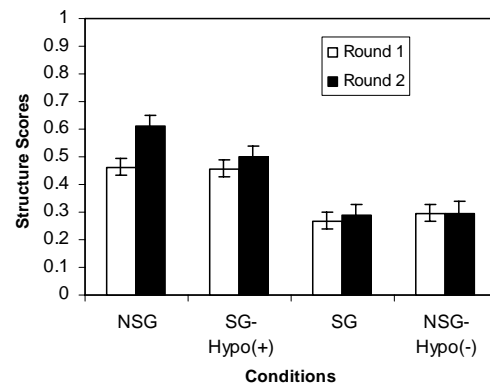

Figure 2: Mean Structure Test scores (±SE) after each block of the learning phase for each condition

A 2x4 ANOVA was conducted using block (Structure Test 1, Structure Test 2) as the within-subjects variable, and condition [NSG, SG, NSG-Hypo(-), SG-Hypo(+)] as the between-subjects variable. There was no significant main effect of block, $F(1, 59) = 2.31$, MSE = .10, and no interactions were significant. There was a significant main effect of condition on structure test scores, $F(3, 59) = 6.19$, MSE = .81, p < .001.

To examine the predictions outlined in Experiment 1, a priori comparisons were conducted using the Bonferroni t test. The analysis revealed significant differences between NSG and SG, $t = .28$, p< .05, and NSG and NSG-Hypo(-), $t = .29$, p< .05. The analysis also revealed significant differences between SG-Hypo(+) and SG, $t = .27$, p< .05, and SG-Hypo(+) and NSG-Hypo(-), $t = .28$, p< .05. There were no significant differences between NSG and SG-Hypo(+), and no significant differences between SG and NSG-Hypo(-). Thus, the evidence confirms the prediction that conditions in which hypothesis testing and self-evaluative thinking during learning are discouraged will show decrements in accuracy of knowledge, compared to conditions in which they are encouraged.

Control Tests 1 and 2 scores. Figure 3 includes the overall mean Control Tests 1 and 2 scores for each condition. Figure 3 shows that the NSG and SG-Hypo(+) conditions made fewer errors in Control Tests 1 and 2 than the SG and NSG-Hypo(-). In addition, Figure 3 suggests that participants made more errors in Control Test 2 than in Control Test 1. To analyze this, a 2x4 ANOVA was conducted to examine the patterns of behavior across conditions for Control Tests 1 and 2 scores, using test (Control Test 1, Control Test 2) as the within-subjects variable, and [NSG, SG, NSG-Hypo(-), SG-Hypo(+)] as the between-subjects variable. There was a main effect of test, $F(1, 59) = 5.74$, MSE =.12, $p < .05$, and a main effect of condition on control performance, $(3, 59) = 16.74$, MSE =.32, $p < .0005$. No other analyses were significant.
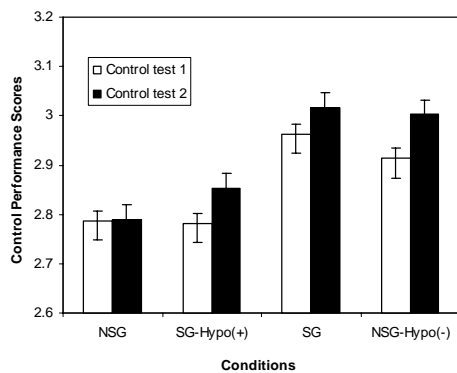

Figure 3: Control Test scores (±SE) at Control Test 1 and Control Test 2 for each condition

There was a main effect of test, $F(1, 59) = 5.74$, MSE =.12, The predictions outlined were examined using the Bonferroni t test. The analysis revealed significant differences between NSG and SG, $t = -.16$, $p < .001$, and NSG and NSG-Hypo(-), $t = -.14$, $p < .005$. The analysis also revealed significant differences between SG-Hypo(+) and SG, $t = -.21$, $p < .001$, and NSG-Hypo(-)(+) and NSG-Hypo(-), $t = -.18$, $p < .001$. There were no significant differences between NSG and SG-Hypo(+), and no significant differences between SG and NSG-Hypo(-). Thus, consistent with the pattern of results from the structure scores, decrements in control performance were found in conditions in which hypothesis testing and self-evaluative thinking were discouraged.

Correlation between Structure Test scores and Control Test scores. Osman (in press) claims that previous findings (Berry, 1991; Lee, 1995) showing dissociations between declarative and procedural knowledge in CDC-tasks are the result of attenuating hypothesis testing behavior and self-evaluating thinking. Therefore, conditions in which these behaviors have been encouraged [NSG, SG-Hypo(+)] should show associations between declarative and procedural measures of knowledge, whereas conditions in which these behaviors have been attenuated [SG, NSG-Hypo(-)] will show no associations. To examine this, a

correlation analysis between mean structure test scores (averaged over Structure Test 1 and Structure Test 2) and mean Control Test scores (averaged over Control Test 1 and Control Test 2) was conducted. The structure scores and control test scores were collapsed across the two conditions in which hypothesis testing was encouraged [NSG, SG-Hypo(+)]. The analysis revealed a significant negative relationship between structure test scores and control test scores, $r(32) = -.54$, $p < .001$. The same analysis was carried out on structure and control test scores collapsed across the two conditions in which hypothesis testing was attenuated [SG, NSG-Hypo(-)]. The findings indicate that there was no relationship between declarative and procedural measures of knowledge, $r(32) = -.28$, $p > .05$.

Discussion

The evidence from the experiment confirmed the main prediction. The evidence suggests that, even under observation-based learning conditions, goal specificity and, in particular, the presence of hypothesis testing and evaluative thinking, influence the acquisition and application of knowledge in dynamic control tasks. Consistent with Osman's (in press) findings, the goal specificity effect was found under observation-based learning conditions in which non-specific goal and specific goal instructions were presented. That is, the SG condition showed poorer control performance and poorer structural knowledge of the system than the NSG condition. However, decrements in control performance and structural knowledge were also found in the NSG-Hypo(-) condition, in which hypothesis testing and evaluative thinking was attenuated; whereas improvements were found in the SG-Hypo(+) condition, in which these behaviors were encouraged. This suggests that it is not specific goal instructions *per se* that lead to poorer problem solving ability in CDC-tasks, but rather that specific goal instructions tend to attenuate hypothesis testing and self-evaluating thinking, and this has damaging effects on the acquisition of declarative and procedural knowledge of a control task. The experiment also examined whether the source of the mixed findings (Berry, 1991; Lee, 1995; Osman, in press) concerning the type of relationship between declarative and procedural knowledge is to be found in whether hypothesis testing and self-evaluative processes are present during learning. The findings revealed that, when attenuated [SG, NSG-Hypo(-)], no association is found between declarative and procedural knowledge, whereas, when encouraged [NSG, SG-Hypo(+)], an association is found.

Thus, the present findings suggest that, in problem solving contexts, observation-based learners are sensitive, in the same way as procedural-based learners, to instructions that affect hypothesis testing and self-evaluative processes. Moreover, contrary to the claims made by Berry (1991), learning does occur under conditions in which there is no direct interaction with a CDC-task, and knowledge transfer from observation to action is achieved.

What therefore might explain the differences found between evidence for dissociations in studies of CDCTs and, as found in the present study, associations between procedural and declarative knowledge? Dissociations are typically found when exploration of the task via hypothesis testing is prevented (e.g., Berry, 1991; Lee, 1995). In addition, dissociations are usually reported in studies in which structural knowledge of the task is examined only after learning takes place (e.g., Berry, 1991; Berry & Broadbent, 1984). Without the opportunity to keep track of one's knowledge of the rule or structure the system operates under, explicit knowledge is found to be poor (Burns & Vollmeyer, 2002; Sanderson, 1989). When information search (i.e., NSG learning) is encouraged problem solvers tend to adopt a hypothesis testing strategy (Burns & Vollmeyer, 2002; Sweller, 1988; Vollmeyer et al., 1996). The present study suggests that information search can occur even under SG-learning conditions, so long as knowledge of the relations between inputs and outputs is examined through self-evaluative processes, or tested directly via hypothesis testing (Burns & Vollmeyer, 2002; Sanderson, 1989; Voss, Wiley & Carretero, 1995).

Conclusion

A number of dichotomies dominate research in problem solving (e.g., non-specific goals vs. specific goals, novice vs. skilled, rule vs. instance based learning). In the study of CDC tasks, one of the most imposing dichotomies is the distinction between declarative and procedural knowledge. One method by which this distinction has been demonstrated contrasts the effects of observation-based learning and of procedural-based learning. In so doing, the evidence suggests that observational learning leads to poorer problem solving ability (Berry, 1991; Lee, 1995). The present study examined observation-based learning of a problem solving task, and investigated the properties of goal instructions that produced both poor and good problem solving ability, and that lead to both dissociations and associations between declarative and procedural knowledge. The findings strongly suggest that claims concerning the detrimental affects of observational learning on problem solving ability have been overstated, and that what leads to successful problem solving in CDC-tasks are the same behaviors necessary when learning is procedural-based: that is, hypothesis testing and self-evaluative thinking.

Acknowledgments

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-406.

Berry, D. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology, 43*, 881-906.

Berry, D., & Broadbent, D. E. (1984). On the relationship between task performance and Associated verbalizable knowledge. *Quarterly Journal of Experimental sychology, 36*, 209-231.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, *81*, 211-241.

Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly urnal of Experimental Psychology, 55*, 241-261.

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, *4*, 3-23.

Kelly, S., & Burton, A. M. (2001). Learning complex sequences: No role for observation. *Psychological esearch, 65*, 15-23.

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

Lee, Y. (1995). Effects of learning contexts on implicit and explicit learning. *Memory and Cognition*, *23*, 723-734.

Locke, E. A., & Latham, G. P. (2000). Building a practically useful theory of goal setting and task motivation. *American sychologist*, 57, 705-717.

Nissen, M. J., & Bullemer, P. (1987). Attentional Requirements of Learning-Evidence from Performance-Measures. *Cognitive Psychology, 19*, 1-32.

Osman, M. (in press). Observation can be as effective as action in problem solving. *Cognitive Science*.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology, 118*, 219-235.

Sanderson, P. M. (1989). Verbalizable knowledge and skilled task performance: Association, dissociation, and mental models. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15*, 729-747.

Squire, L. (1986). Mechanisms of memory. *Science*, *232*, 1612-1619.

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

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. *Cognitive Science, 20*, 75-100.

Voss, J. F., Wiley, J., & Carretero, M. (1995). Acquiring Intellectual Skills. *Annual Review of Psychology*, *46*, 155-181.