

Observation Can Be as Effective as Action in Problem Solving

Magda Osman

Department of Psychology, University College London

Received 3 April 2006; received in revised form 31 October 2006; accepted 7 December 2006

Abstract

This study discusses findings that replicate and extend the original work of Burns and Vollmeyer (2002), which showed that performance in problem-solving tasks was more accurate when people were engaged in a non-specific goal than in a specific goal. The main innovation here was to examine the goal specificity effect under both observation-based and conventional action-based learning conditions. The findings show that goal specificity affects the accuracy of problem solving in the same way when the learning stage of the task is observation-based as when it is action-based. In addition, the findings show that, when instructions do not promote goal specificity, observation-based problem solving is as effective as action-based problem solving.

Keywords: Problem solving; Skill acquisition and learning; Observation vs. intervention

1. Introduction

1.1. Case 1

A food services manager requires, for his food products, a new warming display cabinet that can maintain a more even temperature across the shelf surfaces. The manager must first watch the engineer install and demonstrate the new multifunctioning thermal control system, based only on the precise temperature at which to display their food, and must then restart and control the system himself.

1.2. Case 2

A medical research unit that is conducting a field experiment requests a blood analyzing instrument that requires a specialized miniature air heater that must reach 380°F within 3.5 sec.

A member of the research unit is assigned to examine and learn to use the new custom-built air heater, according to the strict specifications that are outlined.

These cases have in common a precise goal that involves controlling a complex dynamic control (CDC) system (i.e., thermal control system, blood analyzing instrument) by adhering to specific criteria (i.e., temperature regulation). The critical distinction between these cases is that, in the former case, learning to control the system takes place indirectly through observation, whereas in the latter the individual must learn to control the system by interacting with it. This difference raises the question whether the execution of the specific goal (SG) is more accurate after learning in Case 2 in which acquisition of knowledge is an interactive activity, compared to Case 1 in which knowledge acquisition occurs indirectly through observation. It also raises the question whether learning to control a CDC system to specific criteria benefits an overall understanding of the system, in comparison to learning that takes place with no specific outcome set out.

These questions have been separately addressed in studies of CDC system in problem-solving research. Recently, studies (Burns & Vollmeyer, 2002; Vollmeyer, Burns, & Holyoak, 1996) have examined the effects on problem-solving ability when pursuing a SG instruction during learning—that is, when problem solvers are required to learn about a system while also controlling it to a specific criterion. The evidence reveals that this kind of instruction leads to decrements in the underlying knowledge of the structure of the system as well the ability to control the system to criteria other than those originally trained on. This is in contrast with problem solvers that are instructed to pursue a non-specific goal (NSG) during learning: i.e., they simply learn about the system they are presented with. Decrements in control performance have also been reported when learning about the control task is observation-based (i.e., acquiring knowledge indirectly) compared to when it is procedural-based (i.e., acquiring knowledge directly; Berry, 1991; Lee, 1995). Although the effects of goal specificity and mode of learning (i.e., observation-based, procedural-based) on problem solving have been examined using the same paradigm (i.e., CDC system), the two sets of findings remain independent of each other. The present study addresses this by examining whether action is a necessary component in the acquisition of skilled knowledge in a CDC task, as some theorists (e.g., Berry, 1991) have argued. If it is the case that procedural-based learning is necessary for skill learning to take place, then decrements in problem-solving ability should be found when learning is observation-based, regardless of the specificity of the goal in which the learner is engaged.

1.3. CDC tasks

CDC tasks have been a popular task environment for examining motivational and affective processes in complex decision making (Earley, Connolly, & Ekegren, 1989; Locke & Latham, 2002; Vancouver, 1997), skill learning in naturalistic decision making (Brehmer, 1992; Kerstholt, 1996; Lipshitz, Klein, Orasanu, & Salas, 2001), memory and attentional processes in problem solving (Burns & Vollmeyer, 2002; Miller, Lehman, & Koedinger, 1999; Vollmeyer et al., 1996), and implicit learning (Berry, 1991; Berry & Broadbent, 1988; Dienes & Fahey, 1995, 1998). The reason for this popularity is that tasks using CDC system are dynamic learning environments. Rarely is it the case that a decision or plan that is made in

the real world is isolated from prior decisions or plans (Funke, 2001), and in this sense CDC tasks provide a good model of actual situations in the world (Brehmer, 1992; Funke, 2001). This makes them ideal for studying the acquisition and transfer of skill-based knowledge in a variety of complex interactive environments (Campbell, 1988; Cañas, Quesada, Antoli, & Fajardo, 2003; Funke, 2001).

To illustrate, a typical CDC task includes several inputs (e.g., concentration levels of salt, carbon, and lime) that are connected via a complex causal structure or rule to several outputs (e.g., Chlorine concentration, Oxygenation levels, Temperature; see Fig. 1). The CDC system presented in Fig. 1 is taken from Burns and Vollmeyer's (2002) task, which was based on a water purification plant and will be used in the present study.

The process by which a problem solver learns about the system is revealed by the values of the inputs that they change and the strategy that they adopt (e.g., vary all inputs at once, vary one input on each trial, vary one input by one unit on each trial). Through this process, problem solvers acquire knowledge about the underlying structure of the system. They do this by monitoring the continuous feedback that they receive on the output variables that change as a result of the interventions that they perform on the inputs. That is, by manipulating the input values, problem solvers can then 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 (2002) example, the input–output relations are linear, but with a constant value added to each input–output connection.

To examine problem-solvers' knowledge of the system, two types of measures are used. Usually, the difference between the learning phase and the test phase of a CDC task is that, in the latter, problem solvers are required to control the system to a criterion. In the learning phase, the interventions on the system are designed to discover the underlying structure of the system. In the test phase involves making critical interventions that will achieve and maintain specific output values (i.e., the specific criterion) on each trial. Thus problem solvers are able to demonstrate their understanding of the CDC system by applying this knowledge to manipulate the system for the purposes of reaching a specific goal. In addition to this measure,

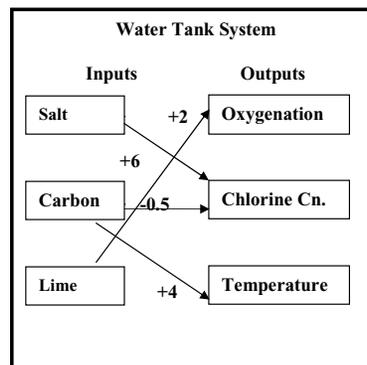


Fig. 1. Water tank system with inputs (salt, carbon, lime) and outputs (oxygenation, chlorine concentration, temperature).

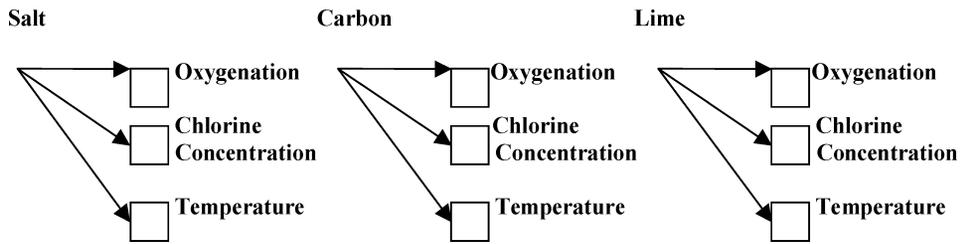


Fig. 2. Structure test.

which is referred to as the indirect measure (Berry & Broadbent, 1987), direct measures of knowledge examine the accuracy of problem-solvers' understanding of the causal structure or rule that underlies the system. This involves reporting the input–output connections (see Fig. 2; e.g., Burns & Vollmeyer, 2002; Vollmeyer et al., 1996), or the underlying rule that relates input to output (Lee, 1995; Lee & Vakoch, 1996), or predicting the output value based on specific input states of the system (Berry & Broadbent, 1987, 1988; Buchner, Funke, & Berry, 1995; Dienes & Fahey, 1995; Marescaux, Luc, & Karnas, 1989).

1.4. Procedural-based versus declarative-based inductive knowledge

Some researchers (Berry & Broadbent, 1987, 1988; Dienes & Fahey, 1995, 1998; Marescaux et al., 1989) have speculated that the knowledge acquired in dynamic control tasks is incidental. That is, learning is non-strategic, and that knowledge is acquired through direct interaction with a CDC task, and repeated exposure to it, which leads to the development of highly specialized specific operations (i.e., manipulating specific inputs to attain particular outputs). The proceduralization of these operations is identified by their being efficiently executed without concurrent conscious awareness of how they come about, and their inaccessibility to conscious control. Dissociation, as evidenced by good performance on indirect measures and poor performance on direct measures, support the claim that the knowledge gained in control tasks is procedural, and inaccessible to conscious awareness (Berry & Broadbent, 1987, 1988; Lee, 1995; Lee & Vakoch, 1996). Along with dissociations, implicit learning theorists have shown that procedural knowledge fails to generalize beyond CDC tasks that are perceptually and structurally similar to the original CDC task on which problems solvers were trained (Berry & Broadbent, 1988; Dienes & Berry, 1997; Sun, Merrill, & Peterson, 2001).

Thus, Berry (Berry, 1991; Berry & Broadbent, 1987, 1988) and others (Lee, 1995; Lee & Vakoch, 1996; Stanley, Mathews, Buss, & Kotler-Cope, 1989; Sun et al., 2001) describe control tasks as proceduralized tasks, because knowledge acquisition and application is incidental, and is neither regulated nor modulated by top-down declarative-based knowledge. This position is empirically supported by Berry's study in which, in a series of experiments, participants were presented with observation-based learning trials (i.e., participants were presented with the learning trials of other participants who had already taken part in the control task), and were tested under procedural-based conditions, (i.e., they were required to control the system to criterion through direct interaction). By so doing, Berry examined the potential conflicts that result during the test phase from declarative processes invoked by intentional

examination of the learning environment through observation, and from procedural processes invoked by direct interaction with the control system. Berry reasoned that the learning and test phases would engage incongruent processes, and that there would be a failure to apply declarative knowledge of the system acquired during learning in order to later control it in the test phase. Consistent with this, the control performance of observation-based learners was impaired compared to that of procedural-based learners. Berry concluded that action is a necessary component of learning in a control task, and that, for maximal control performance, decision-making must be tied to action early in the learning process. The distinction between procedural-based and declarative-based learning has since been reported in a variety of CDC tasks (e.g., Buchner et al., 1995; Lee, 1995; Lee & Vakoch, 1996).

In summary, the evidence from studies examining procedural-based learning in control tasks suggests that the transferability of the knowledge is dependent on the mode of learning in which knowledge acquisition takes place (i.e., observation-based, procedural-based), and the properties of task that knowledge is being applied to (i.e., the perceptual and structural similarity to the training task). By focusing on deliberate strategies to learn about and control a CDC system (i.e., through observation-based learning), the incidental procedural processes that would typically be invoked are disrupted, and the added cognitive effort that is incurred produces poorer control performance. Thus, for successful transfer to take place, knowledge acquisition and knowledge application require procedural based processing.

1.5. Goal specificity

CDC tasks have also been used to examine the presence of skilled rule learning and hypothesis-testing behavior (Burns & Vollmeyer, 2002; Sweller, 1988; Vollmeyer et al., 1996). Expertise in CDC tasks is identified by the ability to call to mind appropriate schemas (i.e., structures of knowledge that allow the expert to recognize a problem state as belonging to a particular category of problem states that typically require a specialized operation). These are developed from past experiences that can be transferred across domains (e.g., Trumpower, Goldsmith, & Guynn, 2004). Sweller (1988) claimed that achieving expertise in problem solving is dependent on the goal-directedness or specificity of the goal in which the solver is engaged. When goal directed, the solver is focused on achieving a particular outcome through means–end analysis (a method of reducing the distance between the current position in the problem and the end state). As a result they are unable to develop a deep understanding of the task. The process of means–end analysis interferes with the uptake of relevant knowledge through hypothesis testing, because during learning the solver is concerned only with serving the immediate demands of a specific goal (Sweller, 1988). Consequently, solvers are able to solve the problem according to the specific goal they are set, but they have poor knowledge of the general structure of that problem. By removing the goal-directed property of the CDC task, schema-based knowledge is promoted. This is achieved by presenting NSG instructions that are characterized as constraint free. This type of instruction is not goalless, because it encourages exploration of the problem, and places responsibility on the problem solver to determine the relevant properties of the CDC task. In contrast, SG instructions are constraint orientated. Typically, they require the solution of specific outcomes, and therefore problem

solvers are concerned with the particular operations that must be executed to achieve those outcomes.

The effects of goal specificity on problem-solving ability were first reported by Sweller (Sweller & Levine, 1982; Sweller, Mawer, & Ward, 1983) and have since been replicated (e.g., Burns & Vollmeyer, 2002; Geddes & Stevenson, 1997; Miller et al., 1999; Owen & Sweller, 1985; Sweller, 1988; Vollmeyer et al., 1996). Typically, the findings show that, when provided with a NSG instruction, problem solvers showed better performance on direct and indirect measures than SG-instructed problem solvers. In addition, they were able to transfer their knowledge across a variety of contexts.

Burns and Vollmeyer (2002) offer an alternative to Sweller's (1988) Cognitive Load theory of the goal specificity effect. Rather than relating the effects of goal specificity to different demands of cognitive load, the Dual Space hypothesis (Klahr & Dunbar, 1988; Simon & Lea, 1974) describes the goal specificity effect in terms of the problem solver's focus of attention in the problem-solving environment. Burns and Vollmeyer claim that a problem can be deconstructed into spaces: the rule space, which determines the relevant relationship between inputs and outputs; and the instance space, which includes examples of the rule being applied. Under SG instructions, the instance space is relevant because it is integral to the goal: The solver's attention is focused primarily on achieving a particular instantiation of the rule and not on discovering the rule itself. Under NSG instructions, because searching through the problem is unconstrained, both rule and instance spaces are relevant. Attention is distributed across all relevant properties of the task, because no one instantiation of the rule is more important than any other. In turn, searching through the rule space encourages hypothesis testing, which leads to a richer understanding of the underlying structure of the problem (e.g., Burns & Vollmeyer, 2002; Geddes & Stevenson, 1997; Renkl, 1997; Trumppower et al., 2004; Vollmeyer et al., 1996).

In summary, there is strong empirical evidence that during learning knowledge acquired under SG instructions is only relevant for those problems that follow the same goal structure as training, whereas knowledge gained under NSG instructions is transferable and so control performance generalizes across different set goals. Sweller's (1988) theory claims that the greater expenditure of cognitive effort incurred under SG instructions helps to guide the problem solver to the problem goal, but not to a deep understanding of the problem. This position differs somewhat from the Dual Space account, which claims that SG instructions focus attention toward specific states of the system, but away from learning about the underlying structure that relates inputs to outputs. The superficial knowledge that is acquired during SG instructions is in contrast with the extensive knowledge that is acquired during NSG instruction. Both theories agree that this type of instruction encourages hypothesis testing behavior, and that it is necessary in the acquisition and application of relevant knowledge in CDC tasks.

1.6. Summary

The claims made by learning theorists studying differences between procedural and declarative processing in CDC tasks conflict with those made by theorists concerned with the effects of goal specificity on skill learning. The former approach to studying CDC tasks suggests that the learning processes that underlie skill acquisition are bottom-up and unavailable for

conscious inspection (e.g., Berry & Broadbent, 1988; Dienes & Berry, 1997). The latter approach claims that top-down processes, such as hypothesis testing behavior, can be skill-based, and that control tasks help identify factors like goal specificity, that mitigate this kind of behavior (e.g., Burns & Vollmeyer, 2002; Sweller, 1988).

2. Present study

The principal objective of the present study was to compare the effects on control performance of a dynamic control task when learning is observation-based with the effects when it is procedural-based; and to establish whether this, rather than the goal specificity of the instructions, impairs problem-solving ability. To examine this, the present study used Burns and Vollmeyer's (2002) water-tank system and included four conditions: two in which the instructions presented in the learning phase of the control task were SG (SG-action, SG-observe) and two in which NSG instructions were presented (NSG-action, NSG-observe). If procedural processes are necessary to accurately control CDC tasks, then performance on indirect measures for both observation-based conditions should be inferior to that of both procedural-based conditions. If, however, hypothesis testing is a necessary component in the acquisition of knowledge in CDC tasks, then, regardless of the mode of learning (i.e., observation-based, procedural-based), SG learning should produce poorer performance on direct and indirect measures than NSG learning. This is because specific goals prevent hypothesis testing behavior, which is critical in the uptake of relevant knowledge in CDC tasks.

2.1. Participants

Sixty-four students from University College London volunteered to take part in the experiment and were paid £4 for their participation. Because participants in the observation-based learning conditions were yoked to a participant from the action-based learning conditions, the action-based conditions were run first. Participants were randomly allocated to one of two conditions (NSG-action, SG-action) and, when these conditions were run, a second randomization procedure was used to allocate participants to one of the two observation-based learning conditions (NSG-observe, SG-observe), with 16 in each of the four conditions. Participants were tested individually and were presented with a fully automated version of Burns and Vollmeyer's (2002) water purification system, which was run on Dell Optiplex computers. The experimental program was written in Visual Basic 6.

2.2. Materials

The study was based on Burns and Vollmeyer's (2002) water purification system. This involved three inputs (Salt, Carbon, Lime) and three outputs (Chlorine Concentration, Temperature, Oxygenation). The input-output relations are depicted in Fig. 1. At the beginning of each trial during the learning phase, and both tests of control, the starting values of the inputs were set to 0, and the values of the outputs were Chlorine Concentration = 500 units, Temperature = 1,000 units, and Oxygenation = 100 units.

2.3. Design

The present study included two types of goal specificity instructions (NSG, SG) and two modes of learning (action, observation), for a total of four conditions (NSG-action, SG-action, NSG-observe, SG-observe). Participants were presented with a learning phase that was divided into two short blocks, each with six trials. After the end of each block, participants were given a structure test in which they were asked to report what the causal structure of the CDC system was. After this, participants were presented with a test phase (Control Test 1, Control Test 2). The difference between conditions was in the instructions received (NSG, SG), and in the method (action, observation) by which learning took place during the learning phase. The effects of this were indexed using a direct measure of knowledge (Structure Test 1, Structure Test 2) during the learning phase, and an indirect measure of knowledge (Control Test 1, Control Test 2) during the control test phase.

2.4. 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 purification system during the learning phase of the task. They were also told that their knowledge of the system would be examined during the learning phase, and that their ability to apply this knowledge would be examined in two tests of control of the system in a later presented control test phase. The critical differences between the four conditions included in the present study were the learning phase, and the specificity of the instructions presented in that phase (for actual instructions used, see the appendix).

2.4.1. Action-based CDC task version

In the learning phase, participants were presented with a computer display (see Fig. 3) with three input variables and three output variables. The underlying structure that connects the inputs and outputs is presented in Fig. 1.

The learning phase comprised 12 trials, which were divided into two blocks of 6 trials. Each trial consisted of participants changing the value of any number of inputs, by using the slider corresponding to each.¹ Each slider ranged on a scale from -100 to 100 units. When participants were satisfied with their changes to the inputs, they clicked on a button labeled "output readings," which revealed the values of all three outputs. When they were ready to start the next trial, they clicked a button "next trial," which hid the output values from view. On the next trial, participants made their changes to the inputs, and these affected the output values from the previous trial: thus, the effects on the outputs were cumulative from one trial to the next. After the first block of 6 trials, participants were presented with a structure test that 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 simply to indicate which input was connected to which output (see Fig. 2). The direction of the input-output connection was implicit in the way that participants interacted with the CDC task, and was also indicated in the instructions to the task, and so it was not necessary to examine the directionality of the input-output relations, only which connections existed. After this, they began the next set of

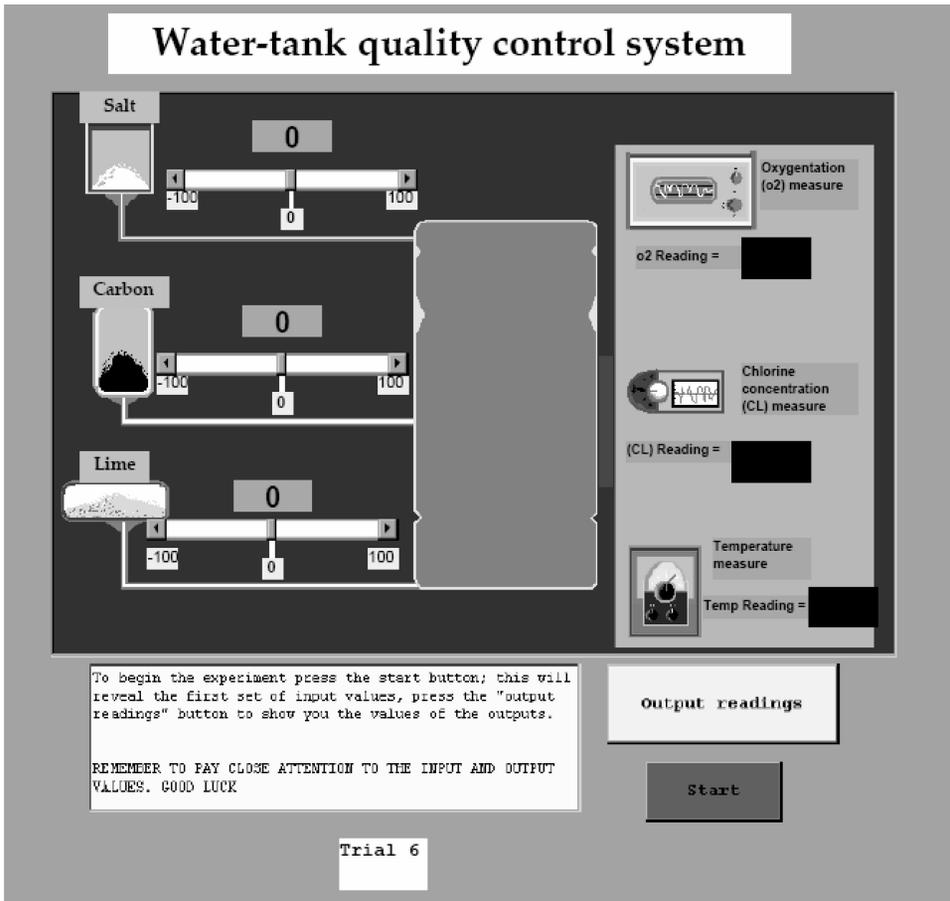


Fig. 3. Screen shot of water tank system.

6 trials, followed by a second structure test: At the beginning of the first trial of the second block, the input values were set to 0 and the outputs were also reset to their respective starting values.

The NSG-action condition was given general instructions about the CDC task and some guidelines as to how to interact with it (see the appendix). In addition to this, the SG-action condition was told that, from the outset, they had to learn about the system by trying to achieve, and then maintain, specific output values (i.e., Oxygenation = 50, Chlorine CL Concentration = 700, Temperature = 900) throughout the two blocks of the learning phase.

2.4.2. Observation-based CDC task version

With the exception that participants themselves could not manipulate input values during the learning phase, but were instead yoked to the values chosen by a participant (hereafter “model”) from the SG-action condition, the observation-based and action-based versions of the learning phase were identical. Observers began by clicking a button to reveal the input values generated by the model for the first trial. (No time limit was imposed on the time

spent studying the input values or output values on each trial.) For example, if the model changed the input Salt by 50 units on Trial 1, this would in turn change the output value of Chlorine Concentration to 556 (i.e., Chlorine Concentration starting value = 500 units, + Salt value change = 50 units, + constant added noise on input–output connection = 6 units). The observer would also see the input Salt change by 50 units. Then, when they were ready, participants clicked a second button to reveal the corresponding output values for that trial: in this case, the starting values of the outputs Temperature and Oxygen remained the same, but the corresponding output Chlorine Concentration changed to 556 units. As soon as they were ready, participants clicked a button to indicate that they were proceeding to the next trial: The button hid the output values from view. The participants then repeated the process of seeing the input values, and then the corresponding changes to the output values. As in the action-based version, after Trial 6, and after Trial 12, participants were presented with a structure test.

The NSG-observe condition was given general instructions as to which features of the system they should attend to when pressing particular buttons. In addition, the SG-observe condition was told that, from the outset, they had to assess how effective the changes made to the system were in achieving and then maintaining specific output values (i.e., Oxygenation = 50, Chlorine CL Concentration = 700, Temperature = 900) throughout the 12 trials. The output values that the SG-observe condition monitored, and the output values that the SG-action condition had to achieve and maintain, were identical.

2.4.3. Test phase (Control Test 1, Control Test 2)

After the learning phase, all participants completed two tests of control. In Control Test 1, they were required to change the input values to achieve and maintain the output values (Oxygen = 50, Chlorine concentration = 700, Temperature = 900) throughout 6 trials. For the NSG conditions (NSG-observe, NSG-action), this test was unfamiliar, whereas the SG-action condition had been performing exactly the same task for 12 trials in their version of the learning phase, and the SG-observe condition had been focused on assessing how effective the changes made to the system were in achieving and then maintaining the same output criteria. In Control Test 2, the output values (i.e., Oxygen = 250, Chlorine Concentration = 350, Temperature = 1,100) that participants had to achieve and maintain for 6 trials were unfamiliar to all four conditions. Thus, for the SG conditions, but not for the NSG condition, this was the first opportunity to apply their knowledge to a different goal to that on which they had been trained.

2.5. Scoring

2.5.1. 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 same 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 could be made). The maximum value

for each structure score was 1. This scoring scheme was applied to score performance on both the structure tests presented during the learning phase.

2.5.2. Control Tests 1 and 2 scores

The scoring procedure used was based on Burns and Vollmeyer's (2002) scoring system. Control performance was measured as error scores. Error scores were based on 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 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 six trials across all three output variables. Success in control performance in both control tests was indexed by the difference between the achieved and target output values. Therefore, lower error scores indicate better performance.

2.6. Results

This section begins with an analysis of performance on Structure Tests 1 and 2 (direct measure of knowledge), and then examines participants' ability to control the CDC task in Control Tests 1 and 2 (indirect measure of knowledge). Finally, correlation analyses were conducted to examine the potential relationship between indirect and direct measures of performance. In all analyses reported in this section, a significance criterion of $\alpha = .05$ was used.

2.6.1. Structure Test scores

Fig. 4 shows that structure test scores increased across all conditions after the second block of six trials in the learning phase. In addition, the mean structure test scores of the NSG-action and NSG-observe conditions were higher than those of the SG-action and SG-observe condition, indicating that the NSG conditions' knowledge of the causal structure of the system tended to be more accurate than that of SG conditions.

A $2 \times 2 \times 2$ analysis of variance (ANOVA) was conducted using block (Structure Test 1, Structure Test 2) as the within-subjects variable, and goal type (NSG, SG) and mode of learning (action, observation) as the between-subjects variables. The analysis revealed a significant main effect of block, $F(1, 60) = 14.04$, $MSE = .71$, $p < .0005$; and of goal type $F(1, 60) = 11.73$, $MSE = 1.37$, $p < .001$. There was no main effect of mode of learning, $F(1, 60) = .59$, $MSE = .07$, $p = .45$; and no interactions were significant. Thus, the evidence confirms the trends suggested in Fig. 4: There was increased accuracy in knowledge of the causal structure in the second block. The evidence also indicates that differences between conditions in performance on structure tests were the result of the specificity of the instruction presented to participants, and not of the mode of learning.

2.6.2. Control Tests 1 and 2 scores

Fig. 5 includes the overall mean Control Tests 1 and 2 scores for each condition. Fig. 5 shows that the NSG-action and NSG-observe conditions made fewer errors in Control Tests

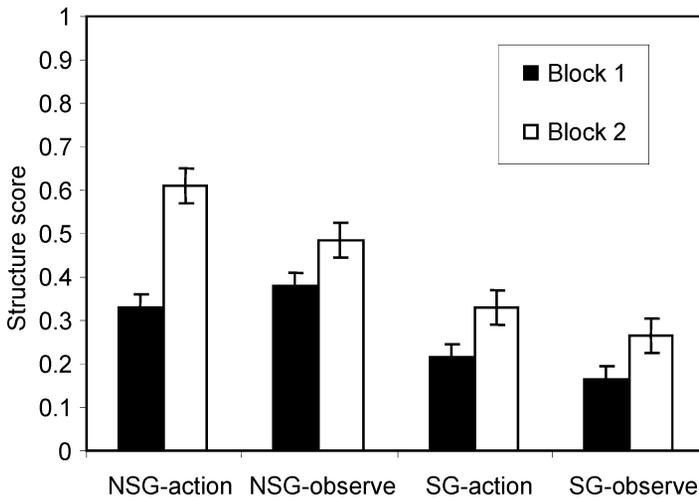


Fig. 4. Mean Structure Test scores ($\pm SE$) after each block of the learning phase for each condition; the lower the score, the better the performance. *Note:* NSG = non-specific goal; SG = specific goal.

1 and 2 than the SG-action and SG-observe conditions. In addition, Fig. 5 suggests that participants made more errors in Control Test 2 than in Control Test 1. To analyze this, a $2 \times 2 \times 2$ ANOVA was conducted to examine the patterns of behavior across conditions for Control Test 1 and 2 scores, using test (Control Test 1, Control Test 2) as the within-subjects variable and goal type (NSG, SG) and mode of learning (action, observation) as the between-subjects

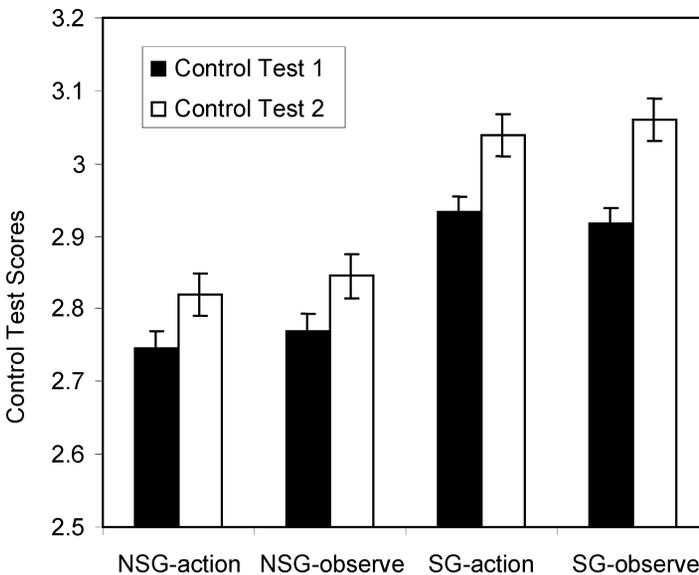


Fig. 5. Control Test scores ($\pm SE$) at Control Test 1 and Control Test 2 for each condition; the higher the score, the better the performance. *Note:* NSG = non-specific goal; SG = specific goal.

variables. The analysis revealed a significant main effect of test, $F(1, 60) = 13.55, p < .0005$; and of goal type $F(1, 60) = 40.33, p < .0005$. However, there was no main effect of mode of learning, $F(1, 60) = .22, p = .64$; nor were any of the interactions significant. Thus, the evidence confirms the trends suggested in Fig. 5: Control performance in Control Test 1 was superior to that in Control Test 2. The evidence also indicates that the difference between conditions in control performance was the result of the specificity of the instruction, and not of the mode of learning.

2.6.3. Correlation between structure test scores and control test scores

Berry and Broadbent (1987, 1988) have argued that dissociation between performance on direct and indirect measures of knowledge of CDC system suggests that the knowledge gained in control tasks is procedural. The basis for this is that individuals show good control of the system, but fail to express their knowledge when asked explicitly to report on the structure or the underlying rule that connects inputs to outputs (Berry & Broadbent, 1987, 1988; Buchner et al., 1995; Lee, 1995). To examine this, a correlation analysis between Structure Tests 1 and 2 scores and Control Tests 1 and 2 scores was conducted. The analysis revealed a significant negative relationship between Structure Test 2 and Control Test 1, $r(64) = -0.49, p < .005$; and between Structure Test 2 and Control Test 2, $r(64) = -0.53, p < .001$. This suggests that, as performance on Structure Test 2 increased, ability to control the system also increased, as indicated by lower control test scores. To fully explore the distinction between declarative and procedural knowledge, the same correlation analyses were conducted separately for the observation-based learning conditions (SG-observe, NSG-observe), and for the action-based learning conditions (SG-action, NSG-action). Focusing only on procedural-based conditions, the analysis revealed a significant negative relationship between Structure Test 2 and Control Test 1, $r(32) = -0.61, p < .0005$; and between Structure Test 2 and Control Test 2, $r(32) = -0.46, p < .01$. In the case of observation-based condition, the analysis revealed a significant negative relationship between Structure Test 2 and Control Test 1, $r(32) = -0.36, p < .05$; and between Structure Test 2 and Control Test 2, $r(32) = -0.60, p < .001$. The findings from these sets of analyses strongly indicate that, for both modes of learning (observation-based, procedural-based), there is a relationship between indirect and direct measures of knowledge.

3. Discussion

The evidence from the present study can be summarized as follows: First, the study successfully replicated Burns and Vollmeyer's (2002) findings, which showed that goal specificity affects the accuracy of performance in a CDC task. Second, consistent with Burns and Vollmeyer's study, the goal specificity effect was located in both indexes of knowledge of the control system (i.e., Structure Test scores, Control Test scores). Third, there was evidence of practice effects in the structure test scores, suggesting that, from Block 1 to Block 2, greater familiarity with the task environment in the learning phase led to increased knowledge of the causal structure of the system. This is consistent with the findings reported in Burns and Vollmeyer's study. However, there was poorer performance in Control Test 2 than in Control

Test 1, suggesting that participants found the criteria of the second control test more difficult to reach and maintain than the first.

Fourth, in the present study the performance of both NSG conditions was superior to the SG conditions in Control Test 1 and 2. This is particularly noteworthy given that the SG conditions had 12 trials of prior experience in evaluating the system to the same criterion as in Control Test 1, whereas the NSG conditions had not. However, in previous studies that used similar simulated systems to the present study (Burns & Vollmeyer, 2002; Vollmeyer et al., 1996) the control performance of NSG and SG conditions in Control Test 1 was equivalent, and only in Control Test 2 did NSG conditions show superior performance. The procedures used in the learning phase differed slightly between the present study and Burns and Vollmeyer's study (see footnote 1), which provides an explanation for the difference in findings. In the present study, the added effort required in accurately recalling the values of all the inputs and outputs imposed on any available resources SG condition had that would have otherwise been used to learn the relationship between the inputs and outputs. For the NSG conditions paying close attention to the input and output values and remembering them for the next trial favoured their method of learning. Through hypothesis testing they had to accurately monitor the consequences of their own/observed manipulations to the system, thus strengthening their knowledge of the system. Therefore, despite the extra experience that SG conditions had in evaluating the system to a criterion did not lead to the same results reported in previous studies because the extra working memory load made learning about the system that much harder than the NSG conditions.

Fifth, there was evidence that participants' knowledge of the structure of the system was related to their ability to control the system in both Control Test 1 and Control Test 2. Moreover, this relationship was found in conditions in which learning was observation-based as well as in conditions in which learning was action based. This result challenges the widely held popular view (e.g., Berry, 1991; Berry & Broadbent, 1988; Lee, 1995; Lee & Vakoch, 1996; Stanley et al., 1989; Sun et al., 2001) that CDC tasks promote proceduralized knowledge that is inaccessible to consciousness, and is therefore independent of declarative knowledge of the CDC system. Finally, the present study provides evidence that hypothesis testing, rather than procedural processes, is a necessary component in the acquisition of knowledge in CDC tasks, and that SG learning produces inferior problem-solving performance to NSG learning. Thus, the findings indicate that observation-based learning is sensitive to goal specificity effects, and can be as effective as procedural-based learning.

The present study provides evidence that declarative and procedural knowledge were associated, and that procedural learning *per se* is not a necessary condition for the acquisition of knowledge in a CDC task. In addition, the evidence showed that differences in performance between action and observation-based learning conditions were the result of goal specificity instructions, and not of how problem solvers interacted with the CDC task during learning. Taken together, the findings from the present study raise the following questions: Why is there a discrepancy between the present findings and previous studies of CDC tasks that show dissociation between procedural and declarative knowledge? Should action be awarded a special status in learning and problem solving? The following discussion examines issues related to addressing these questions.

3.1. *Dissociations and associations between procedural and declarative knowledge*

Funke's (2001) recent review suggests that studies of CDC tasks have followed two separate research histories: One has focused on decision making by experts within their domain of expertise (e.g., Dörner, 1996; Klein, 1990; Lipshitz, 1993; Lipshitz et al., 2001; Randel, Pugh, & Reed, 1996), whereas the other examines the development of expertise in novices (e.g., Berry & Broadbent, 1984, 1987, 1988; Burns & Vollmeyer, 2002; Sweller, 1998; Vollmeyer et al., 1996). One of the many differences found between experts and novices is that, for experts, their declarative knowledge and procedural knowledge are associated (Funke, 2001; Glaser & Bassok, 1989), whereas for novices there is dissociation. The reasons for this are that experts' knowledge is closely tied to their conceptions of the goal structures of a problem space (Glaser & Bassok, 1989), and their ability to organize specific instances of the CDC system (i.e., input–output states of the system), by chunking them into cause–effect sequences that relate goals to sub-goals (Sweller, 1988). Thus, one reason proposed for the dissociation between declarative and procedural knowledge reported in novices is that they do not acquire sufficient experience for an association to occur (e.g., Berry & Broadbent, 1987, 1988; Broadbent, Fitzgerald, & Broadbent, 1986; Glaser, 1976; Randel et al., 1996; Sternberg, Wagner, Williams, & Horvath, 1995).

Increasing experience with CDC systems, to enhance the relationship between novices' declarative and procedural knowledge, has been unsuccessful (e.g., Berry & Broadbent, 1987, 1988; Broadbent et al., 1986; Sanderson, 1989). However, Sanderson's careful manipulations uncovered conditions that revealed dissociations between declarative and procedural knowledge (i.e., increased practice, graphical representations of the CDC system), but also revealed associations (i.e., removal of problem-solver's trial history—a record of input–output changes on each trial during learning). The exclusion of a trial history during learning was inspired by Berry and Broadbent's (1984) finding. They showed that, when problem solvers gave verbal protocols during learning, as a way of tracking the input–output changes in the system and their hypothesis testing behavior, associations between declarative and procedural knowledge were found. Sanderson proposed that, without a record of their interactions with the system, problem solvers would engage in processes that were similar to those used in providing verbal protocols. Consistent with Berry and Broadbent's (1984) finding, Sanderson showed that, by mentally tracking their hypothesis testing strategies, problem solvers continually updated their knowledge of the input–output relations of the CDC task. This led to an association between declarative knowledge and procedural knowledge. However, Burns and Vollmeyer (2002) presented their participants a trial history of their learning phase and found an association between measures of declarative and procedural knowledge. The evidence from the protocols they recorded suggested that the tests of declarative knowledge included after the first and second block of the learning phase served as status checks. Participants examined their knowledge of the system after the first test after the first block of learning and used the remainder of the learning phase to modify this knowledge. Including tests of declarative knowledge during learning may have facilitated an association between declarative and procedural knowledge because, like Sanderson's study, participants were able to keep track of their knowledge of the system in order to learn to control it during learning.

Similarly, in the present study, problem solvers were presented with tests of their structural knowledge during learning, and were not presented with their trial history during the learning phase. Both these procedures may have contributed to the relationship that was found between indirect (i.e., procedural) and direct (i.e., declarative) measures of knowledge of the system. Moreover, the exclusion of a trial history, along with the inclusion of a measure of declarative knowledge during learning may have also contributed to the similar patterns in performance found in observation-based and action-based learning conditions. Without a trial history, both observation-based and action-based conditions had to keep active in memory the input–output changes from trial to trial. They also had to carefully monitor which inputs they were attending to, and their prediction of which outputs would be affected, and then compare this with the outputs that were actually affected. As Burns and Vollmeyer (2002) showed, compared to NSG instructions, attenuating hypothesis testing through SG instructions is detrimental to problem solving. Thus, in the present study, the instruction presented to SG conditions interrupted hypothesis testing behavior during the learning phase because participants' attention was focused only on recalling specific input–output instances that were relevant to the specific goal they had to reach and maintain (e.g., Buchner et al., 1995; Dienes & Fahey, 1998).

By extension, one reason for the evidence of poorer performance by observation based learners, compared to procedural based learners in CDC tasks (Berry, 1991; Lee, 1995), is the inclusion of trial histories during learning along with the presentation of SG type instructions (Berry, 1991; Lee, 1995). For example, in Berry's observation-based problem, solvers were actively discouraged from hypothesis testing: "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" (p. 885). Similarly, in Lee's observation-based learners were required to learn a rule passively without undertaking any evaluative thinking. Berry's and Lee's active discouragement of hypothesis testing was designed to increase control performance in later tests of problem-solving ability because hypothesis testing involves explicit monitoring and tracking of representations that interfere with procedural processes that are implicit (Berry & Broadbent, 1987). In this way, Berry's and Lee's instruction to avoid hypothesis testing functioned like an SG instruction and, in a similar way, produced poorer control performance than procedural-based learners who were not given an explicit instruction to avoid hypothesis testing.

The implication of studies showing associations between declarative and procedural knowledge (e.g., Berry & Broadbent, 1984; Burns & Vollmeyer, 2002; Sanderson, 1989; Sanderson & Vicente, 1986), along with the present study, is that mentally tracking or verbalizing one's hypothesis testing behavior is important in bridging declarative and procedural knowledge. Underlying these two activities is self-monitoring (Bandura, 1991), which has been shown to be critical in the acquisition of expertise (Bandura, 1991; Bandura & Locke, 2003; Ericsson & Lehman, 1996; Karoly, 1993; Rossano, 2003; Stanovich, 2004). In a given problem, selecting relevant information that bears on achieving a goal cannot happen without an internal status check of ongoing problem-solving performance (Bandura, 1991; Bandura & Locke, 2003; Karoly, 1993; Rossano, 2003). Once experts are sufficiently practiced in a task, self-monitoring is no longer needed to regulate goal relevant information, (Karoly, 1993); but it remains active in order to modify behavior to meet potentially new task demands (Rossano,

2003; Stanovich, 2004). In studies of CDC tasks, findings show that experts' declarative and procedural knowledge is associated (e.g., Funke, 2001), and that self-monitoring is necessary in the acquisition of expertise (e.g., Bandura, 1991; Schraw, 1998). When novices are encouraged to adopt behaviors that are akin to self-monitoring, associations between declarative and procedural knowledge are also found (Berry & Broadbent, 1984; Sanderson, 1989; Sanderson & Vicente, 1986). Thus, self-monitoring of one's behavior, its determinants, and its effects is, for experts, an essential method of acquiring, organizing, and applying different types of knowledge in problem-solving domains (Bandura, 1991; Bandura & Locke, 2003; Funke, 2001; Schraw, 1998).

3.2. Should action be awarded a special status in learning and problem solving?

The role of action has been strongly emphasized, not only in learning and problem-solving research, but also in education and instructional psychology (i.e., the methods and conditions under which instruction facilitates human learning and development). The constructivism tradition (Anderson, 1987; Resnick, 1983, 1987; Schauble, 1990; von Glasersfeld, 1989) asserts that action plays a crucial role for learners in constructing their own knowledge. In an educational context, active engagement entails students examining their own ideas, considering alternative explanations for newly taught concepts, and evaluating competing perspectives. Some theorists (e.g., Anzai & Simon, 1979) propose that these processes are found when "learning is by doing." However, a constructivist perspective essentially points to the need for instructional formats that allow for self-monitoring (e.g., Covington, 2000; Pintrich & De Groot, 1990), which includes reflective activities such as describing, explaining, and evaluative thinking (e.g., Covington, 2000; Zimmerman, 1990), and these need not occur through action.

Another example in which the superior status of action over observation is emphasized is the intention-superiority effect (e.g., Goschke & Kuhl, 1993). This refers to the phenomenon in which items related to activities about to be completed are recalled faster and more accurately than observed activities (Goschke & Kuhl, 1993). It has since been shown that memory of intended actions is not only superior to that of observed actions, but also of completed actions (Marsh, Hicks, & Bink, 1998; Maylor, Chater, & Brown, 2001; Maylor, Darby, & Della Sala, 2000). Moreover, there is evidence that the effect is dependent on personality variables (i.e., action-dependent vs. state-dependent dispositions) that influence self-regulation and self-monitoring (e.g., Goschke & Kuhl, 1993; Kuhl & Beckmann, 1994; Kuhl & Goschke, 1994; Penningroth, 2005). Penningroth showed that, for action-orientated participants who show a tendency to reflectivity, retrieval was more proficient for intended acts than for acts that had taken place. In contrast, there was a reversal of this effect in action-orientated participants who show a tendency to be instinctive, spontaneous, and unreflective.

The special status that action has been awarded over observation has been proposed in learning (e.g., Berry, 1991; Kelly & Burton, 2001; Lee, 1995), memory (e.g., Goschke & Kuhl, 1993), problem solving (e.g., Berry, 1991; Lee, 1995), and education (e.g., Anzai & Simon, 1979; Resnick 1983, 1987). This is particularly because observation is characterized as passive and uninvolved. However, as studies of CDC tasks (e.g., Berry & Broadbent, 1984; Burns & Vollmeyer, 2002; Sanderson, 1989) and the present study suggest, the level of involvement of

self-monitoring during learning may be a critical factor in preserving the distinction between action-based and observation-based learning in problem-solving environments.

4. Conclusions

The evidence convincingly replicated the goal specificity effect and demonstrated that this effect generalized to an observation-based learning mode of problem solving. Moreover, the evidence from this study shows that observational learning can promote successful and accurate problem solving in a dynamic control task. In addition, it is sensitive to goal specificity effects in the same way as action-based problem solving. Finally, the evidence suggests that it is hypothesis testing, rather than the procedural element of control tasks, which is necessary for the successful uptake of knowledge and its practical application in mastering a complex system.

Notes

1. In Burns and Vollmeyer's (2002) study, participants were shown the starting values of input and output values before they began the task. They were also presented with their own trial history; this included the inputs and outputs that changed on each trial. In this experiment, participants were shown only the starting values of the input values and not the output values, which were revealed only on the first trial, and not before. In addition, they were not presented with a trial history that they could refer to during the learning phase. The rationale for these changes was to encourage participants to pay special attention to and accurately monitor and recall the effects on the outputs resulting from the manipulations that were made to the system.

Acknowledgments

Preparation for this article was supported by Economic and Social Research Council (ESRC) Grant RES-000-27-0119. The support of the ESRC is gratefully acknowledged. The work was also part of the program of the ESRC Research Centre for Economic Learning and Human Evolution. I thank the reviewers for their incisive and helpful comments on earlier drafts of this manuscript.

References

- Anderson, C. W. (1987). Strategic teaching in science. In B. F. Jones, A. S. Palincsar, D. S. Ogle, & E. G. Carr (Eds.), *Strategic teaching and learning: Cognitive instruction in the content areas* (pp. 73–91). Alexandria, VA: Association for Supervision and Curriculum Development.
- Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, *86*, 124–140.
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*, 248–287.
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88*, 87–99.

- 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 Psychology*, 36, 209–231.
- Berry, D., & Broadbent, D. E. (1987). The combination of implicit and explicit knowledge in task control. *Psychological Research*, 49, 7–15.
- Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit–explicit distinction. *British Journal of Psychology*, 79, 251–272.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81, 211–241.
- Broadbent, D., Fitzgerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology*, 77, 33–50.
- Buchner, A., Funke, J., & Berry, D. (1995). Negative correlations between control performance and verbalizable knowledge: Indicators for implicit learning in process control tasks. *Quarterly Journal of Experimental Psychology*, 48, 166–187.
- Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly Journal of Experimental Psychology*, 55, 241–261.
- Campbell, D. (1988). Task complexity and strategy development: A review and conceptual analysis. *Academy of Management Review*, 13, 40–52.
- Cañas, J. J., Quesada, J. F., Antoli, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics*, 46, 482–501.
- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative view. *Annual Review of Psychology*, 51, 171–200.
- Dienes, Z., & Berry, D. C. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, 4, 3–23.
- Dienes, Z., & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 848–862.
- Dienes, Z., & Fahey, R. (1998). The role of implicit memory in controlling a dynamic system. *Quarterly Journal of Experimental Psychology*, 51, 593–614.
- Dörner, D. (1996). *The logic of failure. Recognizing and avoiding error in complex situations*. Reading, MA: Addison-Wesley.
- Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, 74, 24–33.
- Ericsson, K. A., & Lehman, A. (1996). Expert and exceptional performance: Evidence of maximal adaptation to constraints. *Annual Review of Psychology*, 47, 273–305.
- Funke, J. (2001). Dynamic systems as tools for analyzing human judgment. *Thinking and Reasoning*, 7, 69–89.
- Geddes, B. W., & Stevenson, R. J. (1997). Explicit learning of a dynamic system with a non-salient pattern. *Quarterly Journal of Experimental Psychology*, 50A, 742–765.
- Glaser, R. (1976). Cognitive psychology and instructional design. In M. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. XV–XXVIII). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, 40, 631–666.
- Goschke, T., & Kuhl, J. (1993). Representation of intentions: Persisting activation in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 1211–1226.
- Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology*, 44, 23–52.
- Kelly, S., & Burton, A. M. (2001). Learning complex sequences: No role for observation. *Psychological Research*, 65, 15–23.
- Kerstholt, J. H. (1996). The effect of information costs on strategy selection in dynamic tasks. *Acta Psychologica*, 94, 273–290.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1–55.
- Klein, G. A. (1990). Knowledge engineering: Beyond expert systems. *Information Technologies*, 16, 27–41.

- Kuhl, J., & Beckmann, J. (1994). *Volition and personality: Action versus state orientation*. Seattle, WA: Hogrefe & Huber.
- Kuhl, J., & Goshke, T. (1994). State orientation and the activation and retrieval of intentions in memory. In J. Kuhl & J. Beckmann, (Eds.), *Volition and personality: Action versus state orientation* (pp. 127–153). Seattle, WA: Hogrefe & Huber.
- Lee, Y. (1995). Effects of learning contexts on implicit and explicit learning. *Memory and Cognition*, 23, 723–734.
- Lee, Y., & Vakhoch, D. (1996). Transfer and retention of implicit and explicit learning. *British Journal of Psychology*, 87, 637–651.
- Lipshitz, R. (1993). Converging themes in the study of decision-making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood, & C. Zsombok (Eds.), *Decision-making in action: Models and methods* (pp. 103–137). Norwood, CT: Ablex.
- Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, 14, 332–251.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist*, 57, 705–717.
- Marescaux, P.-J., Luc, F., & Karnas, G. (1989). Modes d'apprentissage selectif et nonselectif et connaissances acquies au control d'un processus: Evaluation d'un modele simule [Selective and nonselective learning modes and acquiring knowledge of process control: Evaluation of a simulation model]. *Cahiers de Psychologie Cognitive*, 9, 239–264.
- Marsh, R. L., Hicks, J. L., & Bink, M. L. (1998). Activation of completed, uncompleted, and partially completed intentions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24, 350–361.
- Maylor, E. A., Chater, N., & Brown, G. D. A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bulletin & Review*, 8, 162–167.
- Maylor, E. A., Darby, R. J., & Della Sala, S. (2000). Retrieval of performed versus to-be-performed tasks: A naturalistic study of the intention-superiority effect in normal aging and amnesia. *Applied Cognitive Psychology*, 14, 83–98.
- Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science*, 23, 305–336.
- Owen, E., & Sweller, J. (1985). What do students do while solving mathematics problems? *Journal of Educational Psychology*, 77, 272–284.
- Penningroth, S. L. (2005). Free recall of everyday retrospective and prospective memories: The intention-superiority effect is moderated by action versus state orientation and by gender. *Memory*, 13, 711–724.
- Pintrich, P. R., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40.
- Randel, J. M., Pugh, L., & Reed, S. K. (1996). Differences in expert and novice situation awareness in naturalistic decision making. *International Journal of Human Computer Studies*, 45, 579–597.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, 21, 1–29.
- Resnick, L. B. (1983). Mathematics and science learning: A new conception. *Science*, 220, 477–478.
- Resnick, L. B. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
- Rossano, M. J. (2003). Expertise and the evolution of consciousness. *Cognition*, 89, 207–236.
- 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.
- Sanderson, P. M., & Vicente, K. J. (1986). *Verbalizable knowledge and skilled task performance: Explaining association and dissociation* (Tech. Rep. No. EPL-86-04). Urbana, IL: University of Illinois at Urbana-Champaign, Engineering Psychology Research Laboratory.
- Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology*, 49, 31–57.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113–125.
- Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L. W. Gregg (Ed.), *Knowledge and cognition* (pp. 105–127). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Stanley, W. B., Mathews, R. C., Buss, R. R., & Kotler-Cope, S. (1989). Insight without awareness: On the interaction of verbalization, instruction, and practice in a simulated process control task. *Quarterly Journal of Experimental Psychology*, *41*, 553–577.
- Stanovich, K. E. (2004). *The robot's rebellion*. Chicago: University of Chicago Press.
- Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist*, *50*, 912–927.
- Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, *25*, 203–244.
- Sweller, J. (1988). Cognitive load during problem solving: Effects of learning. *Cognitive Science*, *12*, 257–285.
- Sweller, J. (2003). Evolution of human cognitive architecture. *Psychology of Learning and Motivation: Advances in Research and Theory*, *43*, 215–266.
- Sweller, J., & Levine, M. (1982). Effects of goal specificity on means–end analysis and learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 463–474.
- Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, *11*, 639–661.
- Trumppower, D. L., Goldsmith, T. E., & Guynn, M. (2004). Goal specificity and knowledge acquisition in statistics problem solving: Evidence for attentional focus. *Memory & Cognition*, *32*, 1379–1388.
- Vancouver, J. B. (1997). The application of HLM to the analysis of the dynamic interaction of environment, person and behavior. *Journal of Management*, *23*, 795–818.
- 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.
- von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, *80*, 121–140.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychology*, *25*, 3–17.

Appendix

Standard action instructions:

You are a trainee laboratory technician working in a water filtration unit. As part of your training you will learn to control the water tank system by managing three water quality measures: Oxygenation; Chlorine CL concentration; Temperature. The quality measures are known as outputs and are used to monitor three system inputs: Salt; Carbon; Lime. In the following task you will be presented a total of 12 trials in which you will see a diagram of the ‘Malwart’ water filtration unit which you will learn to control. You can modify the quality measures by manipulating the amount of Salt, Carbon, or Lime inputs; this can be done by moving the slider corresponding to the input either to the left or to the right.

NSG-action condition also received:

For each trial you should try to change one input, however this is only a recommendation and you may choose to use a different strategy. Once you have changed the value of an input you can then check the output levels by pressing the button labeled ‘show me readings’; this will reveal the concentration levels of the quality measures. After you have studied these you should press the ‘restart’ button to begin the next trial. You should try and pay close attention to the values of the inputs you enter into the system and the output levels because this will help you to learn about the system. Good Luck!

SG-action condition also received:

For each trial you should try to change one input, but this is only a recommendation and you may choose to use a different strategy. Once you have done this you can check the output levels by pressing the button labelled 'show me readings'; this will reveal the concentration levels of the quality measures. After you have studied these you should press the 'restart' button to begin the next trial. Your task will be to change the output levels so that Oxygenation = 50, Chlorine CL Concentration = 700, Temperature = 900. Try to get as close to these levels as possible, and once you have done this try to maintain these levels throughout. Good Luck!

Standard observation instructions:

You are a trainee laboratory technician working in a water filtration unit. As part of your training you will learn to control the water tank system by managing three water quality measures: Oxygenation; Chlorine CL concentration; Temperature. The quality measures are known as outputs and are used to monitor three system inputs: Salt; Carbon; Lime.

In the following task you will be presented with a series of trials in which you will see a diagram of the 'Malwart' water filtration unit which you will learn to control. The system is set so that the quality measures change according to the values chosen by one of the workers of the water plant. You will see the amount of Salt, Carbon, and Lime inputs change automatically according to those set by the worker, this is indicated by a slider corresponding to each input moving either to the left or the right. You will see a total of 12 trials divided into two short sessions of 6 each.

NSG-observe condition also received:

For each trial you should watch carefully the changes to the inputs. When you have examined the changes to the inputs you can check the output levels by pressing the button labeled 'Output readings.' This will reveal the concentration levels of the quality measures. After you have studied these you should press the 'Input levels' button to begin the next trial. You should try and pay close attention to the values of the inputs that are entered and the output levels. Good Luck!

SG-observe condition also received:

For each trial you should watch carefully the changes to the inputs. When you have examined the changes to the inputs you can check the output levels by pressing the button labeled 'Output readings.' This will reveal the concentration levels of the quality measures. After you have studied these you should press the 'Input levels' button to begin the next trial. For each trial, your task will be to assess how well the worker of the water plant successfully achieved the following output levels so that Oxygenation = 50, Chlorine CL Concentration = 700, Temperature = 900. Good Luck!