Anti-Learning: Negative Transfer in an Inductive Learning Task

Magda Osman

University College London

Department of Psychology

University College London

Gower Street

London WC1E 6BT

England

Phone: +4420 7679 7572

Fax: +4420 7436 4276

Email: m.osman@ucl.ac.uk

Abstract

The present study examined how self-regulatory processes influenced problem solving behavior under conditions in which people were either re-exposed to their prior learning, or experienced someone else's. In a series of four experiments, participants solved two complex control tasks that were identical in structure but varied in presentation format. Participants either learnt to solve the second task, based on their original learning phase from the first task, or learnt to solve the second task based on another participant's learning phase. Experiment 1 showed that, under conditions in which participants' learning phase was experienced twice, performance deteriorated in the second task compared to the first. In contrast, when the learning phases in the first and second task differed, performance improved in the second task. Experiment 2 introduced instructional manipulations that induced the same response patterns as Experiment 1, and Experiment 3 showed that participants differentially weighted the learning phase in the second task. In Experiment 4, judgments of self-efficacy were shown to track control performance. The implications of these findings for induction, self-regulation, and insight are discussed.

Keywords

Induction, self-regulation, control, observation vs. action, skill learning

Anti-Learning: Negative Transfer from an Inductive Learning Task to Itself

Example 1: Pilot A is training to fly the Boeing 737NG (next generation) plane. In a flight simulation, the pilot flies the plane on a two-hour night flight. The training schedule includes the tutor replaying Pilot A his flight profile in order to help him assess his performance. Example 2: Pilot B experiences the same initial training routine as pilot A, with one exception: after his flight, he is played the flight profile of pilot A, not his own. A final briefing session reviews both pilots' competence, and assesses how to transfer their training successfully to new flight patterns. These examples share a precise goal that involves accurately and reliably controlling a complex dynamic control task (CDCT—i.e., the aircraft). Control behavior in this, and other types of CDCTs (e.g., air-traffic control, naval mine sweeping, car driving) depends on the development of expertise (e.g., Brehmer, 1992; Funke, 2001; Gonzalez, Lerch, & Lebiere, 2003; Kerstholt, 1996; Randel & Pugh, 1996; Sweller, 1988; 2003).

Expertise involves both *monitoring*—through self-evaluation of goal-directed actions (e.g., Bandura & Locke, 2003; Ericsson & Lehman, 1996; Karoly, 1993; Lerch & Harter, 2001; Rossano, 2003; VanLehn, 1996)—and *control*, or the generation and selection of goal-directed actions (e.g., Lerch & Harter, 2001; Locke & Latham, 2002; Rossano, 2003; Sweller, 1988; VanLehn, 1996; Vollmeyer, Burns, & Holyoak, 1996). Both behaviors are also interrelated (e.g., Bandura & Locke, 2003; Ericsson & Lehman, 1996; Karoly, 1993; Lerch & Harter, 2001; Rossano, 2003; VanLehn, 1996). However, studies of skill learning in CDCTs have overlooked this relationship, particularly how it influences the transferability of expertise (e.g., Berry, 1991; Berry & Broadbent, 1988; Burns & Vollmeyer, 2002). Without an understanding of how individuals monitor their behavior, little can be said about how evaluative processes are employed when transferring learnt skills to achieve unpracticed goals.

For instance, the critical difference between the examples above is that, in the first case, training (i.e., control performance) and assessment (i.e., monitoring) are based on self-generated behavior; whereas, in the second, assessment is based on comparing self- and other-generated training behavior. These types of training procedures are often used in educational (Pintrich & De Groot, 1990; Pressley, & Ghatala, 1990; Zimmerman, 1990), clinical (Bailey, & Sowder, 1970; Giesler, Josephs, & Swann, 1996; Griffiths & Gillingham, 1978), and military domains (e.g., Gandhe, Gordon, Leuski, Traum, & Oard, 2004; Hill, Gordon, & Kim, 2004). How will the two pilots' different learning experiences impact on their later ability to transfer their knowledge to similar and different goals? The present article addresses a related and, thus far, unexplored question: How does monitoring affect the transfer of control behaviors in a complex skill learning task?

Monitoring: Self-Regulatory Mechanisms

Bandura's (1986) Social Cognitive theory places evaluative processes that regulate motivation and actions at the heart of human cognition. Pursuing a goal is guided by selectively attending to qualitative and quantitative aspects of ongoing behavior, which enables people to evaluate the status of their behavior in relation to a goal (Bandura & Locke, 2003; Karoly, 1993). In this way, monitoring or self-regulatory mechanisms track goal-relevant information, modulate motivation, and trigger self-reflective judgments. The mechanism of perceived self-efficacy occupies a central role in the regulation of motivation and action. This refers to people's beliefs in their ability to exercise control over environmental events, with which people regulate motivational (e.g., Litt, 1988; Moritz et al., 2000), affective (e.g., Bandura & Cervone, 1986; DeShon & Alexander, 1996; Elliot & Dweck, 1988; Spering, Wagener, & Funke, 2005), and decisional processes (e.g., Earley, Connolly, & Ekegren, 1989; Kanfer et al., 1994; Tversky & Kahneman, 1974). For instance, perceived self-efficacy is shown to mediate problem solving so that it produces poor performance, irrespective of people's actual capabilities (Bandura & Wood, 1989; Bouffard-Bouchard, 1990; Hogarth et al., 1991; Wood & Bandura, 1989). Equally compelling is evidence that increasing people's belief in their self-efficacy guides attentional processes so that, in problem solving tasks, people's accuracy in detecting and analyzing solutions to problems can be radically improved (e.g., Bouffard-Bouchard, 1990; Jacobs et al., 1984).

Social Cognitive theory proposes that self-regulatory mechanisms operate dual control systems (Bandura, 1991): i.e., goals can be met through discrepancy reduction, such as error detection and correction, which is an example of a reactive control system. In concert with this goes proactive discrepancy production, which involves incrementally setting difficult challenges that broaden one's knowledge and experience of a skill. Studies of expertise show that both systems (Bandura, 1991) are critical in the acquisition of complex behaviors, ranging from athletic and musical performance, to managerial decision making and stockbroking (Bandura, 1991; Bandura & Locke, 2003; Cohen, Freeman, & Wolf, 1996; Ericsson & Lehman, 1996; Karoly, 1993; Rapoport, 1967; Rossano, 2003; Stanovich, 2004). Selecting relevant information bearing on a solution requires keeping an internal status check of ongoing performance through error detection and correction (Bandura, 1991; Bandura & Locke, 2003; Karoly, 1993; Lehmann & Ericsson, 1997; Rossano, 2003). Additionally, through judgments of self-efficacy, experts advance their behavior towards achieving increasingly more and more difficult goals, thus extending their knowledge and experience (e.g., Bandura & Wood, 1989; Bouffard-Bouchard, 1990; Wood & Bandura, 1989).

Regulatory Mechanisms through Self-Observation

Our capacity to learn vicariously (Bandura, 1986, 2002) also means that self-regulatory mechanisms can modulate behaviors learnt from both modeled and active experience. The self-observation technique involves re-exposing individuals to their own self-generated behaviors: In so doing, the technique enables comparison of subjective experiences with objective representations of them. For example, developmental studies (Fireman & Kose, 1991, 2002; Fireman, Kose, & Solomon, 2003; Fosnot et al., 1988) report improvements in children's ability to solve the Tower of Hanoi (TOH) task through video-taped presentations of their previous solutions. In Fireman et al.'s 2003 study, children completed the TOH task and were then shown their own moves, or another child's previous inefficient moves, or another child's correct completion of the task. They were then presented with another TOH task. Children benefited more from observing their own previous inefficient problem solving strategies than the behaviors of other children.

Studies using the self-observation technique (e.g., Albright & Malloy, 1999; Fireman & Kose, 1991, 2002; Fireman, Kose, & Solomon, 2003; Fosnot, Forman, Edwards, & Goldhaber, 1988; Knoblich & Flach, 2001; Knoblich & Prinz, 2001; Loula, Prasad, Harber, & Shiffrar, 2005; Storms, 1973) have shown that individuals can learn from re-exposure to their actions through a host of media (e.g., video recordings, photographic stills, point light displays [videos of people making movements are reduced to lights on the joints]). The technique works because self-correcting procedures allow individuals to detect and improve on previously generated behaviors (e.g., decision making, meta-perception, motor learning). It is thus also used in education (e.g., Covington, 2000; Pintrich & De Groot, 1990; Zimmerman, 1990), and therapeutically (e.g., Bailey & Sowder, 1970; Dowrick, 1983; Giesler et al., 1996; Hung & Rosenthal, 1978).

However, studies using the self-observation technique focus on the accuracy of detecting self-generated behaviors, and any subsequent improvements. They show self-regulatory mechanisms operating over veridical representations, but provide little insight into how people use these mechanisms on internally represented behaviors. To shed light on this, this study examines self-regulatory mechanisms and their effects on subsequent transfer of skilled control behaviors, by re-exposing problem solvers to products of their own strategic thinking (e.g., decision-making and hypothesis testing behavior), rather than visual (i.e., video) presentation of their selves performing. It thus becomes possible to empirically control the information that people's self-regulatory mechanisms are operating on, and to examine the subsequent impact on the transfer of their control behaviors.

Complex Dynamic Control Tasks (CDCTs)

CDCTs have been a popular task environment for examining many phenomena, including 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 et al., 1999; Vollmeyer et al., 1996), and implicit learning (Berry, 1991; Berry & Broadbent, 1988; Dienes & Fahey, 1995, 1998). Their popularity and range make 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). A typical CDCT includes several inputs connected to several outputs via a complex causal structure or rule (See Figure 1). The CDCT in Figure 1 is from Burns and Vollmeyer's (2002) task, which was based on a water tank 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 they change and the strategy 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. That is, by manipulating the input values, problem solvers can then track the effects on the outputs, and then reason from cause (input changes) to effect (output changes), via acquisition of the causal structure or rule that relates inputs and outputs. To examine problem solvers' knowledge of the system, two types of measures (direct, indirect) are used. In the learning phase, changes to the inputs are designed to discover the underlying structure of the system. Indirect measures index the problems solvers ability to control the system by changing the inputs to reach a SG. Direct measures of knowledge also examine the accuracy of problem solvers' declarative knowledge (i.e., structural knowledge of the system).

Control Behavior: Research Histories of CDCTs

Studies of CDCTs have followed two separate research histories that have produced contrary proposals for the type of processes involved in the acquisition and application of knowledge in control tasks.

The first research history originated in work by Broadbent (e.g., Broadbent, 1977; Broadbent & Ashton, 1978) and is an early example exploring dissociations between implicit and explicit learning systems. Good control performance is unaccompanied by declarative knowledge of the task environment (Berry & Broadbent, 1984; Dienes & Berry, 1997; Dienes & Fahey, 1995), and unaccompanied by self-insight into the processes used to control it (Berry & Broadbent, 1984). In addition, practice can lead to further improvements in the controllability of the system, but not to similar increases in declarative knowledge (Berry & Broadbent, 1988). Implicit learning theorists (Broadbent, Fitzgerald, & Broadbent, 1986; Dienes & Berry, 1997; Dienes & Fahey, 1995, 1998) also claim that dissociations are found because the two systems operate over different types of knowledge. Typically, in CDCTs, the input-output relations are non-salient and therefore difficult to acquire. Rather than abstracting the underlying structure of the system, exemplars of associations between specific actions and their consequences are formed. Hence, dissociations between direct and indirect measures occur because similarity is used to match new goal states of a system with previously stored exemplars; this often leads to relevant procedural knowledge facilitating good control performance, but not to structural knowledge of the system (Broadbent, et al., 1986; Dienes & Fahey, 1995, 1998). Consistent with this, studies contrasting observation-based and procedural-based learning show that, on indirect and direct measures, performance was poorer for observation-based learners (Berry, 1991; Lee, 1995). The strong implications are that action has an advantage over observation, because CDCTs are proceduralized tasks in which learning is incidental and results from direct interactions with the system (Berry, 1991; Berry & Broadbent, 1988; Stanley, Mathews, Buss, & Kotler-Cope, 1989; Sun, Merrill, & Peterson, 2001).

The second research history concerns the effects of two related features of goal setting in CDCTs: goal-specificity (e.g., Burns & Vollmeyer, 2002; Owen & Sweller, 1985; Sweller & Levine, 1982; Vollmeyer, Burns, & Holyoak, 1996), and goal-difficulty (e.g., Kanfer & Ackerman, 1989; Kanfer, Ackerman, Murtha, & Nelson, 1994). In studies of the goal specificity effect, participants are either instructed to first learn about a complex control system by controlling it to a criterion (SG: Specific Goal), or are simply told to explore the system (NSG: Non-Specific Goal). Compared to NSG learning, SG learning interferes with the exploration and discovery of these strategies because it constrains attention to states of the system that are relevant only for reaching a particular goal.

The effect of varying the specificity and difficulty of reaching a goal on control performance was first examined by Sweller (Sweller & Levine, 1982; Sweller, Mawer, & Ward, 1983), and has since been replicated (e.g., Berry & Broadbent, 1984; Burns & Vollmeyer, 2002; Geddes & Stevenson, 1997; Miller, Lehman, & Koedinger, 1999; Owen & Sweller, 1985; Trumpower, Goldsmith, & Guynn, 2004; Vollmeyer et al., 1996). Burns and Vollmeyer's (2002) recent account of the phenomena develops on Dual-Space theory (Klahr & Dunbar, 1988; Simon & Lea, 1974). Its main tenet is that a skill is acquired by applying the principles underlying 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 evaluative processes and prior knowledge, whereas search in the experimental space is guided by the current goal. SGs promote exclusive search of the experimental space, whereas NSG involves moving through both spaces via hypothesis testing (e.g., Burns & Vollmeyer, 2001; Geddes & Stevenson, 1997; Renkl, 1997; Trumpower, Goldsmith, & Guynn, 2004; Vollmeyer, Burns, & Holyoak, 1996).

Consistent with this, studies of multi-cue learning and CDCTs show that problem solvers react evaluatively to their problem solving behavior on difficult and simple goal tasks (Bandura, 1986; Cervone et al., 1991). In complex, difficult set goals, they fail to develop a coherent understanding of the system because they do not focus on developing a single successful strategy, but instead find solutions through trial and error (e.g., Cervone, et al., 1991; Earley et al., 1989; Earley et al., 1990), which in turn lowers judgments of perceived efficacy. In simple tasks, the problem solver already possesses the relevant rudimentary knowledge needed to perform the task, and so goals serve to enhance the solver's existing ability, and increase motivation and perceived self-efficacy (Locke & Latham, 1990).

To summarize, the claims made by implicit learning theories of CDCTs conflict with studies concerned with the effects of goal specificity on skill learning. The former approach suggests that the learning processes underlying 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 (e.g., hypothesis testing and evaluative processes) influence the success of control behaviors (e.g., Burns & Vollmeyer, 2002; Sweller, 1988).

Present Study

The present study examines the relationship between monitoring and control, by asking what effect monitoring has on the transferability of control behaviors in CDCTs. Social Cognitive theory claims that monitoring and control impact on all aspects of behavior. There is strong evidence that, in the acquisition of complex skills, people have self-insight into those of their behaviors on which regulatory mechanisms operate. Similarly, Dual-Space theory proposes that monitoring has a role in skill learning, because evaluative processes are invoked through hypothesis testing behaviors. Both theories claim that expertise is goal-directed, and empirical evidence shows the effects of the specificity of goals on skill acquisition. Moreover, both theories assume the involvement of monitoring behaviors in tracking and modulating control performance. Therefore these behaviors should also affect the transferability of control in CDCTs.

In contrast, implicit learning theorists have argued that skilled behaviors, like controlling a complex system, are acquired incidentally through procedural processes, and that declarative processes (i.e., observational learning) that involve examination of one's knowledge impair control performance. Moreover, explicit declarative knowledge of the system cannot be transformed into the implicit procedural knowledge necessary to control it, because the learning mechanisms that operate over these different forms of knowledge are dissociated. This position strongly asserts procedural processes' independence of conscious declarative processes. Therefore monitoring should have a detrimental affect on the transferability of control behaviors in CDCTs.

Given the tension between theoretical accounts of monitoring and control behavior in CDCTs, and the potential effect of monitoring on the transferability of control behavior, the present study examines these questions: (1) Does control performance improve if monitoring is based on one's prior self-generated behavior, rather than the behavior of another individual? (2) Can people discriminate between their own self-generated behavior and that of another individual? (3) Is control performance improved if monitoring takes place online rather than indirectly, through experience of self-generated or other-generated behaviors?

General Method

In the following four experiments, participants performed two problem solving tasks, each consisting of a learning phase and a control test phase. All participants solved the first problem in the same way, and in each experiment all the critical manipulations concerned the contents of the learning phase in the second problem. In 'self'-labeled conditions, participants in the second problem were again exposed to their own learning phase from the first problem. In 'other'-labeled conditions, participants were yoked to a participant in the corresponding 'self' condition, and in the second problem were exposed to that individual's learning phase. In addition, the presentation format of the learning phase in the second problem was varied: i.e., it was either action-based (Experiments 1, 2, 3) or observation-based (Experiments 1, 4), and the cover story was manipulated so that either the second problem was different to the first (Experiment 1, 4) or identical (Experiment 2, 3). A further manipulation concerned the instructions presented prior to the presentation of the second problem (Experiment 2, 3).

Experiment 1

Experiment 1 included four conditions: in each, participants solved two CDCTs. All participants solved the first CDCT problem in the same way, by generating their own learning experiences in the learning phase. However, in the second problem, half the participants re-experienced their original learning phase from the first problem, either through observation-based (Observe-self) or action-based (Act-on-self) learning. The remainder experienced a different learning phase to that they had generated either through observation-based (Observe-other) or action-based (Act-on-other) learning.

Implicit learning theorists (e.g., Berry, 1991; Berry & Broadbent, 1988; Lee, 1995; Sun et al., 2001) propose that procedural processes are necessary in the acquisition and transfer of knowledge in CDCTs. Therefore, in Experiment 1, transfer of control performance should be facilitated if the learning phase of the first and second problem is procedural-based (Act-on-self, Act-on-other), and performance should increase across problems. Additionally, in conditions in which the learning formats of the first and second problem are different (Observe-self, Observe-other), transfer of control performance should be impaired. Decrements in control performance should be found in the second observation-based learning problem, because declarative knowledge is brought to bear during observation-based learning and invokes intentional examination of the learning process. This interferes with implicit procedural processes (Berry, 1991; Berry & Broadbent, 1987).

If, however, people do have insight into the processes invoked during learning in CDCT, self-regulatory mechanisms, such as monitoring, will mediate control behaviors. Therefore, transfer of control behaviors from the first to the second problem should be facilitated, regardless of the presentation format of the learning phase in the second problem. Instead, if self-regulatory processes like monitoring are involved, then, during the learning phase, people will be sensitive to the kind of information presented to them (i.e., the source of the learning phase in the second problem), not its presentation format (observation-based, action-based). In this case, participants will demonstrate knowledge of the difference in the source of the learning phase of the second problem.

Method

Seventy-two students from University College London volunteered to take part in the experiment and were paid £6 for their participation. They were randomly allocated to one of four conditions (observe-self, act-on-self, observe-other, act-on-other), with eighteen in each. Participants were tested individually.

Design and Materials

Experiment 1 was a mixed design that included two between subject variables comparing re-exposure to self-generated learning instances vs. exposure to other-generated learning instances (i.e., Self vs Other), and the effects of learning format on transfer of control performance (Observation, Action). There were also two within subject variables, one examining the transfer of control performance between two CDCTs in two tests of transfer (Control Test 1, Control Test 2), and one examining the transfer of causal structural knowledge between two CDCTs in four tests (Structure Test 1, Structure Test 2, Structure Test 3, Structure Test 4). Each participant was required to solve two CDCT problems, and the order of presentation of the problems was randomized for each participant. For each problem, there was a learning phase of 12 trials, in which participants were given an opportunity to learn about the control system. After the 6^{th} and 12^{th} trials, participants were presented with a direct test (Structure Test 1, Structure Test 2) of their knowledge of the underlying structure of the system. Participants were then presented with two indirect tests (Control Test 1, Control Test 2), which measured their ability to control the system to specific criteria. Control Test 1 consisted of 6 trials, followed by Structure Test 3. Control Test 2 consisted of 6 trials, followed by Structure Test 4.

The critical manipulation in Experiment 1 was the contents of the learning phase that participants received in the second problem. All participants generated their own learning experiences in the learning phase of the first problem but, in the second, half the participants re-experienced their original learning phase (observe-self, act-on-self), and the other half experienced a different learning phase to that they had generated (observe-other, act-on-other).

CDCTs

The design and underlying structure of the two CDCTs used (Water-Tank control system, Ghost Hunting control system) were based on the Water Tank system (see Figure 1). The only differences between the two problems were the visual layout of each system, as presented on screen, and the cover story (See appendix for cover stories and details). In the Water-Tank control system, each participant was told that they were a worker from the water purification plant and it was their job to inspect the new system being used. The system worked by varying the different levels of salt, carbon, and lime (inputs), which in turn changed oxygenation, temperature, and chlorine concentrations (outputs), which are indicators of how pure the water is. In the Ghost Hunting control system, participants were told that they were newly recruited ghost hunters, and had returned from a field experiment. It was their job to learn the relationship between three pieces of equipment that had been used in the field, and the phenomena that each machine detected. The three machines (i.e., GGH Meter, Anemometer, Trifield Meter) represented the three inputs, and the three phenomena that were detected (Electro Magnetic Waves, Radio Waves, Air Pressure) represented the three outputs. The control element of the task was to modify the levels of the readouts of the phenomena, by manipulating the dials on each machine.

For each problem at the start of the learning phase, and at the beginning of each control test, the input values were set to 0, and the output levels were set as follows: Output 1 (Water Tank = Oxygenation, Ghost Hunting = Radio Waves) = 100; Output 2 (Water Tank = Chlorine Concentration, Ghost Hunting = Electro

Magnetic Waves) = 500; Output 3 (Water Tank = Temperature, Ghost Hunting = Air Pressure) = 1000.

Procedure

Participants were told that they would be participating in a problem solving task. On completing the first problem, participants were told that they had to solve a second problem, but at no stage were they informed that the two problems were structurally the same.

First problem: Learning phase. In the first problem, in the learning phase, participants were presented with a computer display (see Figure 2) with three input variables and three output variables. Figure 1 presents the underlying structure connecting the inputs and outputs.

The learning phase comprised 12 trials, divided into two blocks of 6. 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 designed to index knowledge of the causal structure of the control system. A diagram of the system was shown on screen, and participants were asked simply to indicate which input was connected to which output. The direction of the input-output connection was implicit in the way that participants interacted with the CDCT, and was also indicated in the instructions to the task, so it was not necessary to examine the directionality of the input-output relations, only which connections existed. After this, participants began the next set of 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.

Second problem: Observation-based learning phase. In the second problem, the learning phase was observation-based for half the participants. With the exception that participants themselves could not manipulate input values during the learning phase, but were instead replayed their learning phase from the first problem, or yoked to the learning phase of a participant in the Observe-self 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. Participants then 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. 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 Trial 12, participants were presented with a Structure Test.

Second problem: Action-based learning phase. For the remaining participants, the learning phase of the second problem was action-based. In the Act-on-self condition, participants' learning phase from the first problem was logged: This included a trial history of the inputs that had been changed, and the values that they were changed by. At the start of the learning phase, participants were instructed to change the system by following the trial history sheet. This indicated which inputs to change and the corresponding values they should be changed by. The Act-on-self condition were presented with a trial history of their own learning phase from the first problem, and the Act-on-other condition were presented with the trial history of a participant from the Act-on-self condition.

Control Test Phase: Control Test 1. After the learning phase and Structure Tests 1 and 2, participants' ability to control the system was tested. In this phase, all participants were required to change the input values to achieve and maintain set output values. In the first and second problems, the criterion values they had to achieve were the same, and only the labels of the outputs were different: Output 1 (Water Tank = Oxygenation, Ghost Hunt = Radio Waves) = 50; Output 2 (Water Tank = Chlorine Concentration, Ghost Hunt = Electro Magnetic Waves) = 700; Output 3 (Water Tank = Temperature, Ghost Hunt = Air Pressure) = 900, for the course of 6 trials. On completion of this phase, participants were presented with Structure Test 3.

Control Test 2. In this phase, all participants were required to change the input values, to achieve and maintain a different set of output values to Control Test 1. In the first and second problems, the criterion values they had to achieve were the same, and only the labels of the outputs were different: Output 1 (Water Tank = Oxygenation, Ghost Hunt = Radio Waves) = 250; Output 2 (Water Tank = Chlorine Concentration, Ghost Hunt = Electro Magnetic Waves) = 350; Output 3 (Water Tank = Temperature, Ghost Hunt = Air Pressure) = 1100, for the course of 6 trials. On completion of this phase, participants were presented with Structure Test 4.

Post-test question. After completing the second problem, participants were informed that the experiment they had participated in consisted of one key manipulation in the learning phase of the second problem, which was described to them (i.e., exposure to self-generated learning, or exposure to other-generated learning). They were then asked which of the two manipulations they had experienced. This question served as an index of self-insight, and examined whether participants could accurately detect whether the learning phase that they had experienced in the first problem was the same as or different to their own learning phase in the second problem.

Scoring

Structure scores. The method used to score performance on Structure Tests 1-4 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 can be made). The maximum value for each structure score was 1. This scoring scheme was applied to score performance on all structure tests in Experiments 1-4.

Control Tests 1 and 2. The procedure used in Experiments 1-4 was based on Burns and Vollmeyer's scoring system. Control performance was measured as error scores in Control Tests 1 and 2. Error scores were based on calculating the difference between each target's output value (i.e., the criterion according to the control test) and the actual output value produced by the participant for each trial of the transfer test. A log transformation (base 10) was applied to the error scores of each individual participant for each trial, to minimize the skewedness of the distribution of scores. All analyses of error scores for Control Test 1 were based on participants' mean error, averaged over all 6 trials, across all three output variables. The error scores for Control Test 2 were calculated in the same way. Success in control performance on transfer tasks is indexed by the difference between the achieved and target output values, and therefore lower error scores indicate better performance.

Results

This section begins with an analysis of performance on direct and indirect measures between conditions on the first problem. Analyses were then conducted on participants' ability to control the CDCTs in Control Tests 1 and 2 (indirect measure of knowledge), and their knowledge of the underlying structure of the system in Structure Tests 1, 2, 3, and 4 (direct measure of knowledge). Correlation analyses were conducted to examine the potential association between indirect and direct measures of performance. Finally, responses to the post-test question were analyzed. In all analyses reported in this article, a significance criterion of $\alpha = .05$ was used. The results of non-significant findings are not reported.

Performance on direct and indirect measures in the 1st problem. The control performance of both conditions in the first problem was initially compared, to rule out any possibility of initial group differences influencing any later main effects detected as a result of the critical manipulations. A 2x4 ANOVA with control test (Control Test 1, Control Test 2) as a within subject variable, and condition (Observe-self, Observe-other, Act-on-self, Act-on-other) as a between subject variable, was conducted on mean error scores. The analysis revealed a significant main effect of control test, $F(1, 68) = 14.82$, MSE = 0.37, $p < 0.0005$, $\eta^2 = 0.18$; no other main effect or interactions were significant. The same analysis was conducted on scores from the structure tests. Each participant's scores across the four Structure Tests 1, 2, 3, and 4 were averaged and a Univariate analysis was carried out comparing performance across the four conditions. No significant difference between conditions was found based on structure test scores. Overall, there was no significant difference between conditions on indirect and direct measures of performance in the first problem. However, the overall performance of the four conditions differed significantly between control tests, suggesting that the control tests differed in difficulty. This is consistent with the findings reported in Burn and Vollmeyer's study, on which the control test used in the present study was based.

Comparison of control test scores in the 1st and 2nd Problem. Figure 3 presents the mean error score of all four conditions for each control test for each problem. Figure 3 suggests that, for the Observe-self and Act-on-self conditions, mean error scores increased in both control tests in the second problem. The reverse trend is indicated for the Observe-other and Act-on-other conditions.

To analyze the pattern of behavior indicated in Figure 3, a 2x2x2x2 ANOVA was carried out using control test (Control Test 1, Control Test 2) and problem (1st Problem, 2nd Problem) as within subject variables, and condition (self, other) and learning format of the second problem (observation, action) as the between subject variables. The analysis showed a significant main effect of control test, $F(1, 68) =$

29.52, MSE = 0.82, $p < 0.005$, $\eta^2 = 0.30$. There was also a significant main effect of condition, $F(1, 68) = 11.59$, MSE = 0.87, $p < 0.001$, $\eta^2 = 0.15$, and a significant Condition x Problem interaction, $F(1, 68) = 53.27$, MSE = 2.46, $p < 0.0005$, $\eta^2 = 0.44$. No other main effects or interactions were significant.

Given that there was no Condition x Problem x Control test interaction, the scores were collapsed across control tests. The significant decrease in performance between the first and second problem of the Observe-self and Act-on-self conditions (Figure 3) was confirmed by a post hoc comparison of error scores: t(35) = -4.52, $p < 0.001$, $d = -1.53$ and t(35) = -6.25, $p < 0.0005$, $d = -2.11$, respectively. The significant increase in performance between the first and second problem of the Observe-other and Act-on-other conditions was also confirmed by a planned comparison of error scores: t(35) =3.75, $p < 0.001$, $d = 1.27$ and t(35) = 3.65, $p < 0.001$, $d = 1.23$, respectively. Thus, the evidence suggests that the difference in the patterns of transfer of control performance was the result of the content of the learning phase in the second problem, and not of its presentation format.

Comparison of structure test scores in 1st and 2nd Problem. For each participant, the scores from Structure Tests 1, 2, 3, and 4 were averaged across the first problem, and again for the second problem. Figure 4 presents the mean structure test scores of all four conditions for each problem. Figure 4 suggests that, for the Observe-self and Act-on-self conditions, structure scores decreased in all structure tests in second problem. The reverse trend is indicated for the Observe-other and Act-on-other conditions.

To analyze the pattern of behavior indicated in Figure 4, a 2x2x2 ANOVA was carried out on averaged structure test scores, using problem (1st Problem, 2nd Problem) as a within subject variable, and condition (self, other) and format (observation, action) as the between subject variables. There was a significant main effect of condition, $F(1, 68) = 4.37$, MSE = 32.23, $p < 0.05$, $\eta^2 = 0.06$, and a significant Condition x Problem interaction, $F(1, 68) = 35.95$, MSE = 129.34, $p < 0.001$, $\eta^2 = 0.35$. No other main effects and interactions were significant.

The significant decrease in performance between the first and second problem of the Observe-self and Act-on-self conditions (Figure 4) was confirmed by a post-test carried out on structure scores: $t(17) = 5.16$, $p < 0.005$, $d = 2.50$, and $t(17) = 2.32$, $p < 0.05$, $d = 1.13$, respectively. In addition, the significant increase in performance, in the second problem, of the Observe-other and Act-on-other conditions, was confirmed: $t(17) = -3.38$, $p < 0.005$, $d = -1.64$ and $t(17) = -2.40$, $p < 0.05$, $d = -0.57$, respectively. Thus, the evidence suggests a negative transfer of declarative knowledge in the Observe-self and Act-on-self conditions, and a positive transfer in the Observe-other and Act-on-other conditions.

Correlation between indirect and direct measures. Berry and Broadbent (1987, 1988) argued that dissociation between performance on direct and indirect measures suggests that the knowledge gained in control tasks is procedural. To examine this, a correlation analysis was carried out on averaged control error scores (averaged across Control Tests 1 and 2), and mean Structure Test (averaged across Structure Tests 1, 2, 3, 4) scores from the first and second problems. The analysis revealed a significant negative relationship between Structure Test scores and Control Test error scores in the first problem, $r(72) = -0.29$, $p < 0.05$, and between Structure Test scores and Control Test error scores in the second problem, $r(72) = -0.38$, $p < 0.001$. The findings from these sets of analyses strongly indicate that, for both types of learning phases (observation-based, procedural-based), there is a relationship between indirect and direct measures of knowledge.

Post-test question. Eighty-three percent of participants in the Observe-self condition and 67% in the Act-on-self condition were able to accurately report which of the two conditions they were in. Seventy-eight percent of participants in the Observe-other condition and 78% in the Act-on-other condition answered correctly. Pearson's chi-squared analysis revealed that there was no significant difference in correct and incorrect response by condition.

Discussion

The evidence from Experiment 1 can be summarized, as follows: First, there was successful transfer of control performance across problems regardless of the similarity or difference in learning format of the problems. Second, performance on direct and indirect measures was associated, and in both, the pattern of performance across problems was consistent. Third, problem solvers showed insight into the processes used during the acquisition of knowledge in CDCTs, indexed by responses to the post-test question. All conditions were able to correctly discriminate the source of the learning phase in the second problem, and there was no difference in accuracy of responses between the four conditions. Fourth, across problems there was evidence of positive transfer of structural knowledge and control performance in Observe-other and Act-on-other conditions, and negative transfer in Observe-self and Act-on-self conditions.

Taken together, the evidence indicates that control performance is not exclusively dependent on proceduralized learning in CDCTs. Moreover, participants' declarative knowledge of the system corresponded to the control behavior, and they showed accurate self-insight into behaviors invoked during the task, which suggests that declarative and procedural knowledge is accessible to consciousness. Although inconsistent with implicit theorists' claims, the findings suggest that monitoring and control behaviors are associated, and that differences in transfer of control performance from the first and second problem are dependent on the source, rather than the format, of the learning phase.

The negative transfer effect found in direct and indirect measures of performance in the Observe-self and Act-on-self conditions was an unexpected result. One reason for this result may have been that participants accurately detected the origin of the second learning phase: that is, the strategies used for guiding which inputs were changed and by how much. It has previously been shown (Bandura, 1986, 1989; Burns & Vollmeyer, 2002; Cervone et al., 1991; Ericsson, 1996) that problem solvers react evaluatively to their performances through the goals they are set. In Experiment 1, participants may have judged the success of the learning phase of the first problem by retrospectively evaluating it from their control performance in the control test phase. Both *self* conditions may have judged the learning phase to be less effective and consequently performed poorly in the second problem, whereas both *other* conditions remained neutral in their evaluations because the learning phase was not their own.

Although rare, demonstrations of negative transfer in studies of problem solving and rule learning do exist (e.g., Chen & Daehler, 1989; Lee & Vakoch, 1996; Luchins, 1942; Novick, 1988; Woltz, Bell, Kyllonen, & Gardner, 1996; Woltz, Gardner, & Bell, 2000). In many of these studies, negative transfer of skilled learning is taken as an indication of overlearnt behavior. The studies show that well rehearsed memories, for sequences of operations, make it difficult for people to prevent transfer to contexts similar to the conditions in which they were acquired. For example, in studies of rule learning (Luchins, 1942; Woltz et al., 1996; Woltz et al., 2000), participants trained in the discovery and application of rules to specific tasks tend to over-generalize to novel instances in which they are not applicable. Additionally, as expertise in the use of rules increases, or as exposure to the learning environment in which they were acquired increases, so too does the inability to discover new strategies that lead to successful solutions (Lee & Vakoch, 1996; Woltz, Bell, Kyllonen, & Gardner, 1996; Woltz, Gardner, & Bell, 2000). In these examples negative transfer is an index of strong memory of prior learnt instances. However, this is not the case in the present study: Rather, negative transfer appears to be a result of re-exposure to self-generated learning instances. Before fully exploring the basis for the negative transfer effect found in Experiment 1, a further experiment was devised to investigate the reliability of both the negative and positive transfer effects that were found.

Experiment 2

Experiment 2 examined two objectives: the reliability of the negative transfer effect reported in Experiment 1 under practice rather than transfer conditions, and whether disguising the origin of the learning phase in the second problem would interfere with the negative and positive transfer effects found in Experiment 1. Experiment 2 included four conditions: Act-on-self, Act-on-other, disguised-self, disguised-other.

Unlike Experiment 1, Experiment 2 examined the development of skilled performance through practice rather than transfer of knowledge. To fully explore skill learning in CDCTs, participants were presented with perceptually and structurally identical CDCTs, and the learning phase of both problems was proceduralized, and followed the same procedure as the Act-on-self and Act-on-other conditions in Experiment 1. Thus the experience participants had gained during the first problem should have been easily applicable to the second. This change in procedure was based on the hypothesis that general practice effects should be more likely to emerge in the second problem, because participants would then be highly familiar with the problem. If, however, the negative and positive transfer effects found in Experiment 1 were robust, then, regardless of whether the problems were perceptually and structurally identical, in both indirect and direct measures of performance the Act-on-self condition should show decrements across problems, whereas the Act-on-other condition should show improvements across problems.

To complement this, a further manipulation disguised the origin of the learning phase in the second problem. Before presentation of the second problem, the disguised-self condition was told that the trial history sheet was generated from another participant in the same problem, but the trial history was actually of their own learning phase from the first problem. The disguised-other condition was told that the trial history sheet was of their own learning phase from the first problem, whereas it was generated by a participant from the disguised-self condition. This manipulation was intended to examine whether participants negatively evaluate their own learning phase, which then impairs later control performance; and, conversely, whether participants positively evaluate another participant's learning phase. It was hypothesized that, following this instructional manipulation, if monitoring influences control behavior then manipulating belief in the origin of the learning phase would also affect control behavior. If so, then performance in indirect and direct measures should increase in the disguised-self condition, because they now believed that the origin of the learning phase was another participant. Decreases in performance across problems in the disguised-other condition should be found, because they now believed that they were re-exposed to their own learning phase.

Method

Seventy-two students from University College London volunteered to take part in the experiment and were paid £6 for their participation. They were randomly allocated to one of four conditions (Act-on-self, Act-on-other, disguised-self, disguised-other), with eighteen in each. Participants were tested individually.

Design

Experiment 2 was a mixed design that included a between subject variable examining the effects of the second learning phase, or the actual origin of the learning phase that affected transfer of control performance, by comparing four conditions (Act-on-self, Act-on-other, disguised-self, disguised-other), and two within subject variables, one examining the transfer of control performance between two CDCTs in two tests of transfer (Control Test 1, Control Test 2), and one examining the transfer of causal structural knowledge between two CDCTs in four tests (Structure Test 1, Structure Test 2, Structure Test 3, Structure Test 4). In each condition, half the participants were presented with the Water-Tank system problem twice, and the remainder with the Ghost Hunting problem twice. With the exception of the instructional manipulation introduced prior to 2nd Problem, the design of Experiments 1 and 2 was identical.

Procedure

The procedure used in Experiment 2 differed from Experiment 1 in two respects. Participants in all four conditions experienced the identical problem twice. Half the participants were required to solve the Water-Tank control problem twice, and the others solved the Ghost Hunting problem twice. In all four conditions, the critical manipulation occurred in the second problem. The Act-on-self and Act-on-other conditions followed the same procedure as in Experiment 1. The Act-on-self was presented with a trial history of their own learning phase from the first problem, whereas the Act-on-other condition was yoked to a participant from the Act-on-self condition. However, for the disguised-self and disguised-other, an instructional manipulation was introduced prior to presentation of the second problem. The disguised-self condition was informed that they would be presented with a record sheet of the learning phase of a participant who had taken part in the same problem. In actual fact, the record sheet was of their own learning phase from the first problem. Before presentation of the second problem, the disguised-other condition was informed that they would be presented with a record sheet based on the learning phase they had generated in the first problem. However, the record sheet belonged to a participant from the disguised-self condition. To render the instructional manipulation maximally effective, the disguised-other condition was instructed that the record sheet was based on, rather than identical to, their learning phase in the first problem, because it has been shown that it is easier to manipulate belief concerning self-generated behavior, because people assume similarity of behavior across individuals. However, it is difficult to manipulate belief when the learning experiences are not self-generated, because people are reliably able to detect self-generated behaviors (e.g., Knoblich & Flach, 2001; Knoblich & Prinz, 2001; Loula et al., 2005).

Results

Performance on direct and indirect measures in 1st Problem. The same preliminary analyses were conducted as in Experiment 1, to exclude any possibility of initial group differences influencing any later main effects detected as a result of the critical manipulations. The analyses carried out on control test error scores and mean structure test scores (averaged over Structure Tests 1, 2, 3, and 4) revealed, overall, no significant difference between conditions on indirect and direct measures of performance in the first problem.

Comparison of control test scores in 1st and 2nd Problem. Figure 5 presents the mean error score of both conditions for each control test for each problem. Figure 5 suggests that, for the Act-on-self and disguised-other conditions, mean error scores increased in both control tests in the second problem. This pattern was reversed for the Act-on-other and disguised-self conditions.

A 2x2x2x2 ANOVA was carried out using control test (Control Test 1, Control Test 2) and Problem (1st Problem, 2nd Problem) as within subject variables, and Source of second learning phase (Self-generated, Other-generated) and Belief in the origin of the second learning phase (undisclosed, disguised) as the between subject variables. The analysis showed a significant main effect of Control Test, $F(1, 68) = 11.47$, MSE = 0.30, $p < 0.001$, $\eta^2 = 0.14$, and Belief, $F(1, 68) = 6.09$, MSE = 0.20, $p < 0.05$, $\eta^2 = 0.08$. The following interactions were also significant: Source x Problem, $F(1, 68) = 4.86$, MSE = 0.14, $p < 0.05$, $\eta^2 = 0.07$; Belief x Problem, $F(1, 68) = 6.84$, MSE = 0.20, $p < 0.01$, $\eta^2 = 0.09$; and Belief x Source, $F(1, 68) = 9.47$, MSE = 0.31, $p < 0.005$, $\eta^2 = 0.12$. There was also a three-way interaction between Source x Belief x Problem, $F(1, 68) = 23.32$, MSE = 0.68, $p < 0.005$, $\eta^2 = 0.26$.

To locate the source of the interactions, post-test analyses were conducted of the trend in performance between the first and second problems for the four conditions. The analyses confirmed the trends in Figure 5. The decrease in performance across problems in the Act-on-self condition was confirmed: $t(35) = -5.72$, $p < 0.0005$, $d = -1.93$. The increase in performance across problems of the Act-on-other condition, and disguised-self, was also confirmed: $t(35) = 2.49$, $p < 0.05$, $d = 0.84$, and $t(35) = 3.32$, $p < 0.005$, $d = 1.12$, respectively. Thus, the pattern of performance indicated in Figure 6 was confirmed, suggesting that across problems there was evidence of decrements in control performance in the Act-on-self condition, and practice effects in the Act-on-other and disguised-self conditions.

Comparison of structure test scores in 1st and 2nd Problem. Figure 6 presents the mean score averaged over the four structure tests for each problem. Figure 6 indicates that, for Act-on-self and disguised-other conditions, structure scores decreased across problems, whereas performance increased across problems for the Act-on-other and disguised-self conditions.

To analyze this, a 2x2x2 ANOVA was carried out on structure scores averaged across the four tests of each problem, using problem (1st Problem, 2nd Problem) as within subject variables, and Source of second learning phase (Self-generated, Other-generated) and Belief in the origin of the second learning phase (undisclosed, disguised) as the between subject variables. No main effects were significant. The following interactions were also significant: Source x Belief, $F(1, 68) = 5.05$, MSE = 38.23, $p < 0.05$, $\eta^2 = 0.07$, and Belief x Problem x Source, $F(1, 68) = 26.04$, MSE = 97.62, $p < 0.0005$, $\eta^2 = 0.27$. Post hoc comparisons revealed that the decrease in performance between the first and second problem of the Act-on-self condition was significant, $t(17) = 2.37$, $p < 0.05$, $d = 1.15$. The significant increase in performance across problems of the Act-on-other condition and disguised-self was also confirmed: $t(17) = -3.08$, $p < 0.01$, $d = -0.73$, and $t(17) = -4.57$, $p < 0.0005$, $d = -2.22$, respectively. Thus, the evidence suggests negative transfer of causal knowledge of the system in the Act-on-self condition, and positive transfer of structure knowledge in the Act-on-other and disguised-self conditions.

Correlation between indirect and direct measures. A correlation analysis was carried out on averaged control error scores (averaged across Control Tests 1 and 2) and mean structure test scores from the first and second problems. The analysis revealed a significant negative relationship between Structure Test score and Control Test error scores in the second problem, $r(72) = -0.45$, $p < 0.001$. The findings from these sets of analyses strongly indicate a relationship between indirect and direct measures of knowledge in the second problem.

Post-test question. Sixty-seven percent of participants in the Act-on-self and 47% in the disguised-self conditions were able to accurately report which condition they thought they were in. Seventy-eight percent of participants in the Act-on-other condition and 61% in the disguised-other answered correctly. Pearson's chi-squared analysis revealed no significant difference in the accuracy of responses between conditions.

Discussion

The first objective of Experiment 2 was to examine the reliability of the negative transfer effect found in *Self* conditions, and the positive transfer effects found in *Other* conditions. Experiment 2 presented participants with two perceptually and structurally identical problems and replicated the effects reported in Experiment 1, in the Act-on-self and Act-on-other conditions. Experiment 2 also included two further conditions in which the origin of the learning phase in the second problem was disguised. The findings confirmed the prediction that participants have a biased perception of the learning phase they generated. The evidence showed that, although participants in the disguised-self condition were re-exposed to their own learning phase, the belief that it belonged to another participant generated patterns of performance that were consistent with the Act-on-other condition. Control performance and accuracy of structural knowledge increased in the second problem.

Participants in the disguised-other condition were led to believe that the learning phase of the second problem was their own, and may have been less convinced by this manipulation. This may explain why control performance and accuracy of structural knowledge in the second problem were equivalent in both problems. To further explore whether problem solvers react evaluatively to the learning phase they are presented in the second problem, as a result of self-regulatory mechanisms (Bandura, 1986; Cervone et al., 1991), a further experiment was designed. This examined whether biases can be introduced that lead to negative and positive evaluations of self-generated learning instances, which in turn affect later control behavior.

Experiment 3

The evidence from Experiments 1 and 2 indicates that the evaluative processes employed during learning influence the transferability of control performance. Consistent with Social-Cognitive theory, we suggest that participants judged the success of the learning phase of the first problem by retrospectively evaluating it from their control performance in the control test phase. Thus, problem solvers' estimation of their problem solving ability informs the evaluative process and in turn affects their later control performance.

To examine this, Experiment 3 included two conditions. Both re-exposed participants to their own previous learning experiences. In the Act-on-self-high and Act-on-self-low conditions, participants solved the first problem in the same way, and, before the solution of the second, were given advance knowledge that the learning phase was their own. They were also presented with the average performance of other participants that had solved the first problem. Participants in the Act-on-self-high condition were told that average performance was extremely good, and that they were within +/- 20[2] of the target values on both control tests. Participants in the Act-on-self-low condition were told that the average performance was extremely bad, and that they were within +/- 200 of the target values on both control tests.

If the transferability of control performance is strongly influenced by the evaluative processes employed during learning, then the Act-on-self-high would negatively evaluate their control performance from the first problem. This would affect their estimation of the effectiveness of the learning phase when presented again in the second problem, and produce negative transfer of control performance. Conversely, the Act-on-self-low condition should positively evaluate their performance in the first problem, and the effectiveness of the learning phase presented again in the second problem, thus producing positive transfer of control performance.

Alternatively, the differences in the instructions presented to the two conditions may result in differences in attention to the learning phase in the second problem, which might lead to differences in the transferability of control performance between conditions. Given that all participants were told that they would be re-exposed to their learning phase from the first problem again in the second, those in the Act-on-self-high condition may simply fail to attend to the learning phase again, resulting in poorer performance in the second problem. To prevent this, before the presentation of the learning phase in the second problem, all participants were informed that they would be given a test. Although participants were not explicitly informed of the memory test that was presented directly after the learning phase, knowledge of some impending test was designed to motivate them to pay full attention to the learning phase, particularly since they did not know what aspects of the phase were relevant for the test.

Method

Thirty-six students from University College London volunteered to take part in the experiment and were paid £6 for their participation. They were randomly allocated to one of two conditions (Act-on-self-high, Act-on-self-low), with eighteen in each. Participants were tested individually.

Design

Experiment 3 was a mixed design that included a between subject variable examining the effects of the diversity of learning experiences, by comparing two conditions (Act-on-self-high, Act-on-self-low), and two within subject variables, one examining the transfer of control performance between two CDCTs in two tests of transfer (Control Test 1, Control Test 2), and one examining the transfer of causal structural knowledge between two CDCTs in four tests (Structure Test 1, Structure Test 2, Structure Test 3, Structure Test 4). In each condition, half the participants were presented with the Water-Tank system problem twice, and the remainder was presented with the Ghost Hunting problem twice. Both conditions were presented with their self-generated learning experiences in the first problem. The critical difference between the conditions was the information they had been presented with before the presentation of the second problem.

Procedure

Both conditions performed the first problem following the same procedure as in Experiments 1 and 2. Prior to the presentation of the second problem, the experimenter informed participants in both conditions that they would be presented with a record sheet that included the changes to inputs they had made in the learning phase of the first problem. In addition, the Act-on-self-high condition was told that, to motivate them, they would be told the average performance of other participants that had solved the first problem. It was explained that the average performance was extremely good, and that participants were within +/- 20 of each target value for each output on both control tests. Similarly, the Act-on-self-low condition was told that, to motivate them, they would be told the average performance of other participants that had solved the first problem. It was explained that the average performance was poor and that participants were within +/- 200 of each target value for each output, on both control tests. Along with the control and structure measures of knowledge used to examine the effects of presenting participants with extraneous information, a memory test was presented directly after the learning phase in the second problem. The memory test consisted of a blank trial history sheet, and participants took, on average, three minutes to freely recall the inputs they had changed, and the values they had changed them by, for each of the 12 trials of the learning phase. Participants were not told in advance that they would receive a memory test, but were told, before the presentation of the second problem, that they would be tested on their knowledge of the learning phase. The test was included to ensure that participants still paid attention to the learning phase presented in the second problem, despite explicitly knowing that it was their own.

Results

Performance on Control Tests 1 and 2. Preliminary analyses revealed no significant difference between conditions based on performance on control test and structure test scores in the first problem.

Comparison of control test scores in 1st and 2nd Problem. Figure 7 suggests that, for the Act-on-self-high condition, mean error scores increased in both control tests in the second problem, whereas the reverse is indicated for the Act-on-self-low condition.

A 2x2x2 ANOVA was carried out, using control test (Control Test 1, Control Test 2) and problem (1st Problem, 2nd Problem) as within subject variables, and condition (Act-on-self-high, Act-on-self-low) as the between subject variable. The analysis showed a significant main effect of control test, $F(1, 34) = 15.54$, MSE = 0.42, $p < 0.0005$, $\eta^2 = 0.34$. There was a significant main effect of condition, $F(1, 34) = 12.13$, MSE = 0.26, $p < 0.001$, $\eta^2 = 0.38$, and a significant Condition x Problem interaction, $F(1, 34) = 21.14$, MSE = 1.15, $p < 0.001$, $\eta^2 = 0.26$. No other main effects or interactions were significant.

The significant decrease in performance between problems of the Act-on-self-high condition (Figure 7) was confirmed by a post-test on error scores: $t(35) = -4.70$, $p < 0.0005$, $d = -1.59$. The significant increase in performance between problems of the Act-on-other-low condition was also confirmed, $t(35) = 2.54$, $p < 0.05$, $d = 0.86$. Thus, the evidence suggests that control performance improved in the second problem in the Act-on-self-high condition, and decreased across problems in the Act-on-other-low condition.

Comparison of structure test scores in 1st and 2nd Problem. Figure 8 indicates that, for the Act-on-self-high condition, structure scores decreased across problems, whereas performance increased across problems for the Act-on-self-low condition.

To analyze the pattern of behavior indicated in Figure 8, a 2x2 ANOVA was carried out, using problem (1st Problem, 2nd Problem) as the within subject variable, and condition (Act-on-self-high, Act-on-other-high) as the between subject variable. There was a significant Condition x Problem interaction, $F(1, 34) = 8.46$, MSE = 210.98, $p < 0.01$, $\eta^2 = 0.19$. The decrease in performance between the first and second problems of the Act-on-self-high condition (Figure 8) was confirmed, $t(71) = 3.21$, $p < 0.005$, $d = 0.73$. In addition, the increase in performance across problems of the

Act-on-other-low condition was confirmed, t(71) = -3.83, $p < 0.0005$, $d = -0.90$. Consistent with the pattern of control performance, accuracy of structural knowledge decreased in the second problem in the Act-on-self-high condition, and increased across problems in the Act-on-other-low condition.

Memory scores. Responses to the memory test presented at the end of the learning phase in 2nd Problem were scored in two ways: Input change—the proportion of correct input changes recalled; and Input value—the mean difference between actual input values and remembered input values. Input change was scored similarly to structure scores: i.e., (the number of correctly recalled input changes, and incorrect input changes avoided)/ N (the total number of input changes that can be made). The final score for each participant was converted to a percentage. Each participant's Input values were based on subtracting the recalled value for an input that was correctly recalled from the actual value of that input: This procedure was carried out for each trial of the learning phase, and the average difference between recalled and actual input values represented the input value score. The mean Input change scores of the Act-on-self-high condition and the Act-on-self-low condition were fairly low (38% and 36%, respectively). However, when input change scores were based on the first 6 trials, the Act-on-self-high condition and the Act-on-self-low condition scored above 50% (68% and 66%, respectively); but, when input change scores were based on the remaining 6 trials, the figures were lower (19% and 19%, respectively). The trend strongly suggests a primacy effect. The following analysis conducted on Input change scores is based on the full 12 trials of the learning phase.

Given that the data was not normally distributed, non-parametric tests were conducted on the data. However, to provide a more stringent analysis, parametric tests were also conducted. Both analyses revealed no difference between conditions based on their recall of their learning phase.

Correlation between indirect and direct measures. A correlation analysis was carried out on averaged control error scores (averaged across Control Tests 1 and 2) and mean structure test scores from the first and second problem. The analysis revealed a significant negative relationship between Structure Test scores and Control Test error scores in the first problem, $r(36) = -0.46$, $p < 0.005$, and between Structure Test scores and Control Test error scores in the second problem, $r(72) = -0.41$, $p < 0.05$. The findings of these analyses strongly indicate a relationship between indirect and direct measures of knowledge.

Discussion

The evidence from Experiment 3 confirmed the hypothesis that the evaluative processes employed during learning strongly influence the transferability of control performance. Both direct and indirect measures of knowledge revealed negative transfer of control performance and structural knowledge in the Act-on-self-high condition, and positive transfer of control performance and structural knowledge in the Act-on-self-low condition. Thus, transfer of control performance and structural knowledge was radically affected by a single change in the instructions (i.e., +/- 20, vs. +/- 200). These findings are also consistent with studies showing that erroneous feedback influences performance on tests of stamina (Litt, 1988), physical strength (Weinberg et al., 1981), strategic thinking (Bouffard-Bouchard, 1990), and complex decision making (Hogarth, Gibbs, McKenzie, & Marquis, 1991). In these examples, participants were presented with bogus normative standards that either suggested they had performed higher than the mean—which later elevated their performance—or lower than the normative standards—which then impaired performance.

The results of the memory test indicated that the instructional manipulation did not differentially affect recall of the learning phase presented in the second problem. More important, this result suggests that the accuracy of the overall recall of both conditions was independent of the pattern of control behavior of both conditions. This result speaks to an important issue concerning the negative transfer effect reported in self-experience conditions (i.e., Experiments 1 & 2). Lack of attention alone cannot explain a negative transfer effect. If participants failed to pay attention to the learning phase in the second problem, then their control performance should have been equivalent to that in the first problem. However, lack of motivation and attention can lead to forgetting, and this may account for the poor transfer effects reported in Experiments 1 and 2. If participants failed to pay attention to the learning phase in the first problem, then in the second they would have controlled the system based on their memory of what they had learnt in the first. However, if the learning achieved in the learning phase of the first problem was forgotten, then there would be poorer control performance in the second problem. The inclusion of the memory tests was designed to countermand the hypothesis that the difference in the transferability of control performance between conditions was the result of differences in motivation and attention, based on the instructions presented. The results of the memory test also rule out the hypothesis that forgetting the learning that takes place in the learning phase results in negative transfer of control performance and structural knowledge of the system.

Experiment 4

Thus far, the learning phase in each problem has been exploratory. Participants were simply told to interact with the system or observe the changes to it, and learn as much as they could about the relationship between inputs and outputs. The learning phase enabled participants to acquire knowledge of the general properties of the system, but not to gain practice in controlling the system to criterion, which is necessary for the control test phase. In addition, the number of available learning trials used throughout the study is the same as in the original CDCT by Burns and Vollmeyer (2002). This constrains the amount of learning that can occur. Therefore, Experiment 4 included two conditions (Observe-self, Observe-other) that were each exposed to two different CDCTs. In each, the number of learning trials was increased. The learning phase included an SG, in which participants were required to control the system to criterion—the same criterion of the first control test in each problem. By introducing a SG during the learning phase of the problem, participants were now able to assess the success of the various strategies they were developing to understand and control the system. Experiment 4 thus provided an opportunity to examine the generality of the positive and negative transfer effects found in Experiments 1-3, under conditions in which learning is SG based.

Experiment 4 explicitly examined the relationship between monitoring and control, by including judgments of self-efficacy after the learning phase, and again after the control test phase of each problem. If self-regulatory systems like self-efficacy influence control behavior, then judgments of self-efficacy taken before the control test phase should be highly associated with control performance. In addition, judgments of self-efficacy taken after the learning phase of the second problem should be higher in the Observe-other condition than in the Observe-self condition.

Finally, after the learning phase of the second problem, participants were asked to what extent they had based their understanding of the structure of the system on their structural knowledge from the first problem. This question was included to examine whether the way prior experience was used in the second problem discriminated between the conditions. If participants in the Observe-self conditions negatively evaluate the effectiveness of their learning phase from the first problem, then, in the second problem, they should report relying on it less than the Observe-other condition.

Method

Thirty-two students from University College London volunteered to take part in the experiment and were paid £10 for their participation. They were randomly allocated to one of two conditions (Observe-self, Observe-other), with sixteen in each. Participants were tested individually.

Design

Experiment 4 was a mixed design that included a between subject variable examining the effects of the diversity of learning experiences, by comparing two conditions (Observe-self, Observe-other), and two within subject variables, one examining the transfer of control performance between two CDCTs in two tests of transfer (Control Test 1, Control Test 2), and one examining the transfer of causal structural knowledge between two CDCTs in four tests (Structure Test 1, Structure Test 2, Structure Test 3, Structure Test 4). Each participant solved both the Water-Tank and Ghost Hunting CDCT problems, and the order of presentation of the two problems was randomized for each participant.

Each CDCT comprised 40 trials in the learning phase, divided into 4 blocks of 10 trials, after which a structure test was presented. During the learning phase of each CDCT, participants were also instructed to learn about the system whilst trying to control it by achieving and maintaining the following output criteria: Output 1 (Water Tank = Oxygenation, Ghost Hunt = Radio Waves) = 50; Output 2 (Water Tank =

Chlorine Concentration, Ghost Hunt = Electro Magnetic Waves) = 700; Output 3 (Water Tank = Temperature, Ghost Hunt = Air Pressure) = 900. The structure test and control test phases of each problem were identical to those used in all previous experiments.

Procedure

The procedure used was similar to that in Experiment 1. Participants were instructed that they would learn about, and then control, a system. During the learning phase, they were also told that they would have some practice at controlling the system to specific criteria. Participants were unaware that the criteria they had to follow for the duration of the learning phase were the same criteria used in the first control test for that problem. The number of trials included in the learning phase increased to 40 in total, divided into four blocks. A structure test was presented after each block. The control test phase included two control tests each 6 trials long, with a structure test after each.

The learning phase of the second problem was observation-based, comprising 40 trials, blocked into 4. Participants were instructed that they were watching the learning phase of a participant that had taken part in the study. Their job was to carefully observe the changes to the inputs and the outputs on each trial and assess how successful the observed participant was in achieving the criteria output values. The output criteria values were identical to those used in the first test presented in the control test phase. As with the observation-based learning conditions in Experiment 1, participants in the Observe-self condition were re-exposed to the input values they had generated in the learning phase of the first problem. The Observe-other participants were yoked to a participant from the Observe-self condition, and were thus exposed to different input values to those they had generated themselves in the learning phase of the first problem.

For each problem, after the learning phase was completed, participants were presented with the following question: Based on what you have learnt, how well do you think you can now control the system? Participants were told to imagine that they would be tested on their ability to control the system to the same criteria that they had just practiced on. They were told to estimate how close, on average, they could get to the criteria values, by choosing from the following range 1) +/- 25, 2), +/- 50, 3) +/- 75, 4) +/- 100 5) +/- 125, 6) +/- 150, 7) +/- 175, 8) +/- 200. After the control test phase, participants were presented with the following question: Now you have had a chance to control the system to different criteria, how well do you think you controlled the system in general? They were told to estimate their general ability by indicating, how close, on average, they had reached the criteria values by choosing from the following range 1) +/- 25, 2), +/- 50, 3) +/- 75, 4) +/- 100 5) +/- 125, 6) +/- 150, 7) +/- 175, 8) +/- 200. After the learning phase of the second problem, participants were also asked: To what extent did you base your current understanding of the relationship between the inputs … [salt, carbon, lime/GGH Meter, Anemometer, Trifield Meter] and the outputs … [oxygenation, chlorine concentration, temperature/ Electro Magnetic Waves, Radio Waves, Air Pressure] on your understanding of the relationship between the inputs and outputs from the previous problem? Responses were made using a 9-point scale ranging from negative (Not at all) to positive (Mostly).

Results

Learning phase. Because participants were presented with a SG in each of the four blocks of the learning phase, the following analysis examines the control error scores from these blocks in the first problem. A 4x2 ANOVA was conducted on control performance scores using block (Block 1, 2, 3, 4) as the within subject variable, and condition (Observe-self, Observe-other) as the between subject variable. The analyses revealed no significant effect of block or condition, suggesting that across blocks there was no difference in control performance during the learning phase in the first problem.

In addition, for each participant, a simple strategy analysis was conducted, based on the number of input variables changed for each trial recorded. These scores were averaged for each block and the mean for each condition is presented in Figure 9.

The trend suggested in Figure 9 indicates that participants changed fewer inputs in the second and third blocks of the learning phase than in the first and final blocks. To analyze this, a 4x2 ANOVA was conducted using block (Block 1, 2, 3, 4) as the within subject variable, and condition (Observe-self, Observe-other) as the between subject variable. The analysis revealed a significant main effect of block, $F(3, 90) = 4.86$, MSE = 2.17, $p < 0.005$, $\eta^2 = 0.14$. T-tests revealed significant differences in the number of inputs changed between Block 2 and Block 1, $t(31) = -2.20$, $p < 0.05$, $d = -0.79$, Block 2 and Block 4, $t(31) = -3.51$ $p < 0.001$, $d = -1.26$, and between Block 3 and Block 4, $t(31) = -2.44$, $p < 0.05$, $d = -0.88$.

The findings show no difference in control performance and number of inputs changed between conditions, and no difference in control performance across blocks, suggesting no general improvement in participants' ability to control the system. This may have been because participants were developing different strategies across the course of the learning phase, and varied, from block to block, the number of inputs that they changed. This is consistent with findings reported in Burns & Vollmeyer's (2002) study.

Performance on Control Tests 1 and 2. Preliminary analyses revealed no significant difference between conditions based on performance on Control Tests 1 and 2 in the control test phase, and structure test scores, in the first problem.

Comparison of control test scores in 1st and 2nd Problem. Figure 10 indicates that the Observe-self condition tended to perform more poorly in control tests in the second problem, whereas the reverse is indicated for the Observe-other condition.

A 2x2x2 ANOVA was carried out, using control test (Control Test 1, Control Test 2) and problem (1st Problem, 2nd Problem) as within subject variables, and condition (Observe-self, Observe-other) as the between subject variable. The analysis showed a significant main effect of control test, $F(1, 30) = 7.72$, MSE = 0.22, $p < 0.01$, $\eta^2 = 0.21$. There was a significant main effect of condition, $F(1, 30) = 4.99$, MSE = 0.16, $p < 0.05$, $\eta^2 = 0.14$, and a significant Condition x Problem interaction, $F(1, 30) = 13.17$, MSE = 0.28, $p < 0.001$, $\eta^2 = 0.31$. No other main effects or interactions were significant. The decrease in performance between problems of the Act-on-self-high condition (Figure 10) was confirmed by a post-test of error scores, $t(31) = -3.05$, $p < 0.005$, $d = -1.09$. Moreover, the control performance of the Observe-other condition significantly increased in the second problem, $t(31) = 2.31$, $p < 0.05$, $d = 0.83$. Thus, the evidence suggests that across problems control performance increased in the Observe-other condition, and decreased in the Observe-self condition, consistent with the findings in Experiment 1.

Goal Specificity effect. Although cross comparisons are not always advised, Experiments 1 and 4 were identical, with two exceptions: Experiment 4 included more learning trials, and the learning phase was goal specific. Therefore, to examine the presence of the goal specificity effect, Observe-self and Observe-other conditions from Experiments 1 and 4 were compared, based on control error scores from Control Test 1. Control Test 1 was used because during the learning phase participants in Experiment 4 were trained to control the system to the criteria of this test. Burns and Vollmeyer's Dual-Hypothesis theory predicts that, despite extra training, the overall control performance of participants in Experiment 4 should be poorer than in Experiment 1, because they were focused more on generating instances relevant to the SG than hypothesis testing. Columns 1 and 3 in Figure 10 show that the mean error scores of both conditions in Experiment 4 appear to be greater than both conditions in Experiment 1. To analyses this, a 2x2x2 ANOVA was conducted using problem (1st Problem, 2nd Problem) as the within subject variable, and condition (Observe-self, Observe-other) and experiment (Experiment 1, Experiment 2) as between subject variables. There was a significant main effect of Experiment, $F(1, 64) = 14.81$, MSE = 0.56, $p < 0.0005$, $\eta^2 = 0.18$, and a significant Condition x Problem interaction, $F(1, 64) = 21.67$, MSE = 0.51, $p < 0.0005$, $\eta^2 = 0.25$. The evidence confirms the prediction that goal specificity effects were found, and that introducing a SG during learning impairs later control performance. The source of the interaction was also examined, and further tests revealed that, consistent with the findings throughout this study, Observe-self conditions showed poorer control performance in the second problem, whereas Observe-other conditions showed improved control ability.

Comparison of structure test scores in 1st and 2nd Problem. Figure 11 shows that, for the Observe-self condition, structure scores decreased across problems, whereas for the Observe-other condition they increased.

A 2x2 ANOVA was carried out, using problem (1st Problem, 2nd Problem) as the within subject variable, and condition (Observe-self, Observe-other) as the between subject variable. There was a significant main effect of Condition, $F(1, 30) = 6.20$, MSE = 54.89, $p < 0.05$, $\eta^2 = 0.17$, and Condition x Problem interaction, $F(1, 30) = 5.38$, MSE = 33.45, $p < 0.05$, $\eta^2 = 0.15$. The decrease in performance between the first and second problem for the Observe-self condition (Figure 11) was confirmed, $t(15) = 2.27$, $p < 0.05$, $d = 1.17$. The increase in performance across problems of the Act-on-other-low condition approached significance, $t(15) = -2.05$, $p = 0.058$, $d = -1.06$.

Judgments of self-efficacy and structural knowledge. The following analyses examined the association between judgments of self-efficacy and control performance, averaged across Control Test 1 and 2 for each problem. Participants' efficacy judgments prior to the control test phase accurately tracked their control performance in the first problem, $r(32) = 0.53$, $p < 0.005$, and the second, $r(32) = 0.60$, $p < 0.0005$. No other correlational analyses were significant.

A further analysis was conducted to examine the pattern of judgments of self-efficacy between conditions. A 2x2x2 ANOVA on judgments of self-efficacy, using Stage (before control test phase, after control test phase), and Problem (1st Problem, 2nd Problem) as the within subject variables, and Condition (Observe-self, Observe-other) as the between subject variable. The analysis revealed a significant Stage x Condition interaction, $F(1, 30) = 10.42$, MSE = 7812.50, $p < 0.005$, $\eta^2 = 0.26$. T-tests showed that there was a difference between conditions, based on self-efficacy judgments recorded prior to the control phase in the second problem, $t(15) = 2.29$, $p < 0.05$, $d = 1.18$. The mean estimate of how close participants in the Observe-self condition believed they could approach to the criteria values was +/- 151.56 (SD 30.91), whereas for the Observe-other conditions the range was +/- 125 (SD 38.96). Thus, judgments of self-efficacy revealed that, prior to the control phase of each problem, judgments accurately tracked performance, and that, in the second problem, there were differences between conditions, showing that the Observe-self judgments of self-efficacy were lower than those of the Observe-other condition.

Participants were also asked to judge to what extent they used their understanding of the structure of the system in the first problem again in the second problem. Responses were made on a 9-point scale, with 1 indicating not at all, and 9 indicating that they had mostly relied on their previous knowledge. For the Observe-self condition, the mean response was 2.5 (SD 1.93) and for the Observe-other condition it was 5.25 (SD 2.29). The responses between conditions were compared, and indicated that the Observe-other condition relied more on their prior knowledge to help them in the second problem than the Observe-self condition, $t(30) = -3.67$, $p < 0.001$, $d = -1.34$. These judgments were also correlated with mean structure scores in the second problem, and were found to be significant, $r(32) = 0.66$, $p < 0.005$. In addition, the structure judgments were correlated with control performance averaged across Control Tests 1 and 2 in the second problem. The analysis revealed a significant negative relationship, $r(32) = -0.45$, $p < 0.005$, suggesting that participants who indicated that they relied more on their knowledge of the system from the first problem also showed better control performance in the second problem.

Post-test question. Seventy-five percent of participants in the Act-on-self and 81% in the Act-on-other conditions were able to accurately report which condition they were in.

Discussion

The evidence from Experiment 4 replicated the findings from Experiment 1. The Observe-self condition showed negative transfer of control performance and accuracy of structural knowledge across problems, and the Observe-other condition showed positive transfer of performance in measures of control and structural knowledge. The evidence also confirmed both predictions made concerning judgments of self-efficacy. Prospective judgments of self-efficacy of control performance were associated with actual control performance, whereas retrospective judgments taken after the control test phase of each problem were not associated with control performance. In addition, self-efficacy judgments accurately tracked the different patterns in control performance of the two conditions in the second problem: Estimations of control performance were lower in Observe-self conditions than in Observe-other conditions. Although it is not the focus of Experiment 4, the introduction of a SG during the learning phase of both problems enabled a comparison with NSG learning of equivalent conditions in Experiment 1. The findings revealed that, despite the increased number of learning trials, overall control performance in Control Test 1 was poorer in Experiment 4 than in Experiment 1. Thus, Experiment 4 provides compelling support that goal specificity is a robust phenomenon that generalizes to conditions in which learning is observation-based. This has not previously been demonstrated.

After the learning phase of the second problem, participants were asked to estimate to what extent they had used their knowledge of the structure of the first control system as a basis for understanding the structure of the second control system. The evidence confirmed the prediction that participants in the Observe-self condition tended to report relying less on their previous knowledge from the first problem than the Observe-other condition. In addition, the evidence revealed that the estimations made were strongly associated with the accuracy of structural knowledge in the second problem, and control performance, indicating that, for the Observe-other condition, experience with the first problem had a facilitative effect in the second, whereas the Observe-self condition did not use structural knowledge of the first problem in the second. This finding complements prospective judgments of self-efficacy taken in the second problem. Taken together, they indicate that negative evaluations of the effectiveness of the learning instances may also have led participants in the Observe-self condition to discard potentially relevant structural information from the first problem. For the Observe-other conditions, positive evaluation of the learning phase may have given them reason to draw comparisons with their own experiences, which they then viewed as relevant.

General Discussion

The objective of this study was to uncover the effects of monitoring on the transferability of control behaviors across analogical skill learning tasks. The influence of self-regulatory behaviors on control performance was most revealed in conditions where problem solvers were re-exposed to their prior self-generated learning instances. Experiments 1-4 showed that this manipulation produced negative transfer of control performance and structural knowledge across identical and perceptually dissimilar analogical tasks, regardless of whether regulatory mechanisms operated on online or observed behaviors. Experiments 2 and 3 revealed that problem solvers are biased towards their evaluation of the effectiveness of their learning experiences, and this later affected their control of the system. Experiment 4 demonstrated that individuals can make accurate estimations of their control behavior based on the evaluation of the knowledge they had acquired while learning. Furthermore, consistent with Social Cognitive theory, negative and positive transfers of control performance across problems were accurately tracked by judgments of self-efficacy.

The following discussion examines the implications of these findings for Social Cognitive theory, Dual-Space theory, and implicit learning theories of skilled learning. In particular, the issues explored concern knowledge acquisition in CDCTs, transfer of control behaviors, and the relationship between monitoring and control and its effect on transfer of skilled behaviors.

What Kind of Knowledge is Learnt in Complex Skilled Control Tasks?

Implicit learning theorists (Berry, 1991; Berry & Broadbent, 1988; Dienes & Berry, 1997; Lee, 1995; Sun et al., 2001) have proposed that knowledge acquired in CDCTs and experience in controlling them is embedded within the interactions problem solvers have with the system. This is why only learning through action can produce successful control behaviors: This involves storing individual instances or exemplars of specific condition-action states (i.e., the state of the system, the subsequent inputs changed, and their corresponding outputs). Thus, mastering a control task requires matching the goal and the current situation to previously encountered instances, to determine the next appropriate response. Knowledge is conscious only to the extent that the appropriate response in any given situation can be stated, but what generated this response is unavailable to consciousness (Buchner et al., 1995; Dienes & Berry, 1997; Dienes & Fahey, 1998). The empirical foundation of this position is the phenomena showing that declarative knowledge is dissociated from procedural knowledge. The patterns of dissociations reported are often compared to findings from other implicit learning tasks: artificial grammar learning (Reber, 1989), and sequence learning (Nissen & Bullemer, 1987). Common to these implicit learning paradigms is the claim that knowledge is acquired incidentally, because the target goal (i.e., controlling a system to criterion, memorizing a sequence of letters, responding quickly to a moving stimulus) is reached independently of knowledge of the rule or underlying structure of the task.

However, one problem with classifying CDCTs as implicit learning tasks is that, in contrast to studies showing that rule-search interferes with successful control performance (e.g., Berry & Broadbent, 1988), Experiments 1-4 showed that searching for rules did not impede transfer of control behaviors, and rule-based knowledge was associated with procedural knowledge (e.g., Burns & Vollmeyer, 2002; Brehmer, 1992; Gibson et al., 1997; Schoppek, 2002; Vollmeyer et al., 1996). Moreover, problem solvers showed accurate insight into the methods by which they arrived at their responses, by discriminating between self-generated and other-generated learning instances. This occurred whether the learning phase was procedural-based (Experiments 1, 2, 4) or observation-based (Experiments 1, 4). Importantly, accuracy was not affected by the format in which knowledge acquisition took place, suggesting that participants were able to monitor the strategies they developed both online, and indirectly while observing their behaviors. Finally, they detected the origin of the learning phase in CDCTs with only 12 learning trials (Experiment 1, 2) and with extended practice of 40 trials (Experiment 4). This is consistent with studies showing that self-reports, like think aloud protocols, accurately track control performance (e.g., Burns & Vollmeyer, 2002; Gonzalez et al., 2003; Gonzales & Quesada, 2003; Jensen & Brehmer, 2003). Overall, these findings are particularly hard to reconcile within an implicit learning account that claims knowledge is acquired incidentally and is unavailable to consciousness.

What therefore might explain the differences found between evidence for dissociations and, as found in the present study, associations between procedural and declarative knowledge? Dissociations are typically found where learning is SG-based (e.g., Berry, 1991; Berry & Broadbent, 1984, 1987, 1988; Broadbent et al., 1986; Dienes & Fahey, 1995, 1998; Marescaux et al., 1989; Stanley et al., 1989). A well established research history of CDCTs shows that SG learning leads to an impoverished understanding of control tasks, because only states of the system that advance the problem solver to the SG are explored. 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, 1987, 1988; Broadbent et al., 1986; Dienes & Fahey, 1995, 1998; Marescaux et al., 1989). 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; Sanderson & Vicente, 1986).

When information search (i.e., NSG learning) is encouraged (Experiments 1-3), problem solvers tend to adopt a hypothesis testing strategy (Burns & Vollmeyer, 2002; Funke, 2001; Geddes & Stevenson, 1997; Sweller, 1988; Vollmeyer et al., 1996). Without an imposed constraint, a broader range of instances of the system are explored (Burns & Vollmeyer, 2002; Sweller, 1988; Vollmeyer et al., 1996). Under these learning conditions, problem solvers do not advance towards any particular goal during learning (Experiment 1-3), and so instances are unlikely to be stored as successful condition-action exemplars, as described by Broadbent (Broadbent et al., 1986) and others (Dienes & Fahey, 1995, 1998; Marescaux et al., 1989). Rather, as Dual-Space theory claims, knowledge of the system is formed into both rules and exemplars, which are refined as the problem solver develops systematic ways of exploring the system. Rule formation can even be encouraged under SG-learning conditions, so long as knowledge of the relations between inputs and outputs is examined through self-evaluative processes (i.e., verbal protocols), or tested directly during learning (Experiments 4), which has been shown to prompt problem solvers to evaluate the status of their declarative knowledge (Burns & Vollmeyer, 2002; Sanderson, 1989; Sanderson & Vicente, 1986; Voss, Wiley & Carretero, 1995).

So What Knowledge is Acquired in CDCTs?

The findings from this study can be taken as further support for Burns & Vollmeyer's (2002) Dual-Space theory (Klahr & Dunbar, 1988; Klahr, Fay, & Dunbar, 1993; Simon & Lea, 1974) in which a skill is acquired by searching through different spaces of a task. Depending on the goal of the task, problem solvers develop hypotheses of how the control system works, which they evaluate by testing and updating (Burns & Vollmeyer, 2002; Geddes & Stevenson, 1997; Klahr et al., 1993; Sanderson, 1989; Sanderson & Vicente, 1986; Vollmeyer & Rheinberg, 2000). As a result, problem solvers develop rule-based knowledge and exemplar-based knowledge of the system. During SG learning, hypothesis-testing behaviors are attenuated because the solver focuses more on instance-based learning, in much the same way as described by implicit learning theories of control tasks. However, because CDCTS are not treated as purely proceduralized tasks, knowledge is still thought of as available to consciousness.

This account predicts decrements in control performance and structural knowledge in SG (Experiment 4) compared with NSG (Experiment 1) learning environments, and predicts associations between control behaviors and structural knowledge (Experiments 1-4). Finally, for learning to occur, evaluative processes are used to update hypotheses that are tested (Experiments 3, 4). The way in which evaluative processes inform and update hypotheses that are then used to control CDCTs shares some similarities with the relationship between monitoring and control behaviors proposed by Social Cognitive theory. Experiments 3 and 4 showed that, consistent with this theory, self-regulatory mechanisms, such as perceived self-efficacy, strongly influence skilled behavior (e.g., Bandura, 1991; Bandura & Locke, 2003; Earley, Connolly, & Ekegren, 1989; Kanfer et al., 1994; Karoly, 1993; Lehmann & Ericsson, 1997; Rossano, 2003; Tversky & Kahneman, 1974).

Successful Transfer of Complex Control Behaviors

Previous studies that claim that skill learning in CDCTs is procedural, incidental, and implicit also propose that the transferability of these behaviors is limited, because implicit knowledge is inflexible and perceptually bound (e.g., Berry, 1991; Berry & Broadbent, 1988; Dienes & Berry, 1997; Lee, 1995; Sun et al., 2001). For example, tasks that were perceptually and structurally similar were shown to facilitate successful transfer of control behaviors only (Berry & Broadbent, 1988). Moreover, when a hint as to the underlying similarity of the tasks was presented, regardless of whether the tasks were perceptually similar or dissimilar, no transfer was found. Berry (1991; Berry & Broadbent, 1988) has claimed that conditions that encourage explicit search of structural-based knowledge (e.g., introducing a hint, explicitly instructed to search for rules, observation-based learning) prevent the storage of relevant exemplar-based knowledge that can later be matched to transfer tasks. Moreover, consistent with the view that procedural and declarative knowledge are independent, Berry's findings show that control behaviors can transfer without any accompanying structural knowledge, further emphasizing the dependence of control behaviors on implicit learning.

If, as implicit theorists claim, transfer occurs exclusively for control behaviors, which are impeded by explicit processing, and explicit knowledge itself does not transfer across tasks, then why was there evidence of transfer of both types of knowledge in the present study? Both control behaviors and structural knowledge were shown to transfer successfully across perceptually different but analogous control tasks (Experiments 1 & 4), and when learning in the tasks was SG-based (Experiment 4). These different forms of knowledge transferred across problems, independently of whether knowledge acquisition in both problems was in the same format (procedural-based, Experiments 1, 2) or in a different format (procedural-based and observation-based, Experiments 1, 4), or was biased by extraneous information (Experiment 3).

In this study, the approach to understanding skilled learning in CDCTs posits that control behaviors are mediated by evaluative mechanisms that monitor behavior. In particular, this was encouraged because the learning environment of the problems (Experiments 1-3) encourages hypothesis-testing behavior (e.g., Burns & Vollmeyer, 2002; Sweller, 1988; Vollmeyer et al., 1996; Vollmeyer & Rheinberg, 2000). This necessarily involves the exploration, generation, and evaluation of hypotheses (Klahr & Dunbar, 1988; Klahr et al., 1993), which can be biased by extraneous but highly influential information (Jacobs et al., 1984; Wisniewski & Medin, 1991). Burns & Vollmeyer (2002) showed that, even under conditions in which learning is SG-based, problem solvers allocate some learning trials to hypothesis testing. Experiment 4 also showed that the advance knowledge problem solvers had of the number of learning trials encouraged them to explore the system for some blocks of the learning phase.

The exploration and evaluation of hypotheses during learning leads to the development of structural and procedural knowledge, and forges their association (Brehmer, 1992; Burns & Vollmeyer, 2002; Funke, 2001; Sanderson, 1989; Vollmeyer et al., 1996). Knowledge of the system is not perceptually bound to the learning environment it was acquired in, because it is not limited to exemplars of successful outcomes. Taken together, this suggests that successful transfer, of the kind reported in this study, is the result of the formation of structural knowledge, and its association with experiences of controlling the system. This enables problem solvers to match their prior knowledge with the structural and procedural properties of the transfer problem, and, by assessing the similarity, they can bring relevant knowledge to bear on the task. Where dissociations are found (Berry & Broadbent, 1988; Dienes & Fahey, 1995, 1998), problem solvers learn without the benefit of hypothesis testing (e.g., Geddes & Stevenson, 1997; Gonzales, Vanyukov, & Martin, 2005; Sweller, 1988; Vollmeyer et al., 1996), and in some cases, through explicit instruction, are actively discouraged from doing so (e.g., Berry, 1991). Along with preventing associations between declarative and procedural knowledge, constraining hypothesis testing also appears to limit the extent to which control behaviors transfer. Therefore, the involvement of evaluative behaviors during learning explains why problem solvers successfully transferred their experience of controlling a system and their knowledge of its causal structure, and why this was not found in earlier implicit learning studies.

The Relationship between Monitoring and Control and its Effect on Transfer of Skilled Behaviors

First, this discussion does not presuppose that the relationship between monitoring and control behaviors is unidirectional. The direction of causality has been described as going from subjective judgments to behaviors (e.g., Koriat & Goldsmith, 1996; Nelson & Narens, 1990; Son & Schwartz, 2002), and from behaviors to subjective judgments (Benjamin & Bjork, 1996; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Kelley & Jacoby, 1996; Matvey, Dunlosky, & Guttentag, 2001). Recently, Koriat (1998; Koriat et al., 2006) has suggested that monitoring and control alternate in a cascaded pattern, with the feedback from the control operation serving as the input for later monitoring, and so on. Similarly, Bandura (Bandura & Wood, 1989; Bandura & Jourden, 1991; Wood & Bandura, 1989) and others (Seijts & Latham, 2001; Tabernero & Wood, 1999) have shown that prior performance influences perceived self-efficacy, which in turn affects subsequent performance, in much the same way as Koriat described.

Prior research has shown a complex interplay between regulatory mechanisms and control behaviors. Bandura and Locke's (2002) Social Cognitive theory, and Locke and Latham's (2002) theory of goal setting, have proposed that goals are integral to the way skilled behaviors are regulated. This is also compatible with the Dual-Space theory of goal specificity effects in CDCTs (Burns & Vollmeyer, 2002; Klahr & Dunbar, 1988; Vollmeyer et al., 1996). Goals inform how behavior is assessed (Bandura, 1997; Bandura & Locke, 2002; Locke & Latham, 2002). Whether self-set, or externally imposed, goals are essential to learning about controlling CDCTs (Burns & Vollmeyer, 2002; Trumpower, Goldsmith, & Guynn, 2004). Hypothesis testing, which enables problem solvers to learn about CDCTs, is also a goal-directed activity (Glaser & Bassok, 1989; Voss et al., 1995). Bandura's Social Cognitive theory proposes two self-regulatory systems (Bandura, 1991) that differentially affect the way in which goals are reached: reactive and proactive. Consistent with this, the present study shows that the regulation and transfer of control behaviors in CDCTs is arbitrated by goals, and that regulatory mechanisms operate reactively and proactively over the transferability of control behaviors.

In Experiment 3, bogus information was effective in reproducing the patterns of transfer found in Experiments 1 and 2, because control performance during the test phase of the first problem was perceived as diagnostic of the effectiveness of the self-generated learning instances. The control-monitoring-control relationship was also demonstrated in Experiment 4. Here, the learning phase was SG-based, so that participants could now assess their control skills. Their performance during learning informed their judgments of self-efficacy, which were also comparable with performance on subsequent control tests. The benefits of comparing one's own method of learning about a CDCT, with that of another, enabled problem solvers to gauge whether to adjust their own knowledge (Experiments 1, 2, 4). Even when this comparison was based on self-generated learning instances, assumptions about the similarity between the self's and a supposed other's strategies were also enough to enable self-regulatory mechanisms to facilitate transfer (Experiment 2, disguised-self). This is entirely consistent with the behaviors attributed to the proactive regulatory system. Here individuals use performance standards to consolidate both the knowledge they have gained and beliefs in their self-efficacy, which enhance performance (Bandura & Cervone, 1986).

Without external normative standards, such as those presented in Experiment 3, or comparison with another's learning experiences, self-perceptions of the knowledge and control ability of *Self*-conditions (Experiments 1, 2 (Act-on-self), and 4) lead to negative self-assessments. Experiment 4 also revealed that, consistent with poor self-efficacy judgments, individuals undervalued the relevancy of previously gained knowledge in assisting them in the transfer task. This is consistent with behaviors attributed to the reactive regulatory system. However, in this case, error detection and correction occurs had devastating effects on the way in which relevant task information was attenuated.

Conclusion

Previous research on CDCTs has provided an impoverished understanding of the types of knowledge that are transferable and the modulating factors that lead to successful and unsuccessful transfer of skilled behaviour. The present article was designed to address this and, by studying transferability of skills, was able to provide new insights into the learning process that takes place in complex dynamic tasks. The evidence revealed an association between declarative and procedural knowledge (Experiments 1-4), the acquisition of procedural skills through observation (Experiments 1, 4), a goal specificity effect via observation-based learning (Experiments 1 vs. 4), and accurate monitoring of internally represented behaviors and self insight (Experiments 1, 2). Many of these findings are closely aligned with, and extend, previous work on self-regulatory mechanisms and control behaviors, and goal specificity effects in CDCTs and Multiple-Cue learning tasks.

The two most pivotal findings showed successful transfer across analogous CDCTs (Experiments 1, 4), and the atypical negative transfer effect (Experiments 1, 4) in which control performance and structural knowledge in the transfer problem were impaired relative to the original. Both Social Cognitive theory and Dual-Space theory provide foundations for claiming that problem solvers are sensitive to, and influenced by, their assessments of the effectiveness of self-generated learning instances, and that this plays a significant role in facilitating and attenuating the transferability of knowledge in CDCTs.

Footnotes

1. In Burns and Vollmeyer’s study, participants were shown the starting values of input and output values before they began the task. In the present 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. The rationale for this change was simply to encourage participants to pay special attention to the effects on the outputs resulting from the manipulations they made.
2. The mean discrepancy between achieved values and target values for each participant in each control test, problem, condition, and experiment was ranked and used as a basis for generating the values +/-20 and +/- 200 used in Experiment 3. These values were the extreme ends of the range generated.

References

Albright, L., & Malloy, T. E. (1999). Self-observation of social behavior and metaperception. *Journal of Personality and Social Psychology, 77*, 726-734.

Bailey, K. G., & Sowder, W. (1970). Audiotape and videotape self-confrontation in psychotherapy. *Psychological Bulletin*, *74*, 127-137.

Bandura, A. (1986). *Social foundations of thought and action*: *A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*, 248-287.

Bandura, A., & Cervone, D. (1986). Differential engagement of self-active influences in cognitive motivation. *Organizational Behavior and Human Decision Processes*, *38*, 92-113.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88,* 87-99.

Bandura, A., & Wood, R. E. (1989). Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology*, *56*, 805-814.

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 309–338). Hillsdale, NJ: Erlbaum.

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.

Braucht, G. N. (1970). Immediate effects of self-confrontation on the self-concept. *Journal of Consultant Clinical Psychology*, *35*, 95-101.

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

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *Journal of Social Psychology*, *130*, 353-363.

Broadbent, D. (1977). Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology*, *29*, 181-201.

Broadbent, D., & Ashton, B. (1978). Human control of a simulated economic system. *Ergonomics*, *78*, 1035-1043.

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 CDCTs. *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.

Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem-solving by 6-year-olds. *Cognitive Development, 4*, 327-344.

Cohen, M. S., Freeman, J. T. & Wolf, S. (1996). Meta-recognition in time stressed decision making: Recognizing, critiquing and correcting. *Human Factors, 38,* 206–19.

Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative view. *Annual Review of Psychology*, *51*, 171-200.

DeShon, R. P., & Alexander, R. A. (1996). Goal setting effects on implicit and explicit learning of complex tasks. *Organizational behavior and human decision processes*, *65*, 18-36.

Dienes, Z., & Berry, D. (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.

Dowrick, P. W. (1983). Self-modeling. In P. W. Dowrick & S. J. Briggs (Eds.), *Using video: Psychological and social applications* (pp. 105-124). New York: Wiley.

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.

Elliot & Dweck, (1988). Goals: An approach to motivation and achievement. *Journal of personality and social psychology*, *54*, 5-12.

Ericsson, K. A. (Ed.). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games.* Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Lehman, A. (1996). Expert and exceptional performance: Evidence of maximal adaptation to constraints. *Annual Review of Psychology*, *47*, 273-305.

Fireman, G., & Kose, G. (1991). Video training as a means for enhancing self-awareness in problem solving among young children. *Resources in Education*, *330,* 492.

Fireman, G., & Kose, G. (2002). The effect of self-observation on children's problem solving. *Genetic Psychology*, *163*, 410-423.

Fireman, G., Kose, G., & Solomon, M. (2003). Self-observation and learning: The effect of watching oneself on problem solving performance. *Cognitive Development*, *18,* 339-354.

Fosnot, C., Forman, G., Edwards, C., & Goldhaber, J. (1988). The development of an understanding in balance and the effect of training via stop-action video. *Journal of Applied Developmental Psychology*, *9,* 1-26.

Funke, J. (2001). Dynamic systems as tools for analyzing human judgment. *Thinking and Reasoning*, *7*, 69-89.

Gandhe, S., Gordon A., Leuski A., Traum D., Oard D., First steps towards linking dialogues: mediating between free-text questions and pre-recorded video answers. *Army Science Conference,* 2004.

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.

Gibson, F. P., & Fichman, M., & Plaut, D.C. (1997). Learning in dynamic decision tasks: Computational model and empirical evidence. *Organizational Behavior and Human Decision Processes*, *71*, 1-35.

Giesler, B. R., Josephs, R. A., & Swann, W. B. (1996). Self-verification in clinical depression: The desire for negative evaluation. *Journal of Abnormal Psychology, 105,* 358-368.

Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, *40*, 631-666.

Gonzales, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27,* 591-635.

Gonzales, C., & Quesada, J. (2003). Learning in dynamic decision making: The recognition process. *Computational & Mathematical Organization Theory*, *9*, 287-304.

Gonzales, C., Vanyukov, P., & Martin, M. K. (2005). The use of mircoworlds to study decision making. *Computers in Human Behavior*, *21*, 273-286.

Griffiths, R. D., & Gillingham, P. (1978). The influence of videotape feedback of the self assessment of psychiatric patients. *British Journal of Psychiatry*, *133*, 156-161.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 22–34.

Hill, R. J., Gordon, A., and Kim, J., Learning the lessons of leadership experience: Tools for interactive case method analysis. In *Proceedings of the Twenty-fourth Army Science Conference*, 2004.

Hogarth, R. M., Gibbs, B. J., McKenzie, C. R. M., & Marquis, M. A. (1991). Learning from feedback: Exactingness and incentives. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, *17*, 734-752.

Hung, J., & Rosenthal, T. (1978). Theraputic videotape playback: A critical review. *Advances in Behaviour Research and Therapy, 1*, 103-135.

Jacobs, B., Prentic-Dunn, S., & Rogers, R. W. (1984). Understanding persistence: An interface of control theory and self-efficacy theory. *Basic and Applied Social Psychology*, *5*, 333-347.

Jensen, E., & Brehmer, B. (2003). Understanding and control of a simple dynamic system. *Systems Dynamic Review*, *19*, 119-137.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition [Monograph]. *Journal of Applied Psychology*, *74*, 657-690.

Kanfer, R., Ackerman, P. L., Murtha, T. C., Dugdale, B., & Nelson, L. (1994). Goals setting, conditions of practice, and task performance: A resource allocation perspective. *Journal of Applied Psychology*, *79*, 826-835.

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology*, *44*, 23-52.

Kelley, C. M. (1999). Subjective experience as a basis of “objective” judgments: Effects of past experience on judgments of difficulty. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp 515–536). Cambridge, MA: MIT Press.

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.

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

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science, 12,* 467-472.

Knoblich, G., & Prinz, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 456-465.

Koriat, A. (1998). Metamemory: The feeling of knowing and its vagaries. In M. Sabourin, F. I. M. Craik, & M. Robert (Eds.), *Advances in psychological science* (Vol. 2, pp. 461–469). Hove, England: Psychology Press.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Koriat, A., Ma’ayan, H., & Nussinson, R. (2006). The Intricate Relationships Between Monitoring and Control in Metacognition: Lessons for the Cause-and-Effect

Relation Between Subjective Experience and Behavior. *Journal of Experimental Psychology*: *General*, *135*, 36-69.

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

Lee, Y., & Vakoch, D. (1996). Transfer and retention of implicit and explicit learning. *British Journal of Psychology*, *87*, 637-651.

Lehmann, A.C. and Ericsson, K.A. (1997) ‘Research on Expert Performance and Deliberate Practice: Implications for the Education of Amateur Musicians and Music Students. *Psychomusicology, 16*, 40–58

Lerch, F. J., & Harter, D. E. (2001). Cognitive support for real-time dynamic decision making. *Information Systems Research*, *12*, 63-82.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*, 331-352.

Litt, M. D. (1988). Self-efficacy and perceived control: Cognitive mediators of pain tolerance. *Journal of Personality and Social Psychology*, *54*, 149-160.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

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

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human perception and Performance, 31*, 210-220.

Luchins, A. S. (1942) Mechanization in problem solving. *Psychological Monographs, 54*, Whole no. 248.

Marescaux, P.-J., Luc, F., & Karnas, G. (1989). Modes d'apprentissage selectif et nonselectif et connaissances acquies au control d'un processes: 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.

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition, 29,* 222–233.

Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science, 23*, 305-336.

Moritz, S. E., Feltz, D. L., Fahrbach, K. R., & Mach, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport*, *71*, 280-294.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation*: *Advances in research and theory* (Vol. 26, pp. 125–141). San Diego, CA: Academic Press.

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

Novick, L. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 510-520.

Owen, E., & Sweller, J. (1985). What do students do while solving mathematics problems? *Journal of Educational Psychology, 77*, 272-284.

Pintrich, P. R., & DeGroot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33-40.

Pressley, M., & Ghatala, E.S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist*, *25*, 19-33.

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.

Rapaport, A. (1967). Dynamic programming models for decision making. *Journal of Mathematical Psychology*, *4*, 48-71.

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

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1-29.

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* (Technical Report EPL-86-04). Urbana, IL: University of Illinois at Urbana-Champaign. Engineering Psychology Research Laboratory.

Schoppek, W. (2002). Examples, rules and strategies in the control of dynamic systems. *Cognitive Science Quarterly*, *2*, 63-92.

Seijts, G. H., & Latham, G. P. (2001). The effect of learning, outcome, and proximal goals on a moderately complex task. *Organizational Behavior and Human Decision Processes*, *46*, 118-134.

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.

Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect & B. S. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge, England: Cambridge University Press.

Spering, M., Wagener, D., & Funke, J. (2005). The role of emotions in complex problem solving. *Cognition and Emotion*, *19*, 1252-1261.

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 CDCT. *Quarterly Journal of Experimental Psychology*, *41*, 553-577.

Stanovich, K. E. (2004). *The robot's rebellion*. Chicago: University of Chicago Press.

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, *27*, 165-175.

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.

Tabernero, C., & Wood, R. E., (1999). Implicit Theories versus the Social Construal of Ability in Self-Regulation and Performance on a Complex Task. *Organizational Behavior and Human Decision Processes*, *78*, 104-127.

Trumpower, 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.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

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.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, *47*, 513-539.

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.

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, *10*, 293-309.

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

Weinberg, R. S., Gould, D., Yukelson, D., & Jackson, A. (1981). The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task. *Journal of Sport Psychology*, *4*, 345-354.

Wisniewski, E. J. & Medin, D. L. (1994) On the interaction of theory and data in concept learning. *Cognitive Science, 18*, 221–81.

Woltz, D., Bell, B., Kyllonen, P., & Gardner, M. (1996). Memory for order of operations in the acquisition and transfer of sequential cognitive skills. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 438-457.

Woltz, D., Gardner, M., & Bell. (2000). Negative transfer errors in sequential cognitive skills: Strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 601-625.

Wood, R. E., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology*, *56*, 407-415.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychology, 25*, 3-17.

Acknowledgements

Preparation for this article was supported by Economic and Social Research Council ESRC grant RES-000-27-0119. The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The work was also part of the programme of the ESRC Research Centre For Economic Learning and Human Evolution. The author wishes to thank David Shanks, David Lagnado, Maarten Speekenbrink, Belen Lopez, Chris Berry, Bob Hausmann, Bjoern Meder, Momme von-Sydow, York Hagmayer, and Micheal Waldmann for their comments and interest in this work.

Figure Captions.

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

Figure 2. Screen Shot of Water-Tank and Ghost-hunting Control Problems

Figure 3. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 1

Figure 4. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 1

Figure 5. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 2

Figure 6. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 2

Figure 7. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 3.

Figure 8. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 3

Figure 9. Mean number of inputs (±SE) varied for each block of each problem by condition in Experiment 4

Figure 10. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 4 and Experiment 4.

Figure 11. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 4

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

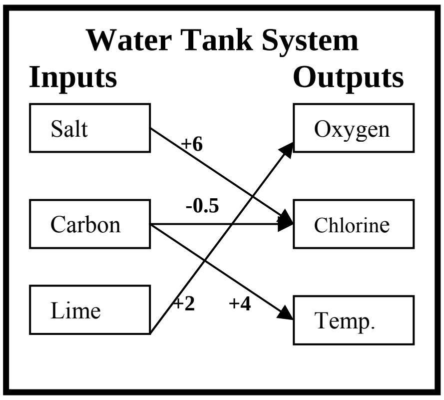
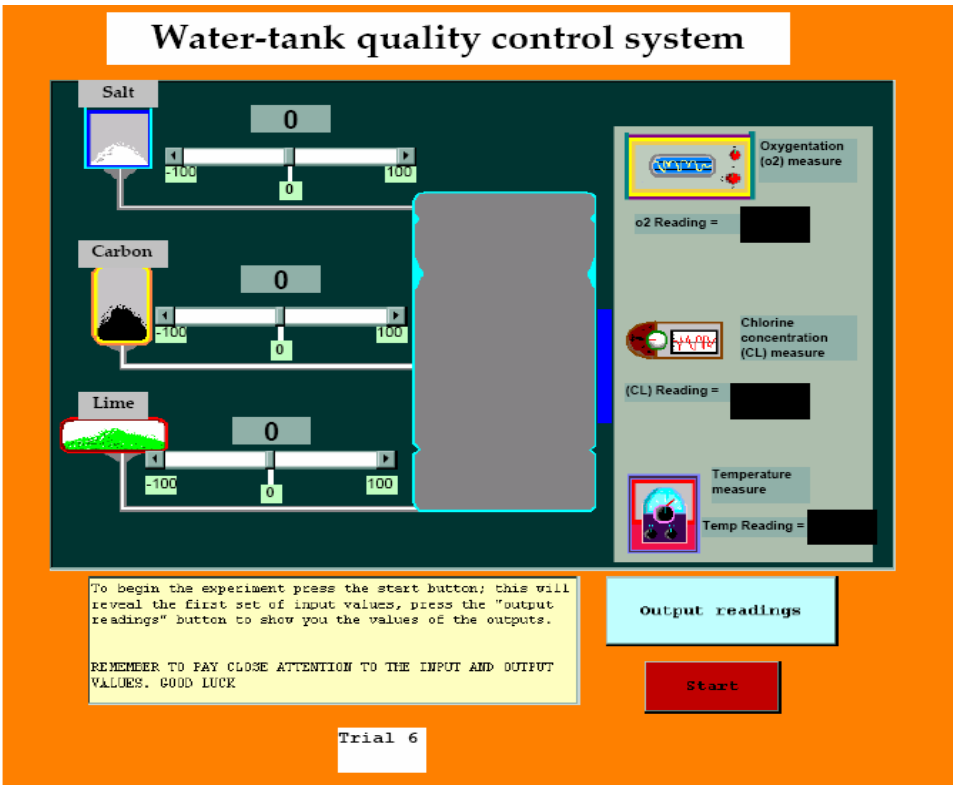


Figure 2. Screen Shot of Water-Tank and Ghost-Hunting Control Problems

Water-Tank Problem



Ghost-Hunting Problem

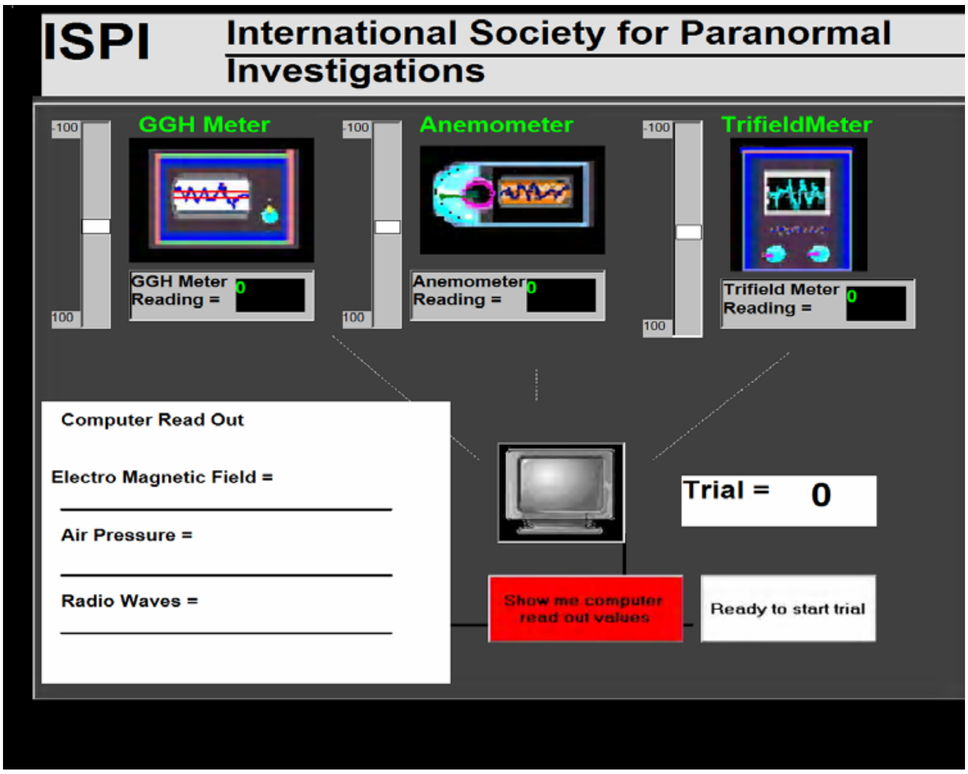


Figure 3. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 1

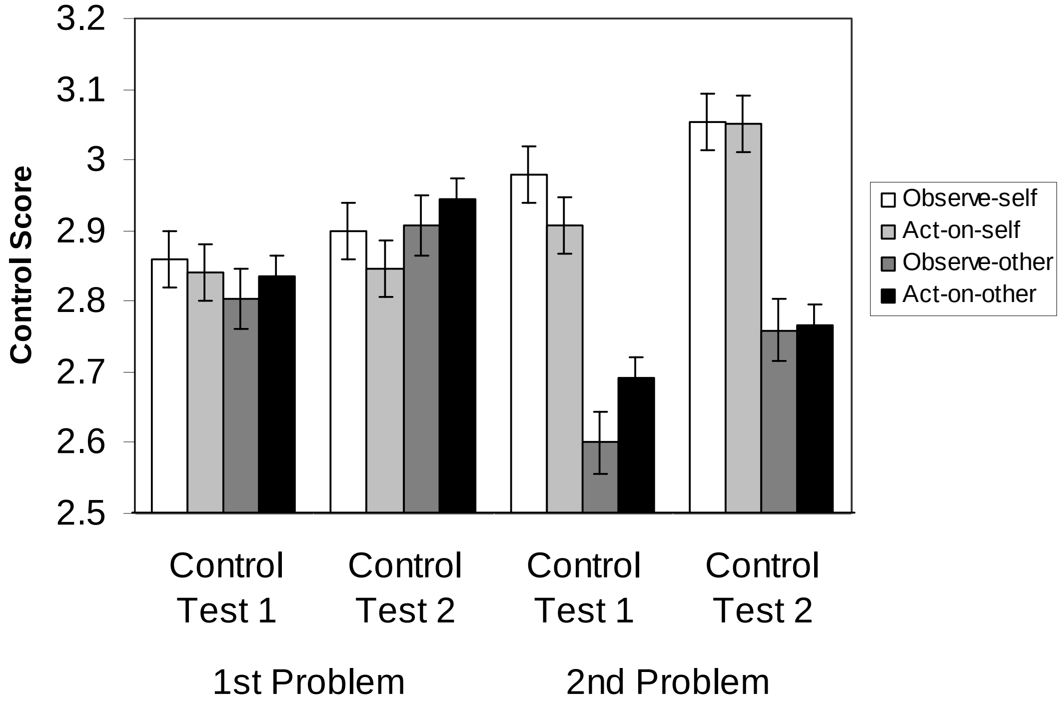


Figure 4. Structure test scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 1

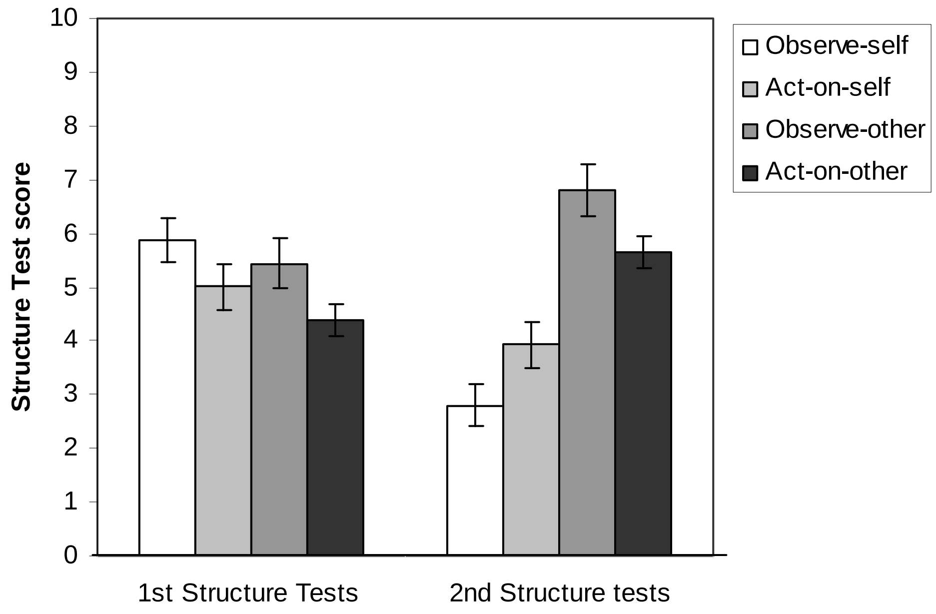


Figure 5. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 2.

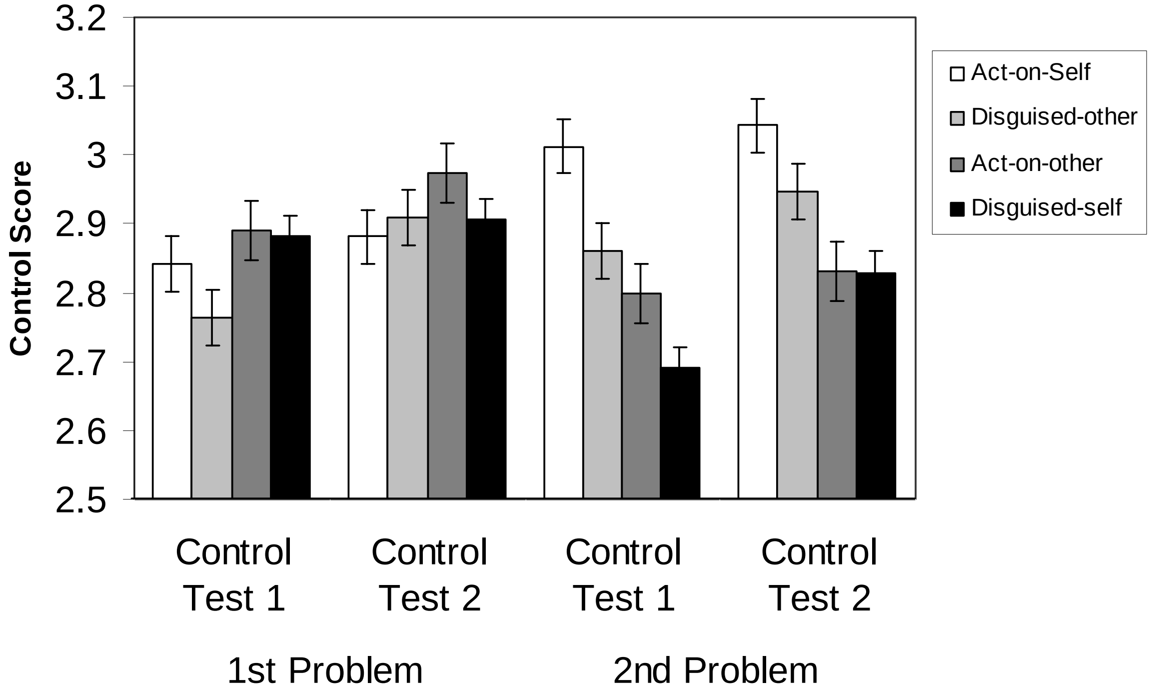


Figure 6. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 2

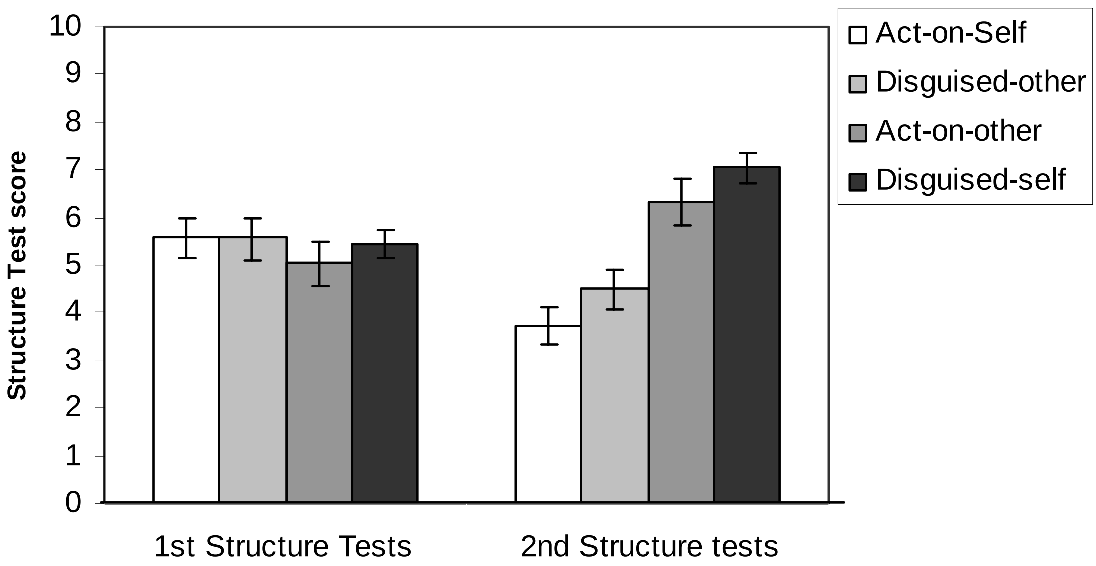


Figure 7. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 3.

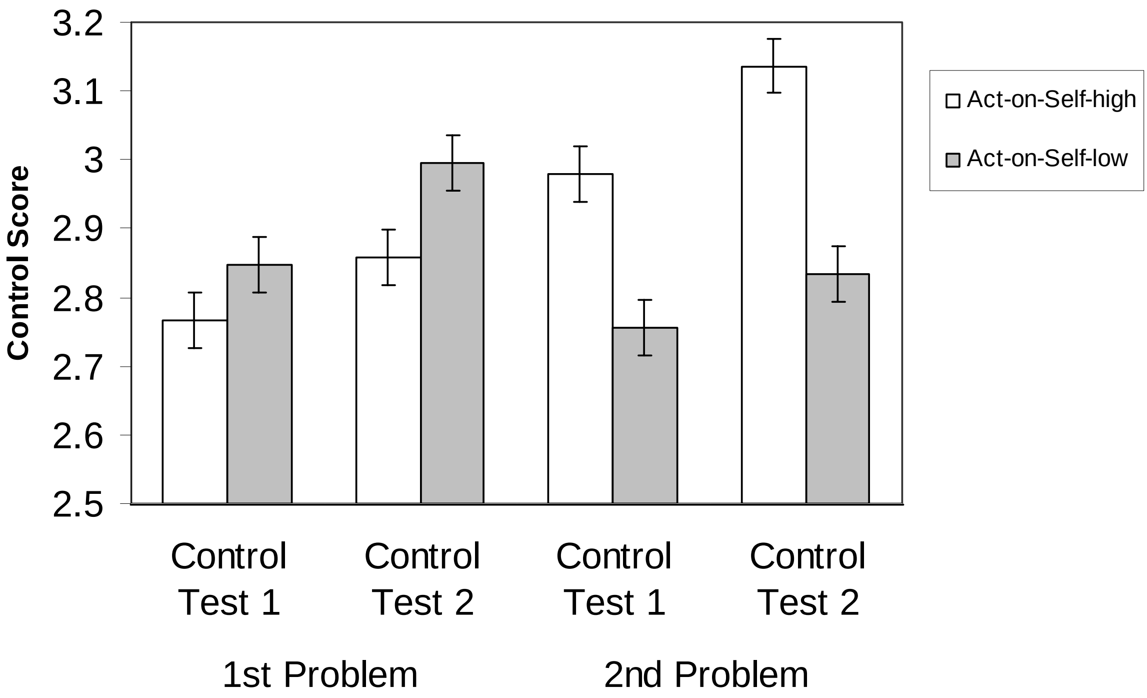


Figure 8. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 3

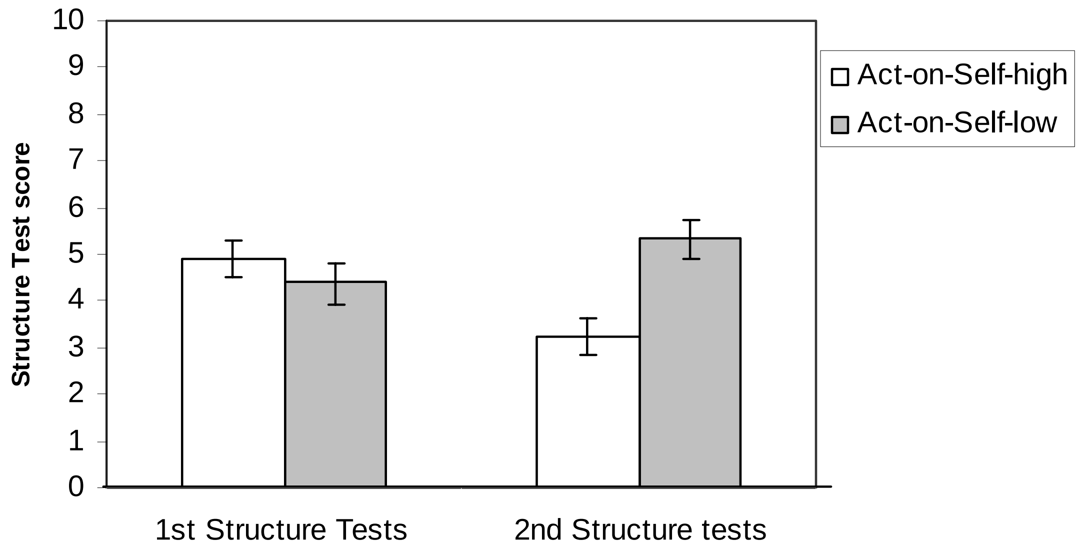


Figure 9. Mean number of inputs varied (±SE) for each block of each problem by condition in Experiment 4

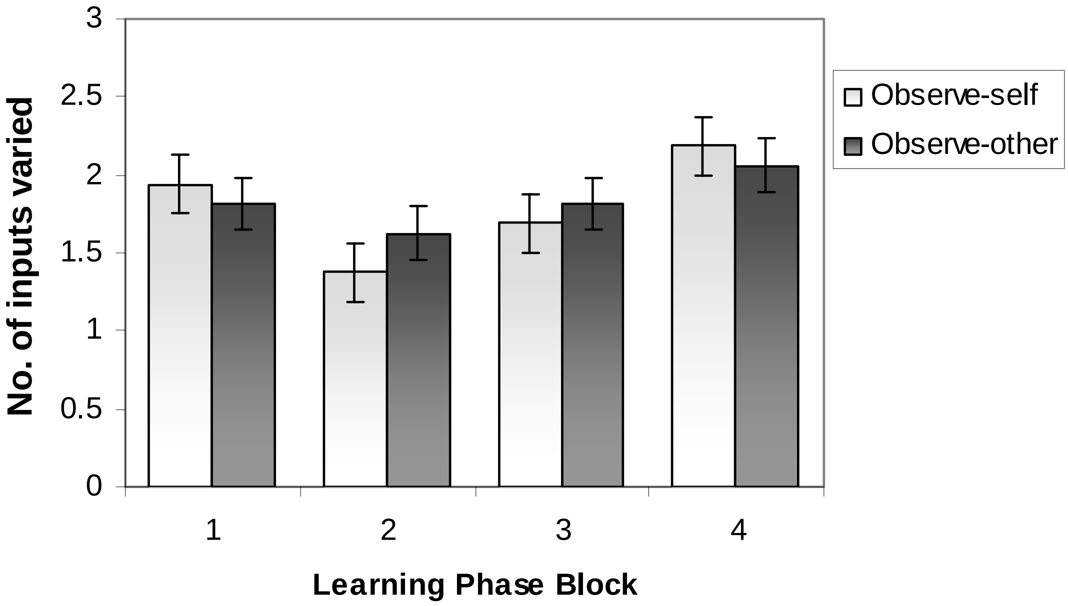


Figure 10. Mean Error scores (±SE) at Control Test 1 and Control Test 1 for each condition in Experiment 4 and Experiment 1.

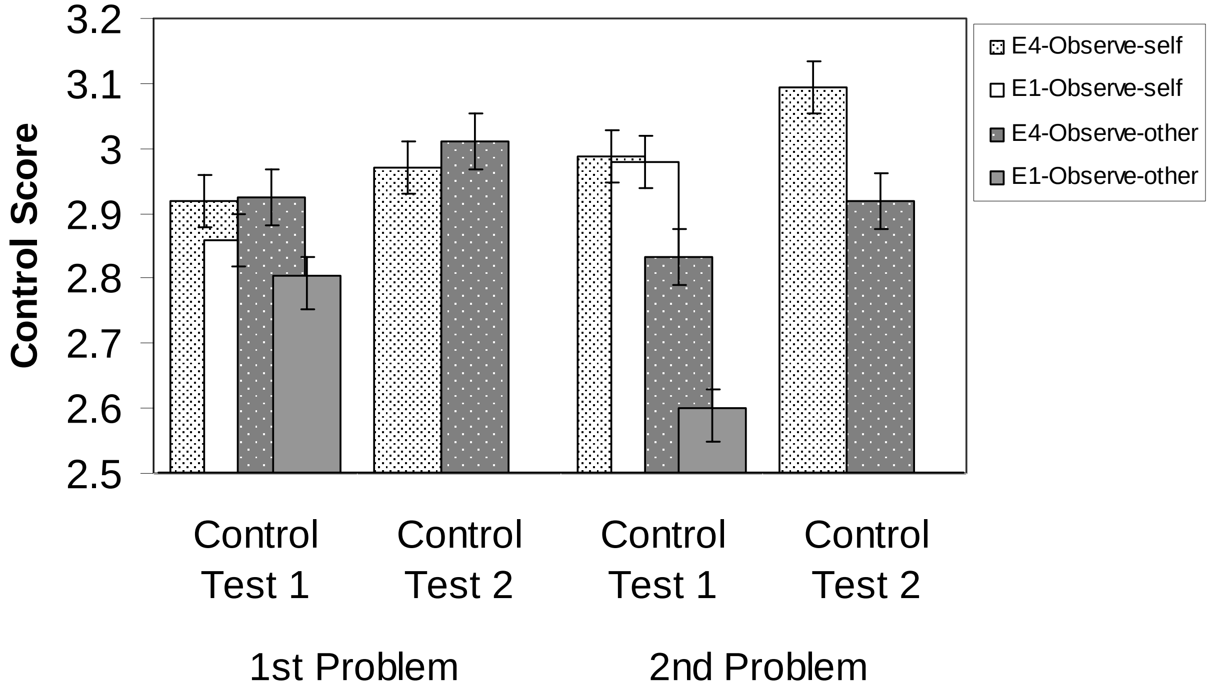
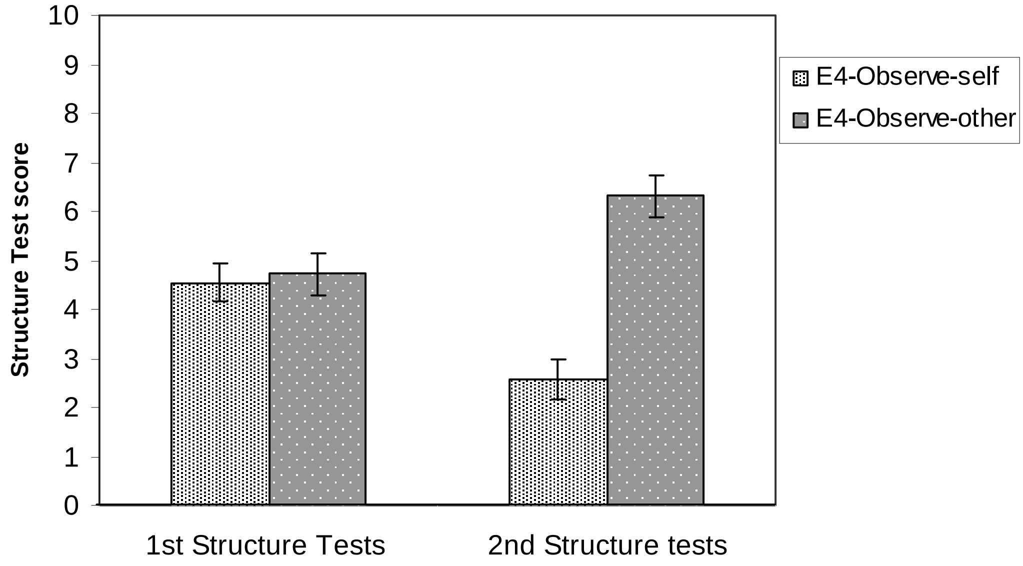


Figure 11. Structure scores (±SE) averaged across Structure Test 1, 2, 3, and 4 for each condition in Experiment 4



Appendix

Water Purification Tank Control System

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.

For each trial you should try to change only 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!

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.

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, this is because you will be required to imitate the worker's behavior later. Good Luck!

Ghost hunting Control System

Standard Instructions

Newspaper Report: Hillside, NJ Investigations, Utah State Library
Library worker John, his brother, and wife all reported seeing odd shadows out of the corner of their eyes. Most unusual was the report of the phone calls that came at 7:15 AM, certain mornings, that were riddles with static and no one on the other end of the call. The team of paranormal investigators went to investigate yesterday and was fully equipped with: a Trifield meter, a Anemomenter, a GGH meter. The investigation took place from 6:30 AM till approx 8:30 AM. Regular recordings were made. You were part of the team. You've done all the hard work and are back at the lab processing the data from the difference pieces of equipment you have used. Since you are new to this you aren't quite sure which of the three pieces of equipment (GGH meter, Anemomenter, Trifield meter) actually registers air pressure, radio waves, and the electro magnetic field – which are all disrupted when a ghost is present.

Action Instructions

You have a total of 12 trials in which you can test the equipment by altering the values of the meters and examining the computer readout for each of the output values: air pressure, radio waves, and electro magnetic field. For each trial you should try to change only 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 readings levels by pressing the button labeled 'show me readings'; this will reveal the computer readings. 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 you chose for the meters and the effects on the output readings. Good Luck!

Observation Instructions

You have a total of 12 trials in which you will observe the equipment being tested by one of your team. This will be done by altering the values of the meters and examining the computer readout for each of the output values of: Air pressure, radio waves and electro magnetic field. You will be presented with the different levels of the meters and the values of the three output readings. 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 computer output readings. 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 that are chosen for the meters and the effects on the output readings. Good Luck!