

DETERMINANTS OF THE CONTROL OF DYNAMIC SYSTEMS:
THE ROLE OF STRUCTURAL KNOWLEDGE

Natassia Goode

SID: 200123934

A thesis submitted for the degree of Doctor of Philosophy

School of Psychology, The University of Sydney

January, 2011

ACKNOWLEDGEMENTS

Firstly, I would like to thank the Accelerated Learning Laboratory and the University of Sydney for their financial support over the course of my PhD.

On a more personal note, I would like to thank Amirali Minbashian for his expert statistical advice and companionship while writing my thesis. In some ways I think our conversations kept me level headed. I would also like to thank Peter Jones for reading my thesis and letting me know when I became too verbose. Also much thanks to Petre Goode and Alira MacKenzie, my mother and sister, who constantly asked me when I was going to finish my thesis and thus encouraged me to get it done.

Finally, and most importantly, I would like to thank my supervisor Jens Beckmann. Without him, I would not have embarked on this project or discovered the world of complex problem solving. Over the past four years, he has been an amazing source of encouragement, inspiration and advice. But most of all, he has really taught me how to think critically and independently, and to strive for a high standard in my experimental work. It really has been an apprenticeship, and I can only hope that I have fully internalised these lessons.

ABSTRACT

In educational and organisational settings it has become common practice to use computer-based complex problems that represent dynamic systems for assessment and training purposes. In the interpretation of performance scores and the design of training programs, it is often assumed that the capacity to effectively control the outcomes of a dynamic system depends on the acquisition of structural knowledge. Control performance scores are generally interpreted as evidence of individual differences in the capacity to acquire and utilise structural knowledge and training programs typically try to improve learners' mental models of the system of interest. However, a causal relationship between the acquisition of structural knowledge and successful system control has not been established, and some findings suggest that it may be possible to control dynamic systems in the absence of structural knowledge.

Therefore, the goals of this project were to determine the conditions that are required to learn how to control dynamic systems and the psychological processes that separate successful from less successful problem solvers in the performance of this task. The main emphasis of this investigation was to clarify the role of structural knowledge in the control of dynamic systems and to identify sources of individual differences in problem solvers' capacity to acquire such knowledge and apply it in a goal-orientated application.

In a series of studies, a combined experimental and differential approach was adopted to address these goals. This consisted of the experimental manipulation of the task and structural characteristics of complex problems combined with the use of process indicators and external psychometric tests. Study 1 examined whether problem solvers need to directly interact with a dynamic system in order to acquire structural knowledge that is useful for system control. Study 2 examined whether increments in structural knowledge lead to improvements in control performance and whether dynamic systems can be successfully controlled without structural knowledge. Study 3 examined whether the relationship between structural knowledge and control performance is moderated by system complexity. Each of these studies also investigated the role of fluid intelligence in the acquisition and application of knowledge. Additional methodological contributions include the

application of Cognitive Load Theory to the design of the instructions used to manipulate structural knowledge, the use of randomly generated control performance scores to evaluate the success of performance and the development of a theoretically driven operationalisation of system complexity.

Across the studies, it was found that structural knowledge was a necessary condition of better than random performance and that there was a causal relationship between structural knowledge and control performance. However, the likelihood that structural knowledge would be acquired and utilised was found to be dependent on the complexity of the system. Small increments in system complexity resulted in floor effects on performance. Fluid intelligence was found to play a crucial role in the acquisition and subsequent application of knowledge. Overall, the results indicate that the complexity of the system determines the amount of knowledge that is acquired by the problem solver, which in turn, combined with their intelligence, determines the quality of their control performance.

TABLE OF CONTENTS

1. Aims and Overview

1.1 Introduction	1
1.2 Key terms	6
1.3 A combined experimental and differential approach	11
1.4 Overview of the thesis	12

2. Complex Problem Solving Research Approaches: A Critical Evaluation

2.1 Introduction	15
2.2 Common features across the approaches	16
2.3 The Dynamic Decision Making approach	18
2.4 The Implicit Learning approach	25
2.5 The Linear Structural Equation approach	29
2.6 Conclusions	36

3. The Role of Knowledge in Controlling a Dynamic System

3.1 Introduction	40
3.2 What type of knowledge is acquired?	40
3.3 Explanations for conflicting results	48
3.4 The acquisition of structural knowledge through direct instruction	55
3.5 Practice	58
3.6 Intervention vs. observation	59
3.7 Conclusions	60

4. Individual Differences in the Acquisition of Structural Knowledge and Control Performance

4.1 Introduction	63
4.2 Individual differences in the acquisition of structural knowledge	65

4.3 Individual differences in control performance	69
4.4 Summary and conclusions	76
5. Study 1: The Effect of Structural Information on the Control of Dynamic Systems	
5.1 Introduction	79
5.2 Method	84
5.3 Results	90
5.4 Discussion	98
6. Study 2: You Need to Know: There IS a Causal Relationship Between Structural Knowledge and Control Performance in Complex Problem Solving Tasks	
6.1 Permissions for published work	102
6.2 Introduction	103
6.3 Method	107
6.4 Results	113
6.5 Discussion	116
7. Study 3: The Effect of System Complexity on the Acquisition of Structural Knowledge and Controlling the Outcomes of a Dynamic System	
7.1 Introduction	120
7.2 Method	137
7.3 Results	140
7.4 Discussion	162
8. Conclusions and Implications	
8.1 Introduction	169

8.2 How does structural knowledge influence the control of dynamic systems?	169
8.3 What differentiates successful from less successful problem solvers in the control of dynamic systems?	173
8.4 Implications for the use of complex problems as tools for the assessment of problem solving competencies	176
8.5 Directions for future research	179
8.6 Conclusion	180
9. References	182
APPENDIX A: Instructions in the complete information condition	210
APPENDIX B: Instructions in the partial information condition	221
APPENDIX C: Instructions in the no information condition	232
APPENDIX D: The conceptualisation of task complexity	240
APPENDIX E: Equations for the systems used in Study 3	259

CHAPTER ONE

AIMS AND OVERVIEW

1.1 Introduction

Many situations require us (as human problem solvers) to interact with, and learn how to control, dynamic systems of causally connected variables. Learning how to heat food in a microwave, respond to emails and buy train tickets are just a few of the many examples that might be encountered in everyday life. On a more complex scale, the successful management of stock levels in a factory, hospital admissions, military operations and inflation in economies, requires problem solvers to manipulate and control the outcomes of highly inter-connected systems of variables. This thesis will examine the conditions that are required to learn how to effectively control the outcomes of dynamic systems and the cognitive processes that separate successful from less successful problem solvers in the performance of this task.

Since the 1970s a myriad of computer-based problems that represent dynamic systems have been constructed to investigate how people behave in complex environments. They are referred to as dynamic decision-making tasks (e.g. Brehmer & Allard, 1991; Gary & Wood, 2007), micro-worlds (e.g. Brehmer & Dörner, 1993; Kluge, 2008; Cañas & Waern, 2005, Funke & Frensch, 2007), simulations (Bühner, Kröner, Ziegler, 2008; Wood, Beckmann & Birney, 2009), computer-simulated scenarios (e.g. Funke, 2001), process-control tasks (e.g. Broadbent, 1977), complex dynamic control tasks (e.g. Osman, 2010), finite state automata (e.g. Buchner & Funke, 1993; Buchner, Funke & Berry, 1995) and complex problems (e.g. Beckmann & Guthke, 1995). In this thesis, for the sake of simplicity they will all be referred to as complex problems.

Complex problems consist of a number of inputs (variables that the problem solver can change) and outputs (outcomes that are generated by the system). The problem solver can change the values of the inputs, which affects the values of the outputs via causal structures that can be described algorithmically. The systems are considered to be “dynamic” because the values of the outputs change in response to the problem solvers’ actions, as well as independently over time (Edwards, 1962;

Brehmer, 1992; Funke, 1992). Typically, problem solvers are required to discover the relationships between the variables in the system; this is referred to as their structural knowledge. Concurrently or subsequently, they are also required to manipulate the inputs to reach certain goal values for the outputs; this is referred to as the quality of their system control, or control performance.

In educational and organisational settings, it has become common practice to use complex problems for assessment and selection purposes (Funke, 1998; U. Funke, 1998; Hornke & Kersting, 2005; Kluge, 2008). For example, in 1999 the Program for International Student Assessment (PISA) used the complex problem RAUMFAHRT (a virtual space shuttle) to assess the capacities to acquire knowledge and learn to control dynamic systems (Wirth & Funke, 2005). A set of similar complex problems will be included in PISA 2012 (The MicroDYN Approach; Greiff & Funke, 2008; 2009). The results of PISA are used to compare the knowledge and skills of students from different educational systems to inform the educational policy of the 60 participating countries (*OECD Program for International Student Assessment*). In organisational settings, complex problems are frequently used to aid the selection of personnel for roles in management, complex machinery operation and industrial research (Funke, 1998; U. Funke, 1998). For instance, in Germany, they are often preferred for the purpose of selection over traditional measures of intelligence (U. Funke, 1998; Kluge, 2008). The widespread adoption of complex problems in assessment and selection settings indicates that the capacity to control dynamic systems is seen as a key factor in job and educational success.

A large number of complex problems have also been developed for the purposes of education and training. Over the last ten years, the use of complex problems that represent scientific principles has become prevalent in primary and secondary education, although they are not typically considered to be a replacement for mainstream teaching and learning practices (de Freitas & Oliver, 2006; de Jong & Van Joolingen, 1998; Hulshof & De Jong, 2006; Goldstone & Sakamoto, 2003). Similarly, complex problems that represent systems in business are used pervasively in MBA programs and for personnel training (Wolfe & Rogé, 1997; U. Funke, 1998; Wood, Beckmann & Birney, 2009; Stainton, Johnson & Borodzicz, 2010). Proponents of this approach to teaching argue that complex problems are advantageous for training and education because they: a) have high face validity, b)

provide a new way of communicating with the “net generation”, c) are more engaging than traditional methods of teaching, d) allow for experimentation in a “low-risk” environment, e) reduce the time needed to expose learners to a wide-variety of situations and f) are adaptable to different training objectives (Wolfe & Rogé, 1997; U. Funke, 1998; Goldstone & Sakamoto, 2003; Hornke & Kersting, 2005; de Freitas & Oliver, 2006; Wood, Beckmann & Birney, 2009; Stainton, Johnson & Borodzicz, 2010). Nevertheless, these assertions are usually tempered by the acknowledgement that there is dearth of empirical evidence to support the claim that the use of complex problems benefits traditional learning outcomes (e.g. the transfer of skills and knowledge to real world settings).

In assessment and training contexts, there is a long held assumption that the problem solvers’ capacity to effectively control the outcomes of dynamic systems depends on their knowledge of the underlying structure of the system. Consequently, differences in control performance scores are interpreted as evidence of individual differences in the capacity to acquire and utilise structural knowledge in dynamic environments (Hornke & Kersting, 2005; Kluge, 2008; Greiff & Funke, 2008; 2009). Similarly, interventions designed to improve the control of dynamic systems often encourage learners to map their mental models, challenge their mental models or improve their mental models of the underlying structure of the system in question (Serman, 1994; Wolfe & Rogé, 1997; U. Funke, 1998; Gonzalez, 1999; de Freitas & Oliver, 2006). In both contexts poor control performance is typically interpreted as evidence that test takers or learners do not possess sufficient knowledge of the underlying structure of the system.

Such an assumption might seem sensible, considering the well-established link between expertise and domain knowledge in naturalistic settings, such as chess and physics. In comparison to novices, experts are characterised by a high level of successful performance in a particular domain. Concomitantly, they possess a large amount of highly structured domain knowledge and tend to focus on the deep structure of problems rather than on their surface features (de Groot, 1978; Chase & Simon, 1973; Chi, Glaser & Rees, 1982; Ericsson & Charness, 1994; Ericsson, Prietula & Cokely, 2007). This suggests that the superior performance of experts is crucially dependent on their abstract domain knowledge (Holyoak, 1991; Ericsson, 2003).

However, the results of studies that have investigated how people learn to control dynamic systems do not clearly support a causal relationship between the acquisition of structural knowledge and successful system control. On the one hand, a number of studies have shown strong associations between the amount of structural knowledge acquired by problem solvers and their capacity to control the outcomes of dynamic systems (Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer, Burns & Holyoak, 1996; Kröner, Plass & Leutner, 2005; Burns & Vollmeyer, 2002; Osman, 2008). On the other hand, a number of studies have demonstrated that problem solvers can learn to control the outcomes of dynamic system seemingly in the absence of structural knowledge (Broadbent, 1977; Broadbent, Fitzgerald & Broadbent, 1986; Berry & Broadbent, 1984; Berry, 1984; Berry, 1991; Stanley, Matthews, Buss & Kotler-Cope, 1989; Marescaux, Luc & Karnas, 1989; Dienes & Fahey, 1995). Most importantly, the provision of structural information appears to have no significant effect of the quality of system control (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996). This suggests that the efficient control of dynamic systems may not depend on the acquisition of structural knowledge (these findings are discussed in detail in Chapter 3).

This possibility has been widely discussed, and a number of alternative views on how people learn to control dynamic systems have been put forward. One perspective is that dynamic systems can be controlled through trial-and-error (Putz-Osterloh, 1993). Another is that problem solvers require a period of practice at controlling the system before they can effectively utilise structural knowledge (Preußler, 1998; Schoppek, 1998; 2002; 2004). Still others put forward the case that problem solvers must acquire knowledge of specific actions and their outcomes through practice, and that the acquisition of structural knowledge has no bearing on the success of this endeavour (Broadbent, Fitzgerald & Broadbent, 1986; Marescaux, Luc & Karnas, 1989; Dienes & Fahey, 1993; Kluwe, Haider & Misiak, 1990; Kluwe, 1993, 1995; Berry & Broadbent, 1995; Gonzalez, Lerch & Lebiere, 2003). Obviously, these arguments are inconsistent with each other and at the present time there is no clear evidence that favours any particular argument.

Clearly, an accurate model of how people learn to control dynamic systems is critical for the valid use of complex problems in applied settings. Currently, the cognitive processes that differentiate successful from less successful problem solvers

in the control of dynamic systems remain unspecified (Kluwe, Misiak & Haider, 1991; Funke, 1992; U. Funke, 1998). Hence, complex problem solving can only be defined as whatever the complex problem solving tests measure (to appropriate Boring's (1923) famous answer to the question of "what is intelligence?"). As a result, we do not know how best to train people to control dynamic systems. Consequently, the development of training programs can only progress through trial-and-error (Gonzalez, 1999). This state of affairs is clearly incongruent with an evidence-based approach to assessment and training.

Therefore, the aim of this thesis is to investigate the conditions that are required to learn how to effectively control the outcomes of dynamic systems, with a particular focus on the role of structural knowledge. The following questions will be addressed:

- Do problem solvers need structural knowledge in order to control dynamic systems effectively?

If so,

- How must this knowledge be acquired?
- Is there is a systematic relationship between the amount of structural knowledge that is acquired and the quality of system control?
- Are there systematic sources of individual differences that might explain the capacity to acquire and utilise knowledge?
- Does the complexity of the underlying structure of the system influence the relationship between structural knowledge and the quality of system control?

The aim of the following sections is to provide a guide for the reader of this thesis. The first section will define the key terms that are used throughout this thesis. The second section will outline the importance of a combined experimental and differential approach to the study of system control, which is the basic methodological approach adopted in the empirical work reported in this thesis. The final section will give an overview of the subsequent chapters.

1.2 Key terms

1.2.1 Tasks and Problems

Firstly, in order to avoid confusion, it is necessary to define what is meant by the terms “task” and “problem”. The use of cognitive tasks is central to the study of human behaviour in psychology in general, and problem solving in particular (Hackman, 1969; Wood, 1986). A task has a recognisable beginning and an end, and contains a certain set of stimuli and guidelines concerning the goals to be achieved (Hackman, 1969). Clearly, problems have these characteristics too; yet problems are usually seen as distinct from cognitive tasks more generally.

A commonly used distinction between a “task” and a “problem” is whether the solution to the desired goal state is known from the outset of performance. For example, Frensch and Funke (1995) review a number of alternative definitions of the term “problem”, and conclude that a problem can be said to exist when an individual does not immediately know how to address the difference between the current state of affairs and the desired state of affairs (the goal state). This definition can be traced as far back to Duncker (1945), who writes: “A problem arises when a living creature has a goal but does not know how this goal is to be reached” (p.1). Thus, in line with this definition, whether a situation constitutes a “task” or a “problem” can only be determined with reference to the knowledge of the particular individual who is required to achieve the desired goal state.

This definition is problematic as it implies that individual and situational characteristics determine whether a situation is to be defined as a “task” or a “problem”. A problem for one person may be a task for another, depending on their level of prior experience. Conversely, a problem may be transformed into a task by providing the individual with the relevant solution to apply. In turn, this suggests that a problem might become a task once an individual has identified the correct solution. In effect, this excludes solution application from the collection of processes that can be considered as part of problem solving. As such, this distinction between “task” and “problem” is insufficient for the purposes of this thesis.

In order to overcome these issues, the terms “problem” and “task” are used more concretely in this thesis. The term “problem” is used to encompass the entire

situation that the individual encounters within the experimental setting. The term “task” is used to refer to an overarching goal that is imposed by the instructions given in the situation. For example, in the context of complex problems, most often the tasks to be performed by the problem solver are “acquire knowledge about the relationships between the variables in the system” and “control the outcomes of the system”. Thus, the task to be performed constitutes one aspect of the overall problem encountered by the individual.

1.2.2 Complex problems

Three major research approaches have been developed that use complex problems to investigate how people learn to control dynamic systems: Dynamic Decision Making (e.g. Dörner, 1975, Brehmer, 1987; Serman, 1994; Busemeyer, 2002; Omodei & Wearing, 1995), Implicit Learning (e.g. Broadbent, 1977; Berry & Broadbent, 1984) and the Linear Structural Equation approach (e.g. Funke, 1985; 1993; 2001; Beckmann, 1994). Each approach entails the use of complex problems that have specific structural, surface and task characteristics. As Chapter 2 presents a comprehensive comparison of the approaches in these terms, the following sections will broadly describe each of these characteristics. This will also serve to highlight the heterogeneous nature of complex problems more generally.

1.2.2.1 Structural characteristics

Structural characteristics describe the properties of the system that is represented in the complex problem. The design of the structure of a system may be guided by a formal framework which specifies the characteristics that the system must contain (e.g. Funke, 1985; 1993; 2001) or informed by data and assumptions about a particular system in the real world (Goosen, Jensen & Wells, 2001; Stainton et al., 2010). Quesada, Kintsch and Gomez (2005) have suggested that the structural characteristics of complex problems can be further described in terms of time-related, variable-related and system-behaviour-related elements:

Time-related elements include whether the system is event-driven or clock-driven, and the degree of time pressure. Event-driven systems change in discrete steps triggered by problem solvers’ actions, while clock driven systems change in response to an internal clock. The degree of time pressure that problem solvers may

be exposed to varies considerably across different systems (Quesada et al., 2005). In some systems, the problem solver may take as much time as they need to make a decision within the reasonable constraints of the experimental session (e.g. CHERRY TREE; Beckmann, 1994). In other systems, such as FIRECHIEF (Omodei & Wearing, 1995), the time pressure is intense and the problem solver must make quick decisions.

Variable-related elements include the type and number of variables, the number of relationships between the variables and whether these relationships are linear or non-linear (Quesada et al., 2005). All systems include input and output variables, but mediating and moderating variables may also be present which cannot be directly observed (Wood, Beckmann & Birney, 2009). The number of variables present in different structures varies considerably; at the lower end of the spectrum CITY TRANSPORTATION (Broadbent, 1977) consists of a single input and a single output variable, while at the extreme upper end of the spectrum LOHHAUSEN (Dörner, Kreuzig, Reither & Stäudel, 1983) consists of over 2000 variables. The number of variables is not indicative of the number of relations in a system, as some structures contain many variables with few connections, while others contain few variables with highly inter-connected structures (Blech & Funke, 2005). While many systems contain purely linear relationships between the variables, DURESS (Christoffersen, Hunter & Vicente, 1996), FIRECHIEF (Omodei & Wearing 1995), LOHHAUSEN (Dörner et al., 1983) and MORO (Dörner, Stäudel & Strohschneider, 1986) contain quadratic and exponential relationships. The effect of the number of variables and relations on system control, and their relationship to the complexity of the system, is discussed in Chapter 7.

System-behaviour-related elements include whether the system is opaque or transparent, stochastic or deterministic, and whether the feedback is immediate or delayed (Quesada et al., 2005). In opaque systems, the underlying structure of the system cannot be fully determined by the problem solver as the system contains layers of hidden variables. In transparent systems, all the variables are displayed in the user interface (Funke, 1991; Quesada et al., 2005). In the literature, transparency also sometimes means that problem solvers are informed as to the underlying structure of the system prior to the instruction to control the system (Putz-Osterloh &

Luer, 1981; Funke, 1991; Putz-Osterloh, 1993; Latzina, 1990). However, in this thesis the term transparency is used strictly in the former sense.

Stochastic systems contain random components in their underlying structures (Quesada et al., 2005). For example, in Broadbent and Berry's (1984) SUGAR FACTORY problem, which consists of a single input and output variable, ± 1000 is added on a pseudo-random basis to the value of the output variable on two-thirds of the trials. This means that no unique output value is associated with any one starting state and input value. In comparison, in deterministic systems, the same starting state and action will always produce the same outcome.

Problem solvers receive feedback about the effect of their actions on the system via the value of the output variables. In systems where the feedback is immediate, the effect of the problem solvers' actions on the outputs is evident as soon as the action is executed. In systems where the feedback is delayed, the effect of the problem solvers' actions may only become apparent after a certain number of trials have passed or a set period of time has elapsed. The presence of feedback delays makes it more difficult to connect specific actions to their outcomes, and distinguish between autonomous changes in the output variables and changes that are the result of direct interventions on the system variables (Brehmer & Allard, 1991; Brehmer, 1995).

1.2.2.2 Surface characteristics

The surface of a complex problem is the user interface that the problem solver interacts with (Wood, Beckmann & Birney, 2009). The most widely discussed surface characteristic is the type of cover story and labels given to the system variables. A distinction is often made between "domain-independent" or "abstract", and "semantically meaningful" or "concrete" cover stories and variable labels (e.g. Hesse, 1982; Beckmann, 1994; Beckmann & Guthke, 1995; Goldstone & Sakamoto, 2003; Lazonder, Wilhelm & Hagemans, 2008; Bühner, Kröner & Ziegler, 2008, Lazonder, Wilhelm & Van Lieburg, 2009). Semantically meaningful or concrete cover stories and variable labels refer to familiar systems in the real world. For example, in LOHHAUSEN the problem solver is instructed to act as the mayor of a virtual small town dealing with variables labelled "living standard of the workforce" and "energy consumption" (Dörner, 1987), while in FIRECHIEF individuals are

required to control variables labelled as “helicopters” or “trucks” to stop simulated forest fires spreading (Omodei & Wearing, 1995). Domain-independent or abstract cover stories and variable labels do not refer to any known or previously experienced system (e.g. COLORSIM, Kluge, 2008; MULTIFLUX, Kröner, 2001). For example, in MULTIFLUX problem solvers are instructed to work out how to control a fictitious machine that consists of four control devices labelled “A”, “B”, “C” and “D” and four instruments labelled “1”, “2”, “3” and “4”.

It could be argued that the important difference between problems such as LOHHAUSEN and MULTIFLUX is not in their “concreteness” or “abstractness”. Rather, the important difference is whether the cover story and variable labels seem familiar or novel to the problem solver, and thus whether prior experience may be useful (Goode & Beckmann, under revision). Therefore, in this thesis, rather than classifying the surface features of complex problems in terms of whether they are “abstract” or “concrete”, the terms “familiar” and “novel” will be used.

The surface of complex problems can also be described in terms of whether the user interface is graphical or numerical, or a combination of the two. In graphical user interfaces, the values of the system variables are represented as lines or bars on graphs. On each trial, problem solvers are able to set the input variables at positive, negative or constant values. The graphs that display the value of the output variables are then updated to show whether each variable has increased, decreased or remains unchanged (e.g. MACHINE, Beckmann, 1994; MULTIFLUX, Kröner, 2001). In contrast, numerical user interfaces display the exact values of the variables in a series of tables. On each trial, problem solvers are able to enter a number for each input variable that is typically constrained within a certain range. The table that displays the output variables is then updated to show the resulting numerical values (e.g. MICRODYN, Greiff & Funke, 2009). In some complex problems, such as COLORSIM (Kluge, 2008), the values of the input and output variables are represented in both graphs and tables.

1.2.2.3 Task characteristics

Task characteristics describe the goals that are imposed by the instructions given in the experimental setting. Knowledge-related goals direct the problem solver to acquire knowledge about the underlying structure of the system. Control-related

goals direct the problem solver to bring the system into a certain state. Each of these goals defines a task to be performed in relation to the system.

With regard to control-related goals, a further distinction is often made between well- and ill-defined goals. Well-defined goals instruct the problem solver to achieve specific values for the output variables. For example, in MACHINE (Beckmann 1994) problem solvers are instructed to match the value of the outputs with lines indicated on the output graphs over a series of seven trials. In contrast, ill-defined goals require the problem solver to interpret the instructions in order to formulate their own specific goals. For example, in MORO (Dörner, Stäudel & Strohschneider, 1986) problem solvers are instructed to improve the living conditions of a fictitious African tribe by manipulating variables such as pastures, cattle, fertiliser and housing. Problem solvers have to decide which of these variables are relevant to the living conditions of the fictitious tribe and how they should be altered. The consequences of well- and ill-defined goals for the interpretation of control performance scores are discussed in Chapter 2.

Another particularly critical task characteristic is whether problem solvers are instructed to simultaneously acquire knowledge and control the system, or whether these tasks are separated experimentally. In many studies, problem solvers are instructed to perform both tasks simultaneously (e.g. Berry & Broadbent, 1984; 1989; Sanderson, 1989; Rigas, 2000; Rigas, Carling & Brehmer, 2002; Süß, 1999; Wittmann & Süß, 1999; Wittmann & Hatstrup, 2004). However, there is much debate over whether it is actually possible for problem solvers to perform both of these tasks concurrently because the optimal strategy required to discover the relationships between variables in a system often differs from the one that is required to control its outcomes (Vollmeyer, Burns & Holyoak, 1996; Burns & Vollmeyer, 2002). This issue is discussed in detail in Chapters 2 and 3.

1.3 A combined experimental and differential approach

Complex problems have been studied from two different research traditions: The differential and the experimental. Differential research in this area has largely been occupied with the question of whether (or not) individual differences in system control can be explained by traditional measures of intelligence (e.g. Dörner, Kreuzig, Reither & Stäudel, 1983; Gediga, Schottke & Tuck-Bressler, 1984; Putz-

Osterloh , 1981; Putz-Osterloh & Lürer, 1981; Reichert & Dörner, 1988; Joslyn & Hunt, 1998; Kröner, 2001; Kröner & Leutner, 2002; Kröner, Plass & Leutner, 2005; Bühner, Kröner & Ziegler, 2008). As yet, findings are rather inconsistent, and this is likely because these studies have to a certain extent ignored the effect of structural, surface and task characteristics on performance. In contrast, while experimental research in this area has made significant progress towards documenting the structural, surface and task characteristics that influence the acquisition of structural knowledge and control performance on average, this research has treated individual differences between subjects as error variance (e.g. Berry & Broadbent, 1984; Funke, 1992; Kluge, 2008). As a result, our ability to explain and predict the performance of individuals on complex problems that have specific characteristics is severely limited.

In this thesis we will attempt to address these limitations through the adoption of a combined experimental and differential approach, as advocated by Cronbach (1957; 1975). In particular, the amount of structural information available to problem solvers and the complexity of the underlying structure of systems will be manipulated experimentally. Process indicators and traditional measures of intelligence will be used to account for individual differences in the effect of these experimental manipulations on the acquisition of structural knowledge and control performance. This approach should allow us to identify whether the control of dynamic systems is causally dependent on the acquisition of structural knowledge, and explain why individuals differ in their capacity to acquire and utilise such knowledge.

1.4 Overview of the thesis

The initial literature review is presented in three short chapters. Overall, the purpose of Chapters 2, 3 and 4 is to identify inconsistencies and gaps in our knowledge of how people control dynamic systems, which then provides the basis for the experimental studies reported in Chapters 5, 6, and 7.

Chapter 2 describes and critically examines the three major approaches that have investigated how people learn to control dynamic systems using complex problems (i.e. Dynamic Decision Making, Implicit Learning and Linear Structural Equations). The purpose of this chapter is to establish a context for the current

research and provide a justification for the selection of the methodological approach adopted in this thesis.

Based on this analysis, Chapter 3 reviews and integrates the experimental findings from the different approaches with regard to the role of structural knowledge in the control of dynamic systems. The purpose of this section is to identify the structural, surface and task characteristics that might influence whether structural knowledge can be effectively acquired and utilised to control dynamic systems, and how this process may take place.

Chapter 4 explicates a model of successful system control based on the findings presented in Chapter 3, and discusses possible sources of individual differences in the tasks of acquiring structural knowledge and controlling the outcomes of dynamic systems.

Chapter 5 reports the data from a study ($N = 91$) that investigated the role of structural information and fluid intelligence in controlling a dynamic system. The aim of this study was to determine whether problem solvers need to directly interact with a dynamic system in order to acquire structural knowledge that can be effectively utilised to control its outcomes, and to identify sources of individual differences in the acquisition and application of knowledge.

Chapter 6 consists of a paper that was published in the journal *Intelligence* (Goode & Beckmann, 2010). This study ($N = 75$) investigated whether there is a causal relationship between the acquisition of structural knowledge and the efficiency of system control. The aim of this study was to establish whether structural knowledge is a necessary component of successful performance, and the extent to which the utilisation of knowledge depends on the problem solvers' level of fluid intelligence.

Chapter 7 reports the data from a study ($N = 293$) that represents a synthesis and expansion of the work reported in Chapters 5 and 6. The aim of this study was to determine whether the results reported in Chapters 5 and 6 generalise to systems of differing complexity.

Chapter 8 integrates the findings presented in Chapters 5, 6 and 7, in order to propose a model of successful system control. We consider the implications of the

current work for the use of complex problems for the purposes of assessment and training.

CHAPTER TWO

COMPLEX PROBLEM SOLVING RESEARCH APPROACHES: A CRITICAL EVALUATION

2.1 Introduction

A recurrent tension exists in psychological research between a desire to maintain both experimental control and ecological validity (Neisser, 1976; Brehmer, 1992; Brehmer & Dörner, 1993; DiFonzo, Hantula & Bordia, 1998; Gray, 2002). For example, Brehmer and Dörner (1993) express their concern that “in field research there is often too much [complexity] to allow for any more *definite* conclusions, and in laboratory research, there is usually too little complexity to allow for any *interesting* conclusions” (p.172). On the one hand, this reflects an on-going concern as to whether laboratory tasks adequately represent the complex demands of real-life situations, and thus whether it is appropriate to generalise the theoretical concepts derived from such tasks to real life performances. On the other hand, there is also a concern that real life situations are far too complex to allow conclusive insights to be drawn about the cognitive processes that underlie performance.

With this dilemma in mind, in the late 1970s researchers criticised traditional problem solving tasks, such as the Tower of Hanoi, for being far too simple, too transparent and static (Dörner, 1975; Broadbent, 1977). It was argued that in comparison, real-life tasks such as flying an aircraft, driving a car or managing an economy, were complex, opaque and dynamic (Buchner, 1995; Wenke, Frensch & Funke, 2004). In response, three major experimental approaches developed to examine how people behave in, and learn to control, complex and dynamic environments: Dynamic Decision Making (e.g. Dörner, 1975, Brehmer, 1987; Serman, 1994; Busemeyer, 2002), Implicit Learning (e.g. Broadbent, 1977; Berry & Broadbent, 1984) and the Linear Structural Equation approach (e.g. Funke, 1985; 1993; 2001; Beckmann, 1994; Vollmeyer, Burns & Holyoak, 1996).¹

Broadly, all researchers working within the constraints of these approaches are concerned with how people learn to control dynamic systems of variables. However,

¹ Computer-based simulations of complex and dynamic systems are also used to investigate discovery learning in conceptual domains. The approach adopted in these studies is similar to the dynamic decision making approach.

researchers from different approaches have addressed markedly different aspects of this question. Researchers who have used the Dynamic Decision Making (DDM) approach have largely been concerned with why people perform poorly in complex and dynamic environments, and whether or not traditional measures of intelligence reflect the same demands as those required to control dynamic systems (e.g. Dörner, 1975, Brehmer, 1987; Sterman, 1994; Busemeyer, 2002). In this research, the effects of structural, surface and task characteristics on performance have largely been ignored. In comparison, researchers who have utilised the Linear Structural Equation approach have typically examined the effect of various problem characteristics on acquiring knowledge about, and controlling, dynamic systems (see Funke, 1991; 1993; 2001). A limited number of researchers have used this approach to examine the role of individual differences in performance. Finally, researchers who have used the Implicit Learning approach have drawn distinctions between different modes of learning that are possible in dynamic environments (e.g. Broadbent, 1977; Berry & Broadbent, 1984; 1987). However, an integrated theory of how people learn to control dynamic systems needs to encompass all of these different perspectives: The cognitive mechanisms that operate during performance, individual differences, as well as structural, surface and task characteristics.

Therefore, the aims of this chapter are to a) determine whether the research findings from the Dynamic Decision making (DDM), Implicit Learning and Linear Structural Equation approaches can be integrated into a holistic model of how people control dynamic systems and b) determine which approach provides the most methodologically sound framework for the study of the processes that underlie system control. Firstly, common elements across the approaches will be discussed. Secondly, each approach will be described and critically evaluated in terms of the type of systems they employ, whether they present complex problems in a familiar or novel context, the tasks set for the problem solver, the operationalisation of dependent variables and the ecological validity of the approach.

2.2 Common features across the approaches

Each approach uses computer-based problems that represent dynamic systems of inter-connected variables. At a mere descriptive level, these problems consist of a number of inputs (variables that the problem solver intervenes on) and outputs

(outcomes that are generated by the system) that are represented in a computer program. The values of inputs can be changed, which affects the values of the outputs via more or less complex causal structures that relate the inputs to the outputs, and which can be described algorithmically. This is referred to as the underlying structure of the system. Systems are considered to be “dynamic” when the values of the outputs change in response to problem solvers’ actions and independently over time (Edwards, 1962; Brehmer, 1992; Funke, 1992). Typically, problem solvers are allowed to interact with the system over a set period of time or for a set number of trials, where a trial consists of setting the value of the input variables and viewing the subsequent effect on the output variables.

In order to “solve” the problem, problem solvers must learn how to manipulate the values of the input variables in order to control the values of the output variables to reach and maintain desired goal states. This is referred to as control performance. Efficient control performance is analogous to identifying the “correct” solution in a traditional laboratory problem-solving task, such as an intelligence test item. However, unlike traditional problems, the application of a solution requires the problem solver to constantly intervene on the system variables to maintain the goal output variable states over a period of time. Therefore, the efficiency of control performance is typically evaluated over a set period of time during which the problem solver continuously interacts with the system to reach and maintain the desired goal states.

In some studies, performance is also evaluated in terms of the amount of structural knowledge that problem solvers acquire while they interact with the system. Structural knowledge refers to verbalisable knowledge about the underlying structure of the system. This knowledge may be represented at different levels of precision. At the broadest level, only the existence of a relationship between two variables is known; this is referred to as qualitative structural knowledge. At a more specific level, the direction of the relationship is known, and optimally, the strength or numerical weight of the relationship can also be specified; this is referred to as quantitative structural knowledge (Funke, 1992; van Joolingen & de Jong, 1997; Kluge, 2008; Kröner, Leutner & Plass, 2005).

The distinction between structural knowledge and control performance is similar to that found in the cognitive skill acquisition literature between declarative and procedural knowledge, respectively. Declarative knowledge is general factual knowledge about a task, while procedural knowledge is how to perform a particular task (Anderson, 1983; 1993). In the case of complex problems, declarative knowledge entails knowing that one variable influences another (e.g. that variable X has a strong negative impact on variable Y) whilst procedural knowledge entails knowing how to achieve a particular result (e.g. knowing how to increase variable X by 3 points to decrease variable Y by 8 points).

2.3 The Dynamic Decision Making approach

2.3.1 Structural characteristics

In this approach, complex problems are referred to as dynamic decision-making tasks, micro-worlds or simulations. The systems that they represent share a number of key characteristics. In particular, they contain a large number of highly interconnected variables (between 10 and 2000), many of the relationships between the variables cannot be directly observed, and may be non-linear, and decisions must be made in real time. It is often argued that these characteristics mirror those of complex, dynamic and uncertain environments in the real world (Brehmer, 1992; Brehmer & Dörner, 1993; Kerstholt & Raaijmakers, 1997; DiFonzo, Hantula, & Bordia, 1998, Gray, 2002; Waern & Cañas, 2003; Cañas & Waern, 2005; Brehmer, 2005).

Two frequently used complex problems in this approach are “MORO” and “TAILORSHOP”. In “MORO” problem solvers must advise a tribe in Africa on how to improve their living conditions (Dörner, Stäudel & Strohschneider, 1986). The system contains 49 highly inter-connected variables, such as pastures, cattle, fertiliser, population and housing. “TAILORSHOP” requires problem solvers to manage the supply chain of a virtual shirt-making factory. The system contains 24 variables, such as raw materials, number of machines and workers. The goal is to maximise the profits of the company over a 12-month period (Putz-Osterloh, 1981). What makes these problems particularly challenging to solve is that the systems are opaque. That is, many of the effects cannot be directly observed because the variables change as a result of complex causal sequences of effects.

Although the complex problems used in this approach share some common structural characteristics, they are far from homogenous. This is problematic because a range of system properties, such as the number of variables (Preußler, 1997), the number of relations (Kluge, 2008; Funke, 1992), the type of relations (Funke, 1992), feedback delays (Sterman, 1989; Brehmer & Allard, 1991), the consistency of the relationships between the inputs and the outputs (Ackerman & Cianciolo, 2002), the type of connectivity pattern (Howie & Vicente, 1998) and whether the relationships in the system are linear or non-linear (Dörner, 1989; Dörner & Scholkopf, 1991) have been shown to have significant impacts on the difficulty of knowledge acquisition and control performance. This state of affairs has led many to argue that a comparison of the experimental findings from studies that use different complex problems is almost impossible (Buchner, 1995; Funke, 1992; Brehmer, 2005; Kerstholt & Raaijmakers, 1997; Buchner & Funke, 1993; Diehl & Sterman, 1995; Mackinnon & Wearing, 1985).

In order to compare the experimental results from different complex problems, Quesada et al. (2005) suggest that system structures should be formally described. In an effort towards the construction of a taxonomy of system characteristics they describe ten different dimensions on which problems used in the DDM literature differ, including the number of variables, the type of variables (continuous or discrete), whether there are feedback delays and whether the system is stochastic or deterministic (see also Gonzalez, Vanyukov & Martin, 2005 for a similar taxonomy). However, it could be argued that this does not resolve the problem, as Quesada et al. (2005) provide no guidelines as to how similar systems should be to yield comparable experimental results.

2.3.2 Type of cover story

As the explicitly stated aim of the DDM approach is to “bring the field into the laboratory” (Brehmer, 2005, p.75) researchers use cover stories that refer to familiar situations in the real world. Many cover stories have been used, including managing an economy (Sterman, 1989), acting as the mayor of a town (Dörner, 1987), directing a fire fighting unit (Brehmer, 1990; 1992) and diagnosing medical problems (Kleinmuntz & Thomas, 1987). The use of familiar cover stories is

designed to encourage problem solvers to bring their prior experience to bear in laboratory tasks, as they would in real life (Brehmer & Dörner, 1993).

This approach has been frequently criticised precisely because it encourages subjects to form assumptions about the underlying structure of the system, and often no attempt is made to ensure that the underlying structure of the system reflects the structure of the real life situation that it is supposed to represent (Buchner, 1995; Funke, 1992; Beckmann & Guthke, 1995). This is problematic because some findings suggest that the acquisition of structural knowledge is facilitated if prior experience is in concordance with the actual underlying structure of the system, but is hindered if it is in conflict (Reither, 1981; Hesse, 1982; Lazonder, Wilhelm, & Hagemans, 2008; Lazonder, Wilhelm & Van Lieburg, 2009). These results suggest that assumptions evoked by the cover story of a problem may confound the interpretation of other experimental manipulations.

Even if experimenters do try to explicitly replicate the structure of real world situations, there is no reason to assume that it is aligned with a particular individual's prior experience. This is clearly demonstrated in the results of Lazonder et al.'s (2008) study, in which the underlying structure of the system was explicitly designed to reflect commonly known relationships in everyday life. The task required subjects to investigate how training frequency, smoking and nutrition affected the time it took an athlete to run 10km. Increased training frequency, eating "sport food" and not smoking decreased the athlete's time, while eating "junk food" increased their time. These relations might seem rather obvious; however, a knowledge test revealed that two subjects (out of forty) had incorrect assumptions about these relationships. These findings suggest that even in carefully designed problems, individual differences in prior knowledge are likely to be a source of potential error variance. It is perhaps surprising then, that prior knowledge is rarely controlled for in experimental studies within the DDM approach (with the notable exceptions of Leutner, 2002; Wittmann & Hatrup, 2004).

2.3.3 Task characteristics

The key task characteristics utilised in the DDM approach are that: a) subjects must simultaneously try to acquire knowledge about the underlying structure of the system while they try to control it and b) the goal states are usually ill-defined. It has

been argued that these demands reflect reality because we (as human problem solvers) rarely have the opportunity to acquire knowledge about a system independently from our attempts to manipulate it and we are often required to decide which goals are relevant to a particular context (Dörner & Wearing, 1995; Rigas, Carling & Brehmer, 2002). However, these characteristics place a serious limit on the validity and reliability of the dependent variables that can be derived as performance measures, and make it difficult to determine which behaviours lead to successful performance.

The requirement that subjects simultaneously control and acquire knowledge about a system presents two main challenges to the assessment of performance. Firstly, the strategy that is necessary to acquire knowledge about a system may be incompatible with the one that is needed to control it. Thus, the problem solver who first undertakes the task to systematically discover the underlying structure of a system before they attempt to control it may appear to have worse control performance than the problem solver who attempts to incrementally manipulate the output variables towards the goal state using a trial-and-error strategy (Kerstholt & Raaijmakers, 1997). Secondly, if we assume that problem solvers do not initially know how to control the outcomes, then this implies that control performance should improve over the course of the task. This will limit the reliability of aggregated measures of control performance because performance at different points in time will reflect different levels of expertise. This is one possible reason as to why the reliability of control performance measures is often low (Funke, 1983; 1984; Kluwe, Misiak & Haider, 1991; Wenke & Frensch, 2003; Wenke, Frensch & Funke, 2004). This is problematic because any researcher using correlations must demonstrate that the variables of interest show sufficient reliability. Thus, the requirement to simultaneously acquire knowledge and control a system poses a significant threat to the validity and reliability of performance measures.

The use of ill-defined goals is also likely to pose a threat to the validity of control performance measures because problem solvers must formulate their own goals and these self-set goals may differ across individuals. For example, in LOUHHAUSEN (Dörner, Kreuzig, Reither & Stäudel, 1983), in which subjects act as the mayor of a virtual small town, the goal is to ensure that the residents are “satisfied with both the town and their own lives” (Dörner & Wearing, 1995, p.67).

Given such instructions, it is clear that different subjects may prioritise different sub-goals, according to their personal values. This makes it difficult to compare the performance of different subjects (Funke, 1993) and any operationalization of solution quality is somewhat arbitrary.

Self-set goals are also problematic because findings from verbal protocols suggest that subjects may not adhere to consistent goals for performance over the course of interacting with the system. Dörner (1980) reports that subjects often jump from one goal to the next, relatively quickly, without achieving any goal, and treating all goals superficially. In the most extreme situation, problem solvers may fail to set any concrete goals at all. Dörner (1980) sees this as a general problem that people confront when dealing with real-world tasks. Alternatively, however, it could be argued that it may reflect the fact that subjects are not sure what they are actually supposed to be doing.

More recently, the proponents of the DDM approach have recognised the problems with ill-defined goals, and have advocated in favour of the use of specific goals (e.g. Rigas, Carling & Brehmer, 2002). However, the complex nature of most of the systems utilised in the DDM approach means that there are multiple ways in which any particular goal might be achieved. For example, in TAILORSHOP (Putz-Osterloh, 1981) subjects are told to maximise the profits of a clothing factory. Subjects who invest heavily in equipment and supplies will produce greater profits in the later half of simulation than subjects who makes no investment in the company. On average, however, there will be no difference in their 12-month profit. Therefore, successful performance cannot be clearly specified on a behavioural level, as no optimal solution exists to the problem (Kluwe et al., 1991). This makes it difficult to distinguish the processes that differentiate successful from less successful problem solvers.

2.3.4 Operationalisation of dependent variables

Within the DDM approach, there is little consensus as how control performance should be operationalised, even within studies that use the same complex problem. For example, using the complex problem TAILORSHOP, Putz-Osterloh (1981) operationalised control performance in terms of the number of months with increased capital assets, while Barth and Funke (2010) calculated the

profits at the end of every month. Most significantly, findings show that when different dependent measures of performance are derived from the same task, different results are obtained with regard to the correlation between traditional measures of intelligence and control performance (i.e. TAILORSHOP e.g. Funke, 1983; Putz-Osterloh, 1981; Süß, Kersting & Oberauer, 1991 and MORO, e.g. Elg, 2005). This may suggest that different performance measures reflect different facets of control performance (Elg, 2005) or it may simply indicate that the measures differ in their reliability (Kluwe et al., 1991; Wenke & Frensch, 2003; Wenke et al., 2004).

Similarly, no general method has been developed to assess what problem solvers have learnt about the underlying structure of such systems. Typically, subjects are given a series of questions designed to assess their knowledge of how the system works (e.g. Wittmann & Hattrup, 2004; Elg, 2005). However, the format and content of these assessments (i.e. whether they assess quantitative, qualitative or both types of structural knowledge) is inconsistent. Obviously, this makes comparisons across studies problematic.

The majority of studies do not directly measure structural knowledge (e.g. Reither, 1981; Diehl & Serman, 1995; Rigas, Carling & Brehmer, 2002; Jobidon, Rousseau & Breton, 2005; Barth & Funke, 2010). In these studies, it is assumed, either explicitly or implicitly, that successful control performance indicates that problem solvers have acquired an accurate mental representation of the underlying structure. Hence, control performance is used as an indirect indicator of structural knowledge. The problem is that a causal relationship between structural knowledge and control performance has not yet been established, and indeed, the results of some studies suggest that structural knowledge and control performance may be unrelated (e.g. Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993). These findings suggest that structural knowledge needs to be assessed independently from control performance.

2.3.5 Ecological validity

Regardless of these problems, proponents of the DDM approach argue that the advantage of studying decision-making in ecologically valid environments far outweighs the disadvantages (Brehmer, 1992; Brehmer & Dörner, 1993; DiFonzo, Hantula, & Bordia, 1998, Gray, 2002; Waern & Cañas, 2003; Cañas & Waern, 2005; Brehmer, 2005). While this may be true, it is questionable whether the

complex problems used in many of the studies within the DDM approach are actually ecologically valid. Complex problems tend to be dramatic oversimplifications of the complex environments that they purport to represent. This is partly because the real world is far too complex to be represented via computer models and also because many of the relationships that govern real world situations are unknown or change over time (Keys & Biggs, 1990; Goosen, Jensen, & Wells, 2001; Gold, 2003). Given such constraints, designers of complex problems within the DDM approach are encouraged to aim for the perception of reality by the problem solver, rather than true fidelity to the real world (Norris, 1986; Brehmer & Dörner, 1993; Gold, 2003; Stanton, Johnson & Borodzicz, 2010).

Thus, the claim that any particular complex problem is ecological valid is usually based exclusively on surface characteristics (i.e. “face validity”). That is, the problem has the appearance of reflecting some situation in reality due to the cover story, and so it is assumed that the psychological demands of the task mirror reality (Beckmann & Guthke, 1995). One way of testing this assumption would be to compare the performance of experts from the relevant domain and novices on a particular task. We would expect that experts should perform better than novices in the real situation, and therefore should show better performance in the simulated situation. The results of two studies illustrate that there is no reason to assume that this is the case, based on the surface features of the complex problem.

Reither (1981) compared the performance of a group of “experts” and “novices” in a computer simulation of a fictitious tribal village in Africa. The experts had 6-8 years experience as aid workers in developing countries, while the novices were preparing for their first assignment. Against expectations, the novices were more successful in promoting the long-term growth of the village. Reither (1981) attributed the experts’ comparatively poor performance to the fact that they believed that they knew what was to be done before they fully understood the system. On the other hand, this also suggests that the task is not ecologically valid, as the behaviours that are successful in the real world situation do not apply to the simulated situation.

In a more recent study, Rolo and Diaz-Caberera (2005) designed a complex problem based on an analysis of a petrol refinery. They then compared performance in the field and experimental settings. In the field study verbal protocol analysis was

used to investigate expert performance in a petrol refinery. In the experimental study, it was found that similar decision sequences were observed in experts in the field and “good” performers in the laboratory. For example, when subjects have to cope with the systems instability, good performers and experts seek information about the variable states in order to diagnose the cause of the problem. The results of these studies illustrate that the surface characteristics of a complex problem provide no indication of its ecological validity, and that ecological validity needs to be demonstrated on a problem-by-problem basis.

2.3.6 Summary of the DDM approach

In summary, the DDM approach utilises problems that claim to simulate a particular situation in reality. They are embedded in a rich semantic network, have highly complex structures, and problems solvers are typically required to formulate, or at least further specify, their own goals, as well as acquire knowledge about the system while they control it. While the proponents of this approach argue that these demands reflect those of real life situations, these demands also result in a reduction of experimental control. The extent to which the complex problems used in this approach actually do mirror real world situations is also questionable. In addition, no standard method has been developed to assess problem solvers’ knowledge about the underlying structure of the system or control performance. These factors make it difficult to compare the results of studies even within this approach.

2.4 The Implicit Learning approach

2.4.1 Structural characteristics

At the opposite end of the spectrum, proponents of the Implicit Learning approach advocate the use of relatively simple and mathematically well-defined systems in order to isolate the underlying psychological processes involved in system control (Broadbent, 1977; Berry & Broadbent, 1984; 1988). The problems used are referred to as process-control (e.g. Broadbent, 1977), complex dynamic control tasks (e.g. Osman, 2008; 2010) or finite state automata (e.g. Buchner & Funke, 1993; Buchner, Funke & Berry, 1995). The problems contain between 2 and 4 variables that are governed by a set of linear equations, which typically include a random error factor. The tasks are usually, but not always, dynamic, so that the current output

value depends on the value of the input selected by the subject and the previous value of the output.

One of the earliest examples is the CITY TRANSPORTATION problem (Broadbent, 1977), in which subjects can control the number of free parking spaces and passenger load on a virtual bus system. This system consists of a two output variables that can be manipulated by changing the value of two input variables. The relationship between the variables is described by two simultaneous linear equations. The system is not dynamic, as the outputs depend only on the current values of the inputs.

Another example is the SUGAR PRODUCTION problem, (Berry & Broadbent, 1984), which contains only one input (the number of workers) and one output (sugar production). Subjects are told that they can control the rate of sugar production by manipulating the number of workers in a factory. The linear equation that describes the relationship between the variables is dynamic and includes a random error factor.

Most of the studies in this approach use the SUGAR FACTORY problem or structurally isomorphic problems with different cover stories (e.g. PERSON INTERACTION, PERSON CONTROL). On the one hand, this makes it easy to synthesize the research findings from this approach with regard to how people learn to control dynamic systems. The results are rather consistent across different studies, as they are not confounded by differences in system properties. On the other hand, the consistent use of the same system begs the question of whether these results are generalisable to different systems, or whether the results obtained are to some degree an artefact of the SUGAR FACTORY system. These issues are discussed in more detail in Chapter 3 when we present the main findings from this approach.

2.4.2 Type of cover story

As in the DDM approach, one criticism is that the problems are typically presented with familiar labels given to the system variables (with the notable exceptions of Buchner & Funke, 1993; Buchner, Funke & Berry, 1995). As previously discussed, this may have an impact on the acquisition of structural knowledge. Indeed, Sanderson and Vicente (1986) found that subjects confronted

with the CITY TRANSPORTATION problem typically started with a number of beliefs about how the system should work and that these beliefs incongruent with the actual underlying structure of the system. This problem has largely been overcome, however, as key findings have been replicated with structurally isomorphic versions of the SUGAR FACTORY problem with different cover stories (Berry & Broadbent, 1984; Berry & Broadbent, 1988).

2.4.3 Task characteristics

In comparison to the DDM approach, the Implicit Learning approach emphasises the use of well-defined goals. For example, in the SUGAR FACTORY problem subjects are instructed to "...reach and maintain a target output of 9,000 tons" (Berry & Broadbent, 1984, p. 212). As the system is mathematically tractable, it is possible to specify an optimal intervention given an arbitrary system state. Therefore, in comparison to the DDM approach, it is possible to validly compare different subjects' performance.

The main similarity to the DDM approach is, however, that subjects are typically required to learn about the underlying structure of the system while they control it. As previously discussed, this makes it difficult to separate the process of acquiring knowledge from controlling the system, and may limit the reliability of measures of control performance. Indeed, Gebauer and Mackintosh (2007) report that if participants are explicitly instructed to infer the rules underlying the SUGAR FACTORY task before they receive specific goals for system control, the reliability of control performance measures is somewhat higher than if they are instructed to infer the rules while they control the task (Spearman-Brown Split-half reliability .87 and .69, respectively).

2.4.4 Operationalisation of dependent variables

No standardised methods have been developed to assess either structural knowledge or control performance. With regard to control performance, multiple methods are often utilised even within the same study. For example, Berry and Broadbent (1987) use both the number of attempts that it takes to reach the goal state and the number of times that the subject moves away from, rather than towards, the goal, as measures of performance. Berry and Broadbent (1984) use the number of

trials on which the goal state is reached. Although the criterion for performance is similar, the use of different measures makes it difficult to directly compare the effect of different task characteristics and experimental manipulations on performance.

The number of trials on which the goal state is reached is the most commonly used indicator of control performance. However, this is particularly problematic as the number of times that any particular goal state can be reached is highly dependent on the characteristics of the underlying system structure. For example, Buchner and Funke (1993) constructed a “small” and a “large” version of a complex problem that were identical in their surface features but the underlying structure of the small problem was designed to be less complex than the large problem. In the small complex problem, given the actions that must be performed to achieve the goal state, the goal state can be achieved on 12 out of 50 trials. In the large complex problem the goal state can only be achieved on 5 trials. Given these limitations, it is impossible to conclude whether the large complex problem is actually more difficult to control than the small complex problem, or whether this is an artefact of the system.

Similarly, a number of different methods have been used to assess structural knowledge. Knowledge is sometimes assessed quantitatively, for example subjects are required to predict an output, given a specific input (Berry & Broadbent, 1984; Berry & Broadbent, 1987), or qualitatively, for example, they must indicate the direction of a change, given a general input direction (e.g. “If the size of the workforce is increased...would you expect the sugar output level to increase/decrease/stay the same/don’t know?”, Berry & Broadbent, 1987, p. 9). However, the measurement of structural knowledge at least allows researchers to test the assumption that control performance and knowledge acquisition are dependent processes.

2.4.5 Ecological validity

Although these tasks have been criticised as overly simple (Quesada, Kintsch & Gomez, 2005), Buchner (1995; Buchner & Funke, 1993) argues that many technical systems, such as “video recorders, computer programs, TV sets, digital wrist watches and banking machines” (p.53) can be adequately described in the same way. All of these systems have a finite number of possible input and output states,

and functions that map the relationships between them. This suggests that although these tasks have the appearance of being highly artificial, they may well be ecologically valid.

2.4.6 Summary of the Implicit Learning approach

In summary, the Implicit Learning approach utilises complex problems that represent only the abstract features of dynamic environments. That is, they have a limited number of input and output variables that are causally connected by a set of linear equations. Subjects are given a specific learning goal and the underlying structures of the systems are mathematically well defined. As a consequence, it is possible to precisely specify the interventions that produce optimal performance. Although the complex problems tend to be embedded in a familiar cover story, this problem is overcome by the replication of experimental findings with complex problems that have isomorphic structures but different cover stories. The main criticisms of this approach are that only a limited number of system structures have been utilised, standardised methods have not been developed to assess control performance and structural knowledge, and that the reliability of such measures are likely to be limited as subjects are required to simultaneously acquire knowledge and control the system.

2.5 The Linear Structural Equation approach

Funke (1985; 1986; 1992; 2001) introduced Linear Structural Equation systems as a formal framework in which to study complex problem solving. This framework entails a formal description of the task environment, assumptions about how knowledge is represented and systems are controlled, and a set of diagnostic procedures for assessing what has been learnt.

2.5.1 Structural characteristics

In this approach, the problems are referred to as complex problem solving (CPS) tasks or micro-worlds. Overall, the problems used in the this approach are similar to those used in the Implicit Learning approach, as they have a limited number of input and output variables that are governed by a set of linear equations. The systems are dynamic, so that the current output value depends on the value of the input selected by the subject, and the previous value of the output. Some systems

also include autonomic changes, so that the values of particular output variables change independently on each trial. In terms of structural characteristics, the main difference between the Implicit Learning and Linear Structural Equation approaches is that the systems used in the latter approach contain more variables (typically between 6 and 8).

The underlying structures of the systems used in this approach typically differ on only a few dimensions: The number of variables, the number of relations and the kind of relations between the variables. One example is “MACHINE” (Beckmann, 1994). In this problem, subjects are initially instructed to work out how three controls, “A”, “B” and “C” (the input variables) affect three displays, “1”, “2” and “3” (the output variables). The underlying structure of the system is described by a set of three linear equations. Each of the outputs change in response to the decisions made by the problem solver and two outputs change independently on each trial. An example of a larger system is “LINAS” (Putz-Osterloh, 1993). LINAS contains four inputs and seven outputs interconnected by fifteen linear relations. The labels given to the system variables do not refer to objects in the real world (e.g. “Bulmin”, “Ordal”, “Trimol”) to control for the influence of prior knowledge. As the systems differ on only a few dimensions and are described by linear equations, these factors make it possible to compare the results obtained using different systems, and systematically manipulate system properties to determine whether results generalise to different systems.

2.5.2 Type cover story

The complex problems used in this approach can be broadly split into two categories with respect to whether they are presented in a novel or familiar context. Novel problems such as ‘SINUS’ (Funke, 1992) and ‘MACHINE’ (Beckmann, 1994; Beckmann & Guthke, 1995) have novel labels of the system variables. For example, in ‘SINUS’, participants are told that they must discover how creatures, called ‘OLSCHEN’, ‘MURKERN’ and ‘RASKELN’ affect other creatures called ‘GASELN’, ‘SCHMORKEN’ and ‘SISEN’. In contrast, other problems contain labels that refer to familiar systems in the real world. For example, in Beckmann’s (1994) CHERRY TREE problem subjects are instructed to work out how “HEAT”,

“LIGHT” and “WATER” affect the production of “CHERRIES”, “LEAVES” and “BEETLES”.

Unlike in the other approaches, the familiarity of the variables and context is seen as a factor that is open to experimental manipulation. The effect of familiarity on performance is somewhat disputed, however, as some studies have found that it hinders the acquisition of structural knowledge (Beckmann, 1994; Beckmann & Guthke, 1995; Funke, 1992), and others have found that it makes no difference (Burns & Vollmeyer, 2002). For this reason, the majority researchers now emphasise the use of novel systems, in order to control for the influence of prior knowledge or assumptions (e.g. Funke, 2001; Blech & Funke, 2005; Kröner, Leutner & Plass, 2005; Kluge, 2008).

2.5.3 Task characteristics

As in the Implicit Learning approach, the goal states are well defined and as the systems are mathematically tractable it is possible to specify an optimal intervention. However, the processes of acquiring knowledge about the system, and controlling the system are separated experimentally. A typical experimental procedure consists of an initial *exploration phase* in which problem solvers are first required to determine how the inputs affect the outputs in the absence of specific goals. In a subsequent *control phase* they are instructed to control the system by manipulating the input variables to reach and maintain defined goal states of the output variables. This means that separate measures of structural knowledge and control performance can be derived. From a methodological perspective, this represents a major advantage over the DDM and Implicit Learning approaches.

2.5.4 Operationalisation of dependent variables

Standardised methods have been developed to assess structural knowledge and control performance. Almost all studies utilise causal or structure diagrams to assess problem solvers’ knowledge of the underlying structure of the system (Blech & Funke, 2005). In this procedure, subjects are given a diagram that depicts all the possible relations in the system and they must indicate which relations exist, as well as the direction and strength of each effect. This diagram is usually presented on paper (e.g. Funke, 1985; Vollmeyer, Burns & Holyoak, 1996; Burns & Vollmeyer,

2002), however, computer-based versions have also been realised (e.g. Beckmann, 1994; Beckmann & Guthke, 1995; Osman, 2008). Subjects are typically required to complete multiple causal diagrams over the course of the exploration phase, and they are later used to compare subjects' structural knowledge to the actual underlying structure of the system. This makes it possible to construct learning curves and to determine how problem solvers' mental models of the system change as they intervene on the system variables.

The major problem with the causal diagram method is that subjects tend to differ in the degree to which they are willing to guess about the structure of the system (Funke, 1992). This problem can be overcome by analysing subjects' causal diagrams using an adaptation of the discrimination index from the two-high threshold model for recognition memory (Snodgrass & Corwin, 1988). In this index, structural knowledge scores are corrected for guessing by subtracting the false alarm rate from the hit rate. The hit rate is the number of correctly identified relations divided by the actual number of relations in the system. The false alarm rate is the number of incorrectly identified relations divided by the number relations that are not present in the system, but that could possibly exist given the number of variables. Thus, as in conventional recognition tasks, the discrimination index corrects for subjects' individual response tendency. This method can be used to generate sensitivity indexes for the detection of the existence of relations in the system, as well as the detection of the strengths and directions, and has been shown to be highly reliable (Müller, 1993; Beckmann, 1994).

A more minor problem with the use of causal diagrams is that there is evidence that their use improves the acquisition of structural knowledge (Blech & Funke, 2006). The negative implication of this finding is that causal diagrams may force problem solvers to represent the system in an artificial manner that is more aligned with the method of assessment than real life processes. The positive implication is that causal diagrams may to some extent homogenise problem solvers' initial models of the system, which Vollmeyer et al. (1996) have found can vary radically from subject to subject. These differences may confound assessments of knowledge acquisition because the problem solvers' initial mental model may partly determine how difficult it is to acquire knowledge about the underlying structure of the system. For example, if the initial mental model includes only direct effects between

variables, then the search space is relatively small and the strategies for detection are fairly obvious. Alternatively, if the initial model includes interaction or random effects, the search space becomes much larger, and it is unclear what strategies might be employed to uncover such effects, or disconfirm their existence (Burns & Vollmeyer, 2000). The provision of a causal diagram informs subjects that only certain relations may be expected in the system, and that in addition, they should test for the direction and strength of the relationships between variables. So, whilst the use of causal diagrams may reduce the ecological validity of the process of knowledge acquisition, it also controls for differences in problem solvers' initial mental models of the system which may confound assessments of structural knowledge.

Other methods have also been used to assess subjects' structural knowledge. In prediction tasks, subjects are asked to predict the values of the output variables, given certain values of the input variables (Funke & Müller, 1988; Beckmann, 1994; Vollmeyer, Burns & Holyoak, 1996; Kröner, Plass & Leutner, 2005; Bühner, Kröner & Ziegler, 2008). Another approach is Preußler's (1996) "pair-task", in which two variables names are presented to the subjects and they must decide whether a relationship exists or not. These measures provide less detailed information about the quality of subjects' structural knowledge than the causal diagram method, as they do not allow for a differentiated analysis of knowledge in terms of relations, directions and strengths.

In early studies, control performance was operationalised by calculating the deviation of the current states of the output values from the goal states of the output values in terms of the root means squares criterion (RMS). In this method, deviations become higher the further away the actual states of the system variables are from the goals (e.g. Funke & Müller, 1988; Putz-Osterloh, 1993). Funke (1992; 1993) criticised this method as it assumes that control performance is on an interval scale, which implies that someone who has missed the goal by 100,000 points has performed 10 times as poorly as someone who misses it by 10,000 points. In order to overcome this problem, Funke (1992; 1993) argues that RMS scores should be logarithmically transformed so that larger distances from the goal states are not weighted more heavily. Most recent studies adopt this method (Funke, 1991; 1992; 1993; Vollmeyer, Burns & Holyoak, 1996; Burns & Vollmeyer, 2002; Schoppek,

2002; Kluge, 2008), and many studies report that is that such measures are highly reliable (Kluge, 2008; Körner, Plass & Leutner, 2005; Bühner, Kröner & Ziegler, 2008).

However, Beckmann (1994) argues that this method unduly focuses on the behaviour of the system rather than the behaviour of the problem solver. In particular, it does not take into account that the values of the input variables are usually constrained. Thus, if a poor decision is made on a given trial, a number of trials may be required to realign the output values with the goal states. Hence, the RMS criterion penalises subjects although they may be making an optimal intervention towards the goal state given the current state of the system.

An alternative method, proposed by Beckmann (1994), is to calculate the Euclidean distance between the optimal input intervention (that would bring the system closest to the goal state) and the actual input intervention made by the subject. Using this method, a score closer to zero reflects a more “optimal” intervention to bring the system in line with the desired goal states.

In order to calculate the Euclidean distance between the actual and the optimal intervention for each trial, the values of the output variables on the previous trial and the goal state values are used to solve the set of linear equations underlying the system. This indicates the *ideal* values of the input variables for that particular trial. If the ideal values are within the possible limits of the input variables then the ideal values of the input variables are equal to the *optimal* values of the input variables. However, if an ideal value is outside the limits of the input variables, then the optimal value is set to the closest limit. For example, if the ideal value is 110, and the inputs are constrained to values between -100 and 100, then the optimal value would be set to 100. Once the optimal value for each input variable is determined, the Euclidean Distance (D_{Euclid}) between the actual and the optimal values of the input variables can be calculated for the trial using the equation:

$$D_{\text{Euclid}} = \sqrt{\sum (X_{it}^{\text{actual}} - X_{it}^{\text{optimal}})^2}$$

where X_{it}^{actual} represents the actual value of input variable i at trial t , and

X_{it}^{optimal} represents the optimal value of input variable i at trial t .

For example, imagine a system that consists of three input variables (X, Y and Z) and three output variables (A, B and C). On the first trial, the optimal values of the input variables are calculated to be $X = 13$, $Y = 5$ and $Z = 8$. The subject enters the values 12, 8 and 6 for X, Y and Z, respectively. The Euclidean Distance would then be calculated as:

$$D_{\text{Euclid}} = \sqrt{(12 - 13)^2 + (8 - 5)^2 + (6 - 8)^2}$$

which gives a control performance score of 3.74 for the first trial.

This method avoids the main limitations associated with the RMS criterion. Firstly, it focuses on the subjects' behaviour (i.e. input interventions) and not on the outcomes of the system. Secondly, it does not penalize the subject for poor decisions made on previous trials as it takes into account that their intervention strategy is limited by the constraints on the values of the output variables. In addition, this performance measure has been shown to have excellent reliability and to correlate with control performance scores based on the RMS criterion (Beckmann, 1994).

One problem with all of the above methods is that they do not take into account the controllability of the system. Controllability reflects the extent to which a goal state can be achieved through random interventions (Strauß, 1993; Beckmann, 1994; Hornke & Kersting, 2006). The controllability of the system is significant because Hornke and Kersting (2006) have found that in some systems it may be possible to reach the desired goal state through random interventions. This suggests that if subjects' control performance scores are not compared to control performance scores generated by random interventions, then it is impossible to conclude whether experimental effects (or lack of them) are in fact psychologically meaningful or are an artefact of the system. Few studies have taken controllability into account (with the notable exceptions of Beckmann, 1994; Kluge, 2008), and obviously this criticism is also relevant to the assessment of control performance in the DDM and Implicit Learning approaches.

2.5.5 Ecological validity

Finally, there is some question over whether the complex problems used in this approach are ecologically valid. Buchner (1995) argues that few systems in the real world will have the "exact properties" (p.48) of linear equations. However, in many

sciences, a general linear model is often used to represent the relationships between variables (Blech & Funke, 2005). So to some extent, these systems can be said to approximate reality.

More importantly, parallels have been identified between hypothesis testing in science and the acquisition of knowledge in the complex problems used in this approach. Both situations involve finding out about the relationships between variables in an unknown environment through the systematic manipulation of variables (Funke, 1992; Burns & Vollmeyer, 2000). In this sense, the demands of complex problems can be said to approximate those required in the context of scientific discovery.

2.5.6 Summary of the Linear Structural Equation approach

In summary, the Linear Structural Equation approach overcomes a number of the problems that are associated with the DDM and Implicit Learning approaches. In contrast to the DDM approach, an optimal solution can be specified for system control as the systems are mathematically well defined and a specific goal is given to problem solvers. This makes it possible to validly compare subjects' control performance scores. Secondly, the use of linear equations limits the number of dimensions on which problems may differ in terms of their system characteristics. This makes it possible to compare the experimental results obtained using problems with different underlying structures. In contrast to both the DDM and Implicit Learning approaches, the experimental separation of the tasks of knowledge acquisition and system control allows for an independent examination of the processes that underlie the performance of each of these tasks. In addition, reliable measures have been developed to assess structural knowledge and control performance which have been consistently used in a large body of prior research.

2.6 Conclusions

One aim of this chapter was to determine whether the research findings from the DDM, Implicit Learning and Linear Structural Equation approaches could be integrated into a holistic model of how people control dynamic systems. On the one hand, the critical evaluation of each approach suggests that a number of factors might preclude a complete integration of the experimental findings in this field, and that

findings from the DDM approach in particular need to be interpreted cautiously. On the other, as the Implicit Learning and Linear Structural Equation approaches share a number of key features their findings should be comparable. In this section, the key issues will be summarised to provide a guide for the subsequent literature reviews presented in Chapters 3 and 4.

Firstly, in terms of structural characteristics, the approaches differ in terms of whether they advocate the use of systems that attempt to represent the complexity of real life or are limited to a number of abstract features. In the former camp, DDM complex problems have highly complex structures that differ on at least ten dimensions. A number of researchers have argued that this makes the comparison of experimental findings that use different complex problems highly problematic, and by extension, this suggests that it will be difficult to reconcile findings from this approach with those reported in the Implicit Learning and Linear Structural Equation approaches. In the latter camp, the dynamic systems utilised within the Implicit Learning and Linear Structural Equation approaches are highly similar: They consist of a small number of input and output variables that are governed by sets of linear equations. The main difference is that the systems used in the Linear Structural Equation approach tend to be larger than those used in the Implicit Learning approach, in that they have more variables and more relations between those variables. This suggests that differences in structural characteristics should not completely confound the comparison of experimental results from these two approaches, and that inconsistent findings might be attributable to differences in system size (the implications of system size for the interpretation of experimental results are discussed in detail in Chapter 3).

Secondly, the use of familiar labels for the system variables may in some cases confound the interpretation of experimental results. The DDM and Implicit Learning approaches utilise familiar cover stories and labels for the system variables, which encourages problem solvers to form assumptions about the underlying structure of the system that may or may not be correct. This suggests that findings from the DDM and Implicit Learning literature with regard to the acquisition of structural knowledge may be confounded by differences in problem solvers' assumptions about the system. In the Implicit Learning literature, however, this issue can be discounted because experimental findings have been replicated with problems with isomorphic

structures but different cover stories. Alternatively, the Linear Structural Equation approach advocates the use of novel cover stories and labels for the system variables to control for the influence of prior experience. The subsequent literature reviews will take into account the possible influence of prior experience in the interpretation and integration of experimental results.

Thirdly, the task(s) set for the problem solver may influence the reliability and validity of control performance scores. In the DDM approach subjects are typically required to formulate, or at least further specify, their own goals. This implies that different subjects may formulate different goals, and thus the comparison of control behaviour to any external criteria of success is somewhat arbitrary. In comparison, the Implicit Learning and Linear Structural Equation approaches advocate the use of well-defined goals. This allows valid comparisons of control performance to be made between subjects, and suggests that what is being measured is due to systematic sources of variability. Consequently, the literature review presented in Chapter 4, which concerns individual differences in knowledge acquisition and control performance, excludes studies in which non-specific goals were given to problem solvers.

Another factor that may influence the interpretation of experimental results is whether the tasks of knowledge acquisition and system control are performed simultaneously or consecutively. In the DDM and Implicit Learning approaches subjects are required to acquire knowledge about the system while they control it. In comparison, in the Linear Structural Equation approach these tasks are separated into an exploration phase and a control phase. The question that remains is whether the tasks of acquiring knowledge and controlling the system involve the same processes when the problem solver is required to do both at the same time. This question is addressed in detail in the literature review presented in Chapter 3.

Finally, only the Linear Structural Equation approach has standardised and reliable measures to assess structural knowledge and control performance. Although the methods of assessment do have some minor problems associated with them, their consistent use makes it possible to directly compare the results of different studies that use the Linear Structural Equation approach. In comparison, researchers within the Implicit Learning and DDM approaches have used a number of different methods

to assess structural knowledge and control performance. Another problem is that reliability coefficients are rarely reported. This suggests that correlational evidence from the Implicit Learning and DDM approaches needs to be interpreted cautiously due to a possible lack of reliability in performance scores.

The issues discussed above also address the second aim of this chapter which was to determine which approach provides the most methodologically sound framework for the study of system control. The evaluation suggests that a methodologically sound approach should entail: Mathematically tractable systems that can be formally described, the use of novel cover stories and variable labels to control for the influence of prior knowledge, the experimental separation of the tasks of knowledge acquisition and system control, well-defined goals for system control and reliable measures of structural knowledge and control performance. In comparison to the DDM and Implicit Learning approaches, the Linear Structural Equation approach can be implemented to meet each of these requirements. Consequently, this is the approach adopted in the empirical work reported in this thesis.

In conclusion, the DDM, Implicit Learning and Linear Structural Equation approaches differ in the extent to which they attempt to replicate the complexity of real world environments in the laboratory. The Implicit Learning and Linear Structural Equation approaches advocate the use of complex problems that represent only the abstract features of complex and dynamic environments. In comparison, the DDM approach attempts to simulate the complexity of the real world in the laboratory. However, although the complex problems utilised in the latter approach may seem to have a closer correspondence to real world situations, the lack of experimental control inherent in their design makes it difficult to draw firm conclusions from the results of studies that utilise them. In contrast, the Implicit Learning and Linear Structural Equation approaches strike a balance between the desire for complexity and experimental control that makes it possible to draw both interesting *and* valid conclusions from the results of their experimental studies.

CHAPTER THREE

THE ROLE OF KNOWLEDGE IN CONTROLLING A DYNAMIC SYSTEM

3.1 Introduction

Currently, there is little consensus regarding the role of that knowledge plays in learning how to control the outcomes of dynamic systems. There are two main points of contention that will be addressed in this chapter. Firstly, there is disagreement over the *type* of knowledge that has to be acquired in order to learn how to control a dynamic system. An argument will be put forward that given certain structural and task characteristics, problem solvers will acquire structural knowledge, and that the acquisition of structural knowledge is preferable if the goal is to develop a flexible skill base for the control of dynamic systems. Secondly, there is disagreement over whether structural knowledge alone is sufficient to produce successful control performance. It will be argued that as yet there is not enough evidence to resolve this debate. As the influence of prior knowledge was discussed in the preceding chapter, this chapter will primarily deal with complex problems that are designed to be novel.

3.2 What type of knowledge is acquired?

With regard to learning to control novel dynamic systems, two main types of knowledge are discussed in the literature: Instance-based knowledge and structural knowledge. Instance-based knowledge, which is also referred to as input-output knowledge (Schoppek, 2002), state transitions (Buchner, Funke & Berry, 1995) or exemplar knowledge (Stanley, Matthews, Buss & Kotler-Cope, 1989) represents particular states of a system. A state of a system consists of the initial value of the output variables, specific input values and the resulting values of the output variables. Structural knowledge represents abstractions of particular instances to rules that describe the relationships between the variables in the system, and may be more or less specific. At the broadest level, only the existence of a relationship between two variables is known. At a more specific level, the direction of the relationship is known, and optimally, the strength – represented in numerical weight values – of the relationship can also be specified (Funke, 1992). Although instance-based and structural knowledge are conceptually distinct, they are also

interdependent because instances can be used to test or infer rules (i.e. acquire structural knowledge).

The distinction between instance-based and structural knowledge is derived from models of rule-induction in problem solving and cognitive skill acquisition (e.g. Simon & Lea, 1974; Dunbar & Klahr, 1988; Anderson, 1983; 1993; Ackerman, 1988; Anderson & Lebiere, 1998; Logan, 1988; 1990). Simon and Lea (1974) describe problem solving that requires rule-induction as a search in two distinct problem spaces: One of instances and one of rules. The instance space consists of all possible states of the problem, while the rule space consists of all possible rules underlying the problem. In complex problems, the instance space corresponds to all the possible states of the system, while the rule space corresponds to all the rules that the problem solver thinks might govern the relationships between the inputs and the outputs. Thus, the rule space is defined by the problem solvers' conceptualisation of the plausible rules within the system. For example, the rule space may or may not include exponential relationships or interactive effects (Burns & Vollmeyer, 2000). Thus, instance-based knowledge consists of information contained in the instance space, while structural knowledge consists of information contained in the rule space.

Similarly, competing models of cognitive skill acquisition emphasise either top-down or bottom-up learning processes that correspond to the acquisition of either structural or instance-based knowledge prior to skilled performance (Taatgen & Wallach, 2002). Top-down models of skill acquisition generally assume that individuals first learn general, abstract rules that through practice turn into specific, usable procedural skills (Anderson, 1983; 1993; Ackerman, 1988; Anderson & Lebiere, 1998). In bottom-up models of skill acquisition, the knowledge base is thought to develop concurrently with performance through practice. Initially, general strategies are applied to reach a desired goal state and the results of each intervention are stored as an instance. After a certain period of practice, skilled action is the result of the storage and retrieval of specific instances that achieve the desired goal state (Logan, 1988; 1990).

A central question then is what type of knowledge is necessary for successful control performance in dynamic systems tasks? Control performance requires

problem solvers to enter specific values of the input variables in order to achieve specific values of the output variables (i.e. generate specific instances). This means that the control behaviour of problem solvers who have acquired either relevant instance-based or structural knowledge is indistinguishable.

However, the key difference in control performance that results from the application of structural knowledge as opposed to the retrieval of instances should be flexibility. Instance-based knowledge should only be useful in situations that are identical to those that have been previously experienced. In comparison, as structural knowledge takes the form of general rules about the inter-relationships between the system variables, it could, at least theoretically, be used to generate any desired system state. Hence, from a training perspective, the acquisition of structural knowledge would be preferable to instance-based knowledge if the goal of training is to develop a flexible skill base. Of course, this is dependent on whether problem solvers can actually acquire and effectively utilise structural knowledge. The following sections will review the evidence for the role of both types of knowledge in system control.

3.2.1 Evidence for the role of instance-based knowledge in system control

In a series of early articles, Berry and Broadbent (Broadbent, 1977; Berry & Broadbent, 1984) drew a comparison between the acquisition of manual skills and learning how to control dynamic systems. They argued that both tasks require a series of inter-related decisions which must be enacted over time. In the case of manual skills, they argued that it is generally accepted that people cannot completely describe how they perform a particular task, and that conversely, knowledge about a task is probably insufficient for successful performance. For example, knowledge of the relationships between the gears and the wheels of a bicycle is unlikely to result in being able to ride a bicycle. Similarly, they argued that knowledge about the relationships in a dynamic system, such as an economy or hospital, is unlikely to be sufficient for successfully controlling its outcomes. Rather, the knowledge base required for successful control performance must be developed as the problem solver attempts to perform the task (i.e. reach the desired goal states).

This argument found support in a series of studies conducted by Broadbent and colleagues (Broadbent, 1977; Broadbent, Fitzgerald & Broadbent, 1986; Berry &

Broadbent, 1984; 1988) who found that problem solvers learnt to control dynamic systems seemingly in the absence of reportable structural knowledge. Berry and Broadbent (1984; 1988) gave subjects the task of learning to control the outcomes of either the SUGAR PRODUCTION or PERSON INTERACTION complex problems. As discussed in Chapter 2, these problems require subjects to learn how to manipulate the value of an output variable by changing the value of a single input variable while concurrently attempting to acquire structural knowledge. In these initial studies, after they interacted with the complex problem, subjects were required to complete questionnaires that assessed their knowledge of the relationship between the system variables by asking them to predict the value of the output variable, given certain values of the input variable. They found that practice increased the subjects' ability to control the value of the output variable but not their ability to answer the questionnaire. Using the CITY TRANSPORTATION task, this dissociation between structural knowledge and control performance was also observed by Broadbent (1977; Broadbent, Fitzgerald & Broadbent, 1986). These results were interpreted as evidence that the development of skilled control performance is independent of the acquisition of knowledge regarding the underlying structure of the system.

One criticism of these studies is that subjects may have had knowledge about the underlying structure of the system that was not adequately assessed by the particular questionnaires used in Berry and Broadbent's studies (Sanderson, 1989; Shanks & St John, 1994; Berry & Broadbent, 1995). However, Berry (1984; 1991) used a number of different question types and still found evidence for the dissociation. Most strikingly, this dissociation was observed in Stanley et al.'s (1989) study, even when subjects were asked to describe what they knew about the system in their own words. They asked subjects to practice either the SUGAR PRODUCTION or the PERSON INTERACTION complex problems and to subsequently explain to someone unfamiliar with the complex problem how to control it. They found that subjects' performance improved long before they could tell someone else how to control it.

The dissociation between verbalisable knowledge and control performance, however, was not entirely supported by the results of Stanley et al.'s (1989) study. They found that after 570 trials, subjects were able to express information that would be useful in helping others control the system. In comparison, subjects in Berry and

Broadbent's (1984; 1989) studies completed a maximum of 60 trials. Similarly, Sanderson (1989) found that structural knowledge was associated with control performance in the CITY TRANSPORTATION complex problem after a significant amount of practice. These results demonstrate that knowledge of the underlying structure of the system is acquired at a much slower rate than knowledge of how to control the system. This supports the claim that controlling the outcomes of a system is not dependent on the amount of knowledge acquired about its underlying structure (Sanderson, 1989; Stanley et al., 1989).

Berry (Berry & Broadbent, 1984; 1988) initially interpreted these results as evidence of implicit learning, which has been characterised as "... a process whereby a person learns about the structure of a fairly complex stimulus environment, without necessarily intending to do so, and in such a way that the resulting knowledge is difficult to express" (Berry & Broadbent, 1995, p.132). That is, the problem solver acquires structural knowledge but they are unable to verbalise it. This knowledge is, however, reflected in their ability to control the system. This explanation has found little support, however, as it has proved quite difficult to demonstrate that subjects acquire "unconscious" structural knowledge.

An alternative interpretation, first put forward by Broadbent et al. (1986), is that these results can be explained in terms of a *look-up table*, which stores the correct actions to be taken in certain situations. The concept of a look-up table corresponds closely to Logan's (1988; 1990) model of cognitive skill acquisition. In Broadbent et al.'s model, the most appropriate action is determined by comparing the current situation to previously experienced situations. The same response is given to situations in which a previous response led to the goal state and a random response is given in new situations. If the random response leads to the goal state then this response is stored for future retrieval. This is a form of instance-based knowledge, as it implies that specific values of input and output variables are stored and retrieved to produce successful performance.

This model of control performance has found considerable empirical support. Marescaux, Luc and Karnas (1989) found that in the SUGAR FACTORY subjects tended to perform better in situations that they had previously experienced and not so well in situations that they had not previously experienced. Subjects also tended to

give the same response in situations in which they had been correct rather than incorrect. Dienes and Fahey (1995) replicated these results using the SUGAR FACTORY and PERSON INTERACTION complex problems. These results suggest that subjects do not need to have any knowledge, either implicit or explicit, of the underlying structure of the system. Rather, performance can be explained by the memorisation and retrieval of specific instances that produce the goal state (Marescaux, Luc & Karnas, 1989; Dienes & Fahey, 1995).

Computational models of system control also support this conclusion. Dienes and Fahey (1995) developed two models of how people learn to control the SUGAR FACTORY; one that stored instances of specific actions that achieve the goal state, and another that started off with a number of different rules to be tested. The former model produced the closest match to the actual data generated by subjects. Similar studies have been conducted which also show that computer models of system control that store instances, rather than rules, most closely fit the data generated by subjects as they learn to control the SUGAR FACTORY (Lebiere, Wallach & Taatgen, 1998; Gibson, Fichman & Plaut, 1997).

More recently, Gonzalez, Lerch and Lebiere (2003) have argued that instance-based knowledge might also account for performance in more complex dynamic systems. They investigated how subjects learnt to control a virtual water distribution system called the WATER PURIFICATION PLANT. The task requires subjects to schedule five water pumps to open and close over a period of time to meet a series of production deadlines. Subjects are only instructed to meet these specific goals and not to try to acquire knowledge about the relationships in the system.

In their computer model of system control, general heuristic strategies are first used to determine the most appropriate action. With exposure to the task, instances that achieve the desired goal state are stored in long-term memory. Gradually, performance shifts from decisions based on the application of general heuristic strategies to the retrieval of appropriate actions from memory. Gonzalez et al. (2003) do not clearly define what is meant by a general heuristic strategy. However, in their computer model, if a situation has not been previously experienced then the water pump with the closest deadline is to be opened. A series of experiments showed that this computer model of task performance closely approximates the learning curves

and performance of human problem solvers. Gonzalez et al. (2003) argue that this indicates that human problem solver rely on a mix of heuristic strategies and instance-based knowledge to control dynamic systems.

A number of other researchers have also proposed that successful control performance might be achieved through the application of heuristic rules, without necessitating the remembrance of instances or the acquisition of structural knowledge (Stanley et al., 1989; Buchner, Funke & Berry, 1995). Buchner et al. (1995) define a heuristic rule as a simple strategy that does not take into account the underlying structure of the system but nevertheless results in adequate performance when it is consistently applied. Like instance-based knowledge, the efficacy of such strategies is highly context specific. Buchner et al. (1995) ran a series of simulation studies that tested the adequacy of different heuristic rules on 30 intervention trials in the SUGAR FACTORY with 500 simulated subjects. In the first simulation study, for 30 trials a workforce of either 800, 900 or 1000 was randomly chosen. In a second simulation study, a workforce of 900 was consistently employed across 30 trials. The success of these strategies was considerable, and in both cases the number of trials on target was close to or better than the number of trials on target of real subjects who interacted with the task for 30 trials. Buchner et al. (1995) argue that if subjects used heuristic strategies, they would not have to store instance-based knowledge and that knowledge of heuristic strategies would not be reflected in post-task questionnaires that assessed structural knowledge.

Similarly, Stanley et al. (1989, Experiment 3) instructed subjects dealing with the SUGAR FACTORY to “Always select the response level half-way between the current production level and the target level, and you will get as close to the target level as possible” (p. 566). They found that such a strategy had an immediate positive impact upon control performance. Thus, it appears that control performance can be successfully achieved through the application of heuristics. This suggests that perhaps the dissociation between structural knowledge and control performance occurs because subjects develop and use heuristic strategies to reach the desired goal states.

There are, however, a number of problems with this argument. Firstly, it seems implausible that subjects should be able to develop efficient heuristic strategies

without instruction or considerable experience with the system. Secondly, if subjects were to apply heuristic strategies, we would expect them to respond consistently to every situation that they encounter. Instead, findings show that subjects respond consistently to situations in which they have been correct, and randomly to new situations and situations in which they have been incorrect (Marescaux, Luc and Karnas, 1989; Dienes & Fahey, 1995). Thus, the dissociation observed between structural knowledge and control performance cannot be attributable to the use of heuristic strategies.

3.2.2 Evidence for the role of structural knowledge in system control

In consideration of the evidence presented above, it may perhaps come as a surprise that there are a significant number of studies that have found that problem solvers actually *do* acquire structural knowledge about dynamic systems. In these studies, a strong association is found between the quality of system control and the amount of structural knowledge that is acquired by problem solvers (e.g. Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer et al., 1996; Kröner et al., 2005; Burns & Vollmeyer, 2002; Osman, 2008a; Kluge, 2008). These studies use Funke's Linear Structural Equation approach (1985; 1986; 1992, see Chapter 2 for more details). In this approach, the complex problems consist of input and output variables that are causally connected through a set of linear equations. A novel cover story and labels are given to the system variables in order to control for the influence of prior experience. As described in Chapter 2, the general experimental procedure consists of an initial exploration phase, during which problem solvers must first try to determine the underlying structure of the system. At certain points during the exploration phase, subjects are asked to create a diagram of the causal relationships that they have detected within the system, including the direction and estimated strength of each effect. A measure of structural knowledge is derived from a comparison of subjects' causal diagrams and the actual underlying structure of the system. This is then followed by a control phase in which problem solvers must try to control the outputs to reach specific goal values by manipulating the inputs.

In direct conflict with the results reported by Broadbent and colleagues (Broadbent, 1977; Broadbent, Fitzgerald & Broadbent, 1986; Berry & Broadbent, 1984; 1988), these studies suggest that the quality of problem solvers' control

performance is dependent on the amount of knowledge that they have acquired about the underlying structure of the system. For example, Funke and Müller (1988) report that control performance and knowledge of the underlying structure of the system are significantly negatively correlated ($r = -.41$), as did Beckmann and Guthke ($r = -.51$, 1995), Vollmeyer and Burns ($r = -.48$ and $r = -.61$, 1995), Vollmeyer et al. ($r = -.57$ and $r = -.65$, 1996), Kröner, Plass and Leutner ($r = -.77$ and $r = -.61$, 2005), Burns and Vollmeyer ($r = -.28$, $r = -.40$, 2002), Osman ($r = -.48$ and $r = -.53$, 2008a) and Kluge ($r = -.68$, $r = -.66$ and $r = -.50$, 2008). The correlation coefficients are negative, as a lower control score indicates better performance.

How can these results be reconciled with those reported by Broadbent and colleagues? In the following sections, we will discuss structural and task characteristics that may influence whether problem solvers are more likely to acquire instance-based or structural knowledge.

3.3 Explanations for conflicting results

3.3.1 The size of the problem space

Schoppek (2002) argues that the size of a systems' problem space determines whether problem solvers will use instance-based or structural knowledge to control the outcomes of the system. In systems where the input variables are independent of each other and there are no side effects, the size of the problem space can be calculated by multiplying the number of possible input values with the number of possible output values. The complex problem used by Broadbent and Berry (1984; 1988) has one input variable and one output variable, each with twelve possible states. Therefore, the number of possible states is $12 \times 12 = 144$. In comparison, in studies that report a relationship between structural knowledge and control performance (e.g. Funke & Müller, 1988; Vollmeyer et al., 1996; Kröner et al., 2005; Burns & Vollmeyer, 2002; Osman, 2008a; Beckmann & Guthke, 1995), the problems have three inputs and three outputs, and as such the number of possible states is much larger. For example, Schoppek (2002) gives an example of a system with three input variables and three output variables, with integer ranges from -10000 to 10000. This would result in $20000^3 = 8 \times 10^{12}$ possible states. He argues that in these larger systems it is implausible that problem solvers would be able to store the amount of instance-based knowledge necessary to successfully control the

system. Hence, they must reduce the amount of information to be stored by abstracting general rules from the data.

The problem with this argument is that problem solvers do not need to store all possible states in the problem space; rather, they only need to store instances that achieve the desired goal state. In the SUGAR FACTORY complex problem, if the goal is to reach a target production of 9,000 tonnes there are only 6 possible input values that would realise this target state, given various starting values (this is the goal given in all the studies that use the SUGAR FACTORY e.g. Berry & Broadbent, 1984; 1988; Sanderson, 1989; Stanley et al., 1989; Marescaux et al., 1989; Dienes & Fahey, 1995; Buchner et al., 1995). Similarly, in larger systems, the number of on-target instances will still only represent a small fraction of the entire problem space. Essentially, we would argue that the size of the problem space is not indicative of the number of instances that must be stored by problem solver and is therefore not informative as to whether problem solvers will acquire instance-based or structural knowledge.

Moreover, it could be speculated that if a system contains a large number of highly inter-connected variables it may be more efficient to remember instances, rather than derive rules, which is directly the opposite of Schoppek's (2002) argument. For example, LOUHAUSEN (Dörner et al., 1983) contains over 2000 highly inter-connected variables. In such systems, it may be impossible to derive an accurate model of the underlying structure of the system because of the information processing limits imposed by the human cognitive system. Given these limits, remembering and re-enacting particular actions that achieve the desired goal state may be the only way to perform successfully, unless one indeed has structural knowledge. Thus, we argue that it seems unlikely that there is any relationship between the size of the systems' problem space and whether problem solvers will use instance-based or structural knowledge.

3.3.2 Random error factor

In the SUGAR FACTORY the random error factor makes it difficult to calculate the relationship between the variables from the data. The underlying structure of the system is given by the equation:

$$P_t = 20 \times W_t - P_{t-1}$$

where P_t is the current sugar production, W_t is the number of workers and P_{t-1} is the sugar production on the previous trial. On two thirds of the trials ± 1000 is added to the results derived from the above equation as a random error factor. This means that no one input state is associated with a particular output state. This makes it difficult to calculate the exact relationship between the variables from the data, unless the random error factor is known.

3.3.3 Constraints on the values of the output variables

In addition, constraints on the values of the output variables may also make it difficult to derive the underlying structure of the system. The values of W and P are constrained to values between 1000 and 12000. This means that many interventions do not lead to any change in the output value, and thus that there are only a limited number of system states that are informative as to the underlying structure. For example, if the current output value is 3000 and an input value of 300 is entered, this will lead to an output value of 3000 (assuming that the error factor is not active on that trial). This reflects the rule underlying the system, as $(20 \times 300) - 3000 = 3000$. Alternatively, if an input value of 1000 is entered this will produce an output of 1000 due to the constraints on the values of the outputs. Given the equation however, we would expect an output value of -2000. Based on this system state, the underlying equation cannot be successfully derived. The problem solver, however, is naive to the constraints on the system variables and to which of the system states are informative. Hence, it could be speculated that if they were trying to derive a linear equation governing the behaviour of the system based on the generated system states they may come to believe that the relationship between the variables is random. In addition to the random error factor, the constraints on the system variables are likely to make inducing the underlying structure of the system difficult, if not impossible.

In comparison, the complex problems used in studies that have found that problem solvers do acquire structural knowledge do not include random error factors or constraints on the output variable values (e.g. Funke & Müller, 1988; Vollmeyer et al., 1996; Kröner et al., 2005; Burns & Vollmeyer, 2002; Osman, 2008a; Beckmann & Guthke, 1995). Consequently, any system state that is produced will reflect the equations that govern the behaviour of the system, and therefore problem

solvers will have more opportunities to correctly induce the rules underlying the system. This may explain why problem solvers are able to acquire structural knowledge about these systems.

3.3.4 The type of learning goal given to problem solvers

The third significant point of difference between studies that find an association between structural knowledge and control performance and those that do not is the type of learning goal that is given to subjects. Specifically, in the studies that report that structural knowledge and control performance are not associated, subjects received a specific learning goal from the beginning of the task. That is, they were instructed to learn to produce and maintain a specific output value (e.g. Broadbent, 1977; Broadbent et al., 1986; Berry & Broadbent, 1984; Berry, 1984; Berry, 1991; Stanley et al., 1989; Marescaux et al., 1989; Dienes & Fahey, 1995). In Berry and Broadbent's (1988) study, subjects were *additionally* told to look for the pattern underlying the system and to do this whilst reaching and maintaining the specific goal. The latter type of instruction means that subjects essentially have two goals to pursue at the same time; a goal to reach and maintain the target outcome and a goal to test the accuracy of their knowledge. Trying to pursue both goals might result in only a narrow search of the problem space in order to find the solution path to the specific goal, with the result that general structural knowledge is not acquired (Geddes & Stevenson, 1997; Vollmeyer et al., 1996; Burns & Vollmeyer, 2002; Wirth, Künsting & Leutner 2009).

In comparison, in studies that report an association between the amount of structural knowledge acquired and the quality of control performance, subjects initially received a non-specific learning goal (Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer & Burns, 1995; Vollmeyer et al., 1996; Kröner et al., 2005; Kluge, 2008; Burns & Vollmeyer, 2002; Osman, 2008a). That is, subjects were first instructed to explore the system in order to look for relationships between variables. During the exploration phase, they were not informed of the goals that they later had to reach in the separate control phase. This type of instruction might result in a wider search of the problem space and so foster rule learning (Vollmeyer et al., 1996; Geddes & Stevenson, 1997; Burns & Vollmeyer, 2002; Wirth et al., 2009).

Findings from a number of studies support the claim that the type of knowledge that is acquired is dependent upon the learning goal given to subjects. Using Berry and Broadbent's (1984) PERSON INTERACTION complex problem, Geddes and Stevenson (1997) instructed subjects to explore and induce the rules underlying the system (non-specific goal), practice reaching a specific goal state (specific goal), or practice reaching a specific goal state while trying to induce the rules (dual goal). They found that subjects who were given a non-specific learning goal showed evidence of rule learning, as they performed well on familiar and unfamiliar prediction questions and could answer general questions about the underlying structure of the system. Subjects who received a specific goal or dual goals performed poorly on this task. In addition, subjects who received a non-specific goal had better control performance than subjects in the other conditions in a later control phase in which new goals were set for performance. Most importantly, and in contrast to the results previously reported with this complex problem (e.g. Broadbent & Berry, 1984; 1989), structural knowledge was significantly positively correlated with control performance, but only in the non-specific goal group. This suggests that if the underlying structure of the system is induced, then it will be used to control the outcomes of the system.

Burns and Vollmeyer (2002) tested the hypothesis that the presence of specific goals during the exploration phase hinders the acquisition of structural knowledge. The complex problem was based on the Funke's (1985; 1986; 1992) Linear Structural Equation approach and the system consisted of three inputs and three output variables. All subjects were instructed to learn about the relationships between the variables during an initial exploration phase. In addition, half the subjects were told the values of the output variables that they would have to reach during a later control phase. Consistent with the results of Geddes and Stevenson (1997), subjects who knew about the goals during the exploration phase acquired less structural knowledge than subjects who did not know about the goals. This suggests that problem solvers cannot acquire structural knowledge while they attempt to control the system.

The results of this study also suggest that subjects who have specific goals during the exploration phase acquire instance-based, rather than structural knowledge. When the goals for performance were the same as those given during the

exploration phase, subjects who had received a specific goal performed similarly to those who had received a non-specific goal. However, when novel goals were set for performance then subjects who had been given a specific goal performed more poorly than those who had been given a non-specific goal. This shows that the method used to control the outcomes by subjects who had been given a specific goal did not transfer to new situations, and suggests that the knowledge they acquired was instance based. In a study utilising the same complex problem, Osman (2008a) replicated these results, as did Wirth et al., (2009), who found the same results using a complex problem that represented the principles of buoyancy in liquids.

Some evidence also suggests that, in addition to having a non-specific goal (e.g. explore and induce the rules), direct instructions to test hypotheses also foster rule learning. In a series of experiments, Vollmeyer and Burns (1995) attempted to promote structural knowledge by giving subjects a specific hypothesis to test. The complex problem used in the experiments was based on the Funke's (1985; 1986; 1992) Linear Structural Equation approach and the system consisted of three inputs and three output variables. The first experiment examined the effect of goal specificity and hypothesis instruction on the acquisition of structural knowledge and control performance. Subjects in a specific goal condition were informed as to the goals that they would have to reach from the outset of the exploration phase, while subjects in a non-specific goal condition were given the goals at the start of the control phase. In addition, half of all subjects were given a hypothesis and instructed to test it to determine whether it was correct. Unbeknownst to the subjects, the hypothesis was indeed correct. Overall, the results showed that subjects with a non-specific goal acquired more structural knowledge, as did subjects who received a hypothesis to test. There was no interaction between these variables, and Burns and Vollmeyer argue that this suggests that a non-specific goal and hypothesis testing instructions have a similar effect; That is, they both encourage a search of the hypothesis space. Burns and Vollmeyer then conducted a follow-up experiment to test whether the positive effect of hypothesis instruction that was initially observed was actually because subjects were given a correct piece of information about the system. This time, subjects were either given a correct hypothesis to test, an incorrect hypothesis to test or just told there was a possible link between two variables. The results showed that the instruction to test any hypothesis (correct or incorrect) lead to

the acquisition of more knowledge than just information that a link might be present. Hence, these results suggest that problem solvers will acquire structural knowledge when the task characteristics are conducive to searching the hypothesis space.

In summary, the type of knowledge that is acquired about complex problems appears to be dependent on certain structural and task characteristics. In particular, the random error factor and constraints on the values of the output variables in the complex problems used by Broadbent and colleagues (Broadbent, 1977; Broadbent et al., 1986; Berry & Broadbent, 1984; Berry, 1984; Berry, 1991; Stanley et al., 1989; Marescaux et al., 1989; Dienes & Fahey, 1995) limit the number of instances that can be used to abstract general rules. This may explain why rule-induction is so difficult even though the equation that governs the behaviour of these systems appears to be very simple. This subsequently suggests that problem solvers may have no alternative but to remember instances that achieve the desired goal state. Secondly, Broadbent and colleagues gave subjects a specific goal from the outset of the task, and findings suggest that a specific goal inhibits the acquisition of structural knowledge. The evidence suggests that subjects must first explore the complex problem, in the absence of a specific goal, in order to acquire structural knowledge (Geddes & Stevenson, 1997; Vollmeyer & Burns, 1995; Vollmeyer et al., 1996; Burns & Vollmeyer, 2002; Osman, 2008a; Wirth et al., 2009). This suggests that structural knowledge must be acquired *prior* to the instruction to control the system.

The question that now arises is whether structural knowledge alone is sufficient in order to effectively control the outcomes of dynamic systems. Namely, is there a direct causal relationship between the amount of knowledge acquired by problem solvers and the quality of their system control? Or is an additional process required before knowledge can be translated into effective control actions? Studies that have found an association between structural knowledge and control performance do not adequately answer this question because subjects had the opportunity to interact with the system prior to controlling it. Successful control performance may then be attributable to a combination of structural knowledge and activity, rather than structural knowledge alone. Therefore, in the following section, we will examine whether there is evidence for a direct causal relationship between structural knowledge and control performance.

3.4 The acquisition of structural knowledge through direct instruction

One way to address the question of whether the acquisition of structural knowledge is sufficient for successful system control is to attempt to promote structural knowledge through direct instruction and observe the effect on control performance. If structural knowledge is sufficient to control the outcomes of dynamic systems, then the provision of structural information should provide an immediate benefit to control performance (Blech & Funke, 2005).

In a series of studies Putz-Osterloh (Putz-Osterloh & Lürer, 1981; Putz-Osterloh, 1993) attempted to promote structural knowledge through the provision of a causal diagram. Putz-Osterloh and Lürer (1981) gave an experimental group a diagram of the underlying structure of the complex problem TAILORSHOP, while a control group received no supporting information. The diagram was presented on a sheet of paper, as a system of arrows linking the input and the output variables. It was found that the experimental group were no better at controlling the system than the control group. Although this suggests that structural knowledge cannot be directly translated into effective control actions, this interpretation is problematic as TAILORSHOP has a familiar cover story and labels for the system variables. This may obviate any beneficial effect that structural information has on control performance because problem solvers' assumptions about the system might make it difficult to incorporate new information into their existing mental model of the system.

This problem was overcome in a later study conducted by Putz-Osterloh (1993). The design was the same as that used by Putz-Osterloh and Lürer (1981), however, this study used the complex problem LINAS which has a novel cover story and labels for the system variables. The results replicate those found by Putz-Osterloh and Lürer (1981): Subjects who received the causal diagram did not perform significantly better than those who were required to control the system without such information. Putz-Osterloh (1993) argues that these results indicate that successful control performance can be accomplished without the aid of specific knowledge and that a simple strategy of trial-and-error may be as efficient in reaching the goal states as the application of the rules underlying the system.

However, this implies that all subjects were able to effectively control the system regardless of the amount of information that they received. This conclusion seems unjustified given that a criterion for successful performance was not used in either study. An equally plausible explanation is that those who were given structural information simply failed to apply it, and that all subjects performed poorly. This suggests that structural knowledge cannot be directly translated into effective control actions.

This is the conclusion reached by Preußler (1996), who also found that structural information did not have the expected beneficial effect on control performance. Subjects in a control group explored the complex problem LINAS without assistance, while subjects in an experimental group were instructed using standardised examples as to how each input affected each output with text-based explanations. No differences in control performance were detected. This result is particularly surprising considering that findings show that most problem solvers are unable to acquire a complete or accurate representation of the underlying structure of a system through an unguided exploration of the system variables (Funke & Müller, 1988, Müller 1993; Beckmann, 1994; Beckmann & Guthke, 1995; Vollmeyer et al., 1996; Burns & Vollmeyer, 2002; Kröner, 2001; Schoppek, 2002; Kröner et al., 2005; Kluge, 2008; Osman, 2008a). These results suggest that structural knowledge, when it is acquired through direct instruction, does not provide an advantage to control performance over knowledge that is acquired through a free exploration of the system variables.

In summary, the results suggest that the acquisition of structural knowledge through direct instruction may be insufficient to promote control performance. The results of Putz-Osterloh's (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993) and Preußler's (1996) studies show that subjects who acquire structural knowledge through direct instruction do not have better control performance than subjects who perform without knowledge or incomplete knowledge. In contrast, a number of studies have found that the amount of knowledge acquired by problem solvers through an exploration of the system variables is strongly associated with the quality of their control performance (Funke & Müller, 1988; Vollmeyer et al., 1996; Kröner et al., 2005; Burns & Vollmeyer, 2002; Osman, 2008a; Beckmann & Guthke, 1995; Kluge, 2008). This inconsistent pattern of results suggests that problem solvers may

need to interact with the system variables before structural knowledge can be effectively utilised.

One alternative explanation that should be considered is that the instructional methods used to inform subjects of the underlying structure of the system in these studies may not have been sufficient to promote structural knowledge. Firstly, in Putz-Osterloh's studies (1993; Putz-Osterloh & Lürer, 1981), subjects may not have understood how the diagram presented on paper related to the input and output variables presented on screen. In order to understand the meaning of the diagram, subjects may require a direct demonstration of how the inputs affect the outputs. Secondly, although subjects in Preußler's (1996) study did receive an explanation as to how the inputs affect the outputs, they did not receive a structural diagram that they could refer to during the control task. Therefore, they may have been unable to recall this information during the control task. Findings show that information that is permanently available, or provided as it is needed, is much more effective in promoting the acquisition of knowledge than information that is given prior to the interaction with a task (Kotovsky, Hayes & Simon, 1985; Berry & Broadbent, 1987; Leutner, 1993; Hulshof & De Jong, 2006). If these conditions were met then structural information may well benefit control performance.

These limitations indicate that as yet there is insufficient evidence to conclude whether there is a direct causal relationship between the acquisition of structural knowledge and control performance. We can conclude, however, that structural information in the form of a causal diagram or standardised examples does not benefit control performance. One reason for this may be that these instructional methods are inadequate to promote problem solvers' structural knowledge. Another reason for this may be that declarative knowledge must first be translated into procedural knowledge in order to be useful for system control. This in turn implies that the relationship between structural knowledge and control performance may be moderated by another factor. In the following sections we will examine the conditions that may allow problem solvers to translate structural knowledge into effective control actions.

3.5 Practice

A number of researchers argue that a period of goal-orientated practice may be required before problem solvers can effectively translate structural knowledge into control actions (Putz-Osterloh, 1993; Preußler, 1996; Preußler, 1998; Schoppek, 2004). As discussed above, Putz-Osterloh (1993) found that subjects who were given structural information performed no better than those who did not receive such information. However, differences were observed in a follow-up study conducted six months later. This included only those subjects who had reached the highest level of control performance in the initial study. Subjects dealt with a complex problem that was identical to that used previously, except that one new relationship was added to the system. This time, subjects who had received the causal diagram had better system control, and were more likely to detect the difference between the initial system and the new system. Given that the advantage to control performance was only evident after subjects had considerable exposure to the system, Putz-Osterloh (1993) suggests that in order to benefit from structural information problem solvers may need to practice applying it.

This claim seemingly finds further support in a study conducted by Preußler (1998), which used the complex problem LINAS to examine the effect of structural information and practice on control performance. An experimental group were given a causal diagram and completed practice tasks in which goal values had to be attained by manipulating the input variables. Each task was repeated until the subject reached the target values. The control group had to perform the same practice tasks, although without having the diagram available and without having the chance of retries until the correct response was found. In a subsequent control phase in which new goals were set for performance, subjects who received the diagram had better control performance than those who performed without such information. Again, it has been argued that these findings suggest that structural information does not directly benefit control performance, as it seems knowledge needs to be practiced in the context of application (Schoppek, 2004; Preußler, 1998).

However, one significant limitation of Preußler's (1998) study is that the additional instructions given to the experimental group confounds the results. Both the experimental and control group completed the practice tasks. As the experimental

group had also to find the correct solutions to the practice tasks, this setting provided them with more opportunity to: a) interact with the system, b) evaluate the accuracy of their acquired knowledge and c) acquire more knowledge about the system. The observed advantage in control performance may be the result of any of these differences. Therefore, it is impossible to conclude from the results of this study that subjects benefit from structural information if they practice using it.

In summary, although the aforementioned studies have some significant limitations, they do suggest that under some conditions structural information may benefit system control. More research is required to determine whether a period of goal-orientated practice is required before problem solvers can utilise knowledge, or whether problem solvers may simply need to interact with the system in order to better understand how to apply the information that they have been given. To the author's knowledge, no studies have tested the latter hypothesis. However, a number of studies have investigated whether problem solvers who interact with a system have better structural knowledge and system control than those who observe the same pattern of interventions. These studies will be reviewed in the following section.

3.6 Intervention vs. observation

Controlling the outcomes of a system requires declarative as well as procedural knowledge; declarative knowledge about the dependencies between the system variables and procedural knowledge about how to change input variables in order to achieve the desired effect on the output variables. There is an ongoing debate, however, as to whether subjects must directly interact with a system in order to acquire procedural as well as declarative knowledge.

In order to address this debate, Funke and Müller (1988) allowed subjects in one condition to explore the complex problem SINUS. In a yoked-control design, subjects in a second condition observed the interventions made by their experimental twin in the first condition. Both groups had to complete causal diagrams and subsequently control the system to reach specific goals. The system consisted of three inputs and three outputs linked by a set of linear equations and had a novel cover story and labels for the system variables. Subjects who observed interventions acquired more knowledge of the underlying structure of the system, yet those who

interacted with the system variables had better system control. These findings demonstrate that advantages in terms of structural knowledge do not necessarily translate into advantages in controlling the system, and that problem solvers who interact with a system may acquire some additional sort of knowledge that benefits control performance in comparison to those who only observe the outcomes of a system.

Beckmann (1994) found a similar pattern of results using a method that was designed to encourage subjects to acquire complete knowledge of the underlying structure of the system. Using the complex problem MACHINE, which is similar to SINUS, Beckmann (1994) instructed subjects in one condition to make the interventions that were considered optimal to diagnose the underlying structure of the system. That is, on the first trial the inputs were to be set at zero, so that any autonomic changes in the output variables could be detected. Subsequently, the inputs were to be changed one at a time, so that the effect of each input could be observed on each output. Subjects in another condition observed the same optimal interventions being made on the system variables. Although subjects in each condition did not differ in terms of the structural knowledge that they acquired, subjects who interacted with the system variables had significantly better control performance than those who observed interventions, as in Funke and Müller's (1988) study. Beckmann (1994) argues that this advantage is due to the acquisition of procedural knowledge, as well as declarative knowledge in relation to the system variables.

In a series of studies, however, Osman (2008a; 2008b) has found that "observation can be effective as action" in learning to control dynamic systems. As in Funke and Müller's (1988) study, Osman (2008a) manipulated whether subjects explored the task or observed the exploration of another subject. In addition, subjects were either given a specific or non-specific learning goal. The system consisted of three inputs and three outputs linked by a set of linear equations and had a novel cover story and labels for the system variables. Whether or not subjects intervened on the system variables appeared to have no impact on either knowledge or control performance. The same result was observed in a second study with a slightly different manipulation of the type of learning goal that subjects received (Osman, 2008b). In contrast to previous findings (Funke & Müller, 1988; Beckmann, 1994),

these results suggest that interacting with the system variables confers no additional advantage for system control over observing the same system states.

One possible explanation for this inconsistency could lie in differences in the complexity of the systems used in these studies. The main difference between these studies is that the systems used by Funke and Müller (1988) and Beckmann (1994) appear to be much more complex than used by Osman (2008a; 2008b), in that they have more relationships between the variables (all the systems have three input and three output variables). It may be that intervening on the system variables provides an advantage over observing the same interventions only when the amount of information to be processed exceeds certain limits. In other words, activity may promote the processes of chunking and efficiently organising information in functional units that otherwise would overload the cognitive system. However, this explanation is purely speculative, and requires further investigation.

3.7 Conclusions

In conclusion, there is sufficient evidence to answer the question of whether structural knowledge is *necessary* to control the outcomes of dynamic systems. A number of studies suggest that dynamic systems can be controlled through the application of heuristic strategies or instance-based knowledge (Broadbent, 1977; Broadbent & Berry, 1984; 1987; Berry, 1984; 1991; Broadbent et al., 1986; Stanley et al., 1989; Buchner et al., 1995; Gonzalez et al., 2003). The success of heuristic strategies and instance-based knowledge is, however, highly context specific, as it does not generalise to new goals for system control (Marescaux et al., 1989; Dienes & Fahey, 1995). Flexible system control requires the acquisition of structural knowledge (Geddes & Stevenson, 1997; Burns & Vollmeyer, 2002). Thus, the acquisition of structural knowledge is necessary to promote the development of transferrable control skills.

The question that then arises is whether the acquisition of structural knowledge is *sufficient* for successful system control. Some findings suggest that structural knowledge either has to be acquired through an exploration of the system variables or practiced in the context of application in order to be useful for system control (Putz-Osterloh & Lür, 1981; Funke & Müller, 1988; Putz-Osterloh, 1993; Beckmann, 1994; Preußler, 1996; 1998). This would suggest that structural

knowledge must be translated into procedural knowledge through an additional process before it can be used to control the outcomes of a system. Alternatively, it may be that the methods used to inform subjects of the underlying structure of the system in previous studies (e.g. Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996) may have been inadequate to effectively promote structural knowledge, and thus directly influence control performance. The empirical study reported in Chapter 5 addresses these issues.

A second issue is that a causal relationship between the amount of structural knowledge that is acquired by the problem solver and the quality of their control performance has not been sufficiently established. While the findings of a number of studies show an association (correlation) between structural knowledge and control performance, these results do not imply causality (e.g. Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer & Burns, 1995; Vollmeyer et al., 1996; Burns & Vollmeyer, 2002; Kröner et al., 2005; Osman, 2008a; Kluge, 2008). Systematic manipulations of the amount of knowledge that problem solvers are able to acquire are needed to examine the nature of this relationship. Thus far, previous studies have only evaluated whether different methods of control (e.g. trial-and-error as opposed to knowledge-based) or different methods of acquiring knowledge (e.g. structural information as opposed to exploration) result in differing levels of control performance (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1981; Preußler, 1996). As such, it is unclear whether increments in structural knowledge result in improvements in control performance. This issue is addressed in the empirical studies reported in Chapters 6 and 7.

Finally, the evidence suggests that whether activity or practice is required to promote the application of knowledge may be dependent on the complexity of the system. In the literature, system complexity is often discussed as a potential moderator variable (Dörner, 1987; Brehmer & Dörner, 1993; Kerstholt & Raaijmaker, 1997; Kluge, 2008). However, experimental results are rarely replicated with different systems. This state of affairs makes it difficult to determine whether the experimental results obtained with complex problems are scalable to real world systems which are likely to be much more complex. Therefore, the empirical study reported in Chapter 7 attempts to replicate the results derived in Chapters 5 and 6 with systems of differing complexities.

CHAPTER FOUR

INDIVIDUAL DIFFERENCES IN THE ACQUISITION OF STRUCTURAL KNOWLEDGE AND CONTROL PERFORMANCE

4.1 Introduction

Thus far our discussion has focussed on how knowledge influences the control of complex problems at an aggregate level. Although a direct causal link is yet to be established, findings suggest that given certain structural and task characteristics, problem solvers who acquire more knowledge about the underlying structure of a system will have better control performance than problem solvers who acquire less knowledge. The outcome of performance can then be seen as a function of a) individual differences in the capacity to acquire knowledge and b) individual differences in the capacity to apply knowledge. This section will be concerned with possible sources of individual differences in performing these two tasks.

A number of researchers have argued that the role of intelligence in system control remains unclear (Kluwe, Misiak & Haider, 1991; Wenke & Frensch, 2003; Wenke, Frensch & Funke, 2004). The nature of this debate will be described in more detail in Section 4.3, however, its resolution seems critical considering that traditional measures of intelligence have been shown to be the best predictors of job and academic performance (Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg & Urbina, 1996; Schmidt & Hunter, 1998) and that complex problems are now used as additional predictors of success in such settings (Funke, 1998; U. Funke, 1998; Hornke & Kersting, 2005; Kluge, 2008). Therefore, in order to contribute to this debate, our search for the sources of individual differences will be restricted to the cognitive or intellectual abilities that are likely to influence performance.²

In this thesis, we adopt the cognitive approach to understanding intelligence, which regards intellectual ability as a set of general cognitive resources that are

² We do, however, acknowledge that individual differences in the acquisition of structural knowledge and system control is likely to be the result of both cognitive and non-cognitive factors. A number of studies have found that non-cognitive factors such as motivation, metacognition and emotions have a significant impact on the cognitive processes involved in the acquisition of structural knowledge and system control (e.g. Vollmeyer, Rollett & Rheinberg, 1997, 1998; Vollmeyer & Rheinberg, 1999, 2000; Spering, Wagener & Funke, 2005; Barth & Funke, 2009).

available to an individual (Snow & Lohman, 1989; Lohman, 2000)³. According to this view, performance on intelligence tests and experimental tasks can be understood in terms of the information processing demands that they impose on the cognitive system (e.g. Carroll, 1993; Sternberg, 1988). Obtained correlations between intelligence test scores and experimental tasks can then be explained on the basis of shared information processing demands (Carroll, 1993), and there is now a large body of research that has detailed the information processing demands of intelligence tests (for a review see Lohman, 2000). Therefore, the comparison of performance on intelligence tests and the acquisition of structural knowledge and control performance in complex problems should allow us to identify consistent sources of individual differences in the cognitive processes that discriminate between successful and less successful problem solvers.

Few studies have separately considered the cognitive demands of acquiring knowledge and controlling the outcomes of dynamic systems at an individual differences level (with the exceptions of Beckmann, 1994; Beckmann & Guthke, 1995; Kröner et al., 2005; Bühner, et al., 2008). This is in part because until recently few studies had experimentally separated these two tasks, and/or measured structural knowledge independently from control performance (see Chapter 2 for more details). The main limitation of this approach is that these two tasks are likely to impose different demands on the problem solver, and yet control performance also appears to be related to the acquisition of structural knowledge. Therefore, to understand what makes control performance difficult, some idea is needed of the distinctive processes that underlie the acquisition of structural knowledge and control performance.

Thus, in the following sections firstly a framework will be introduced that describes at a process level how problem solvers might acquire knowledge about dynamic systems. This will be mapped onto what is known about the demands of traditional intelligence tests. This exercise will then be repeated for control performance. The empirical relationships between performance on intelligence tests and performance on complex problem solving tasks, demonstrated in previous

³ This approach can be seen as complementary to the psychometric approach to the study of intelligence, which is primarily concerned with the organisation of abilities within a nomological network (e.g. Vernon, 1965; Cattell, 1971; Horn, 1968).

research, will then serve as a test of our hypotheses about possible sources of individual differences in the acquisition of structural knowledge and control performance.

4.2 Individual differences in the acquisition of structural knowledge

4.2.1 Understanding the knowledge acquisition process: Dual-space Search Model

Klahr and Dunbar's (1988) model of Scientific Discovery as a Dual-space Search (SDDS) has frequently been used to describe the process of knowledge acquisition in dynamic systems (e.g. Shute, 1990; Vollmeyer et al., 1996; de Jong & van Joolingen, 1997; Geddes & Stevenson, 1997; Burns & Vollmeyer, 2000; 2002). This model is an extension of Simon and Lea's (1974) model of problem solving that requires rule-induction to the domain of scientific discovery. In scientific discovery the main task for the learner is to discover the properties of a given domain through hypothesis testing. Conceptually, this is identical to knowledge acquisition in complex problems as it also involves the exploration of a novel environment (Funke, 1992).

The main assumption in the SDDS model is that hypothesis testing involves a search in two distinct, but related, spaces: A hypothesis space and an experiment space. The hypothesis space consists of all the possible relationships between the variables in a task, while the experiment space consists of all the possible states within a task. In Simon and Lea's (1974) original model, these were referred to as the rule space and instance space, respectively. Three performance components are specified: Search hypothesis space, test hypothesis and evaluate evidence. In the context of acquiring knowledge about a dynamic system, a search in the hypothesis space involves generating a hypothesis about the underlying structure of the system, and making a prediction about the behaviour of the system based on this hypothesis. Testing hypotheses involves a search in the experiment space by designing and implementing experiments through the manipulation of the input variables. The resulting evidence is then evaluated with respect to the initial prediction, and hypotheses are confirmed or modified on the basis of the results. This is an iterative process that continues until the problem solver decides that he or she has correctly induced the underlying structure of the system. Extensions of this model have been proposed that deal with how knowledge about the system is represented, but they do

not propose any additional performance components (Burns & Vollmeyer, 2000; de Jong & van Joolingen, 1997).

To test this model, Klahr and Dunbar (1988) examined subjects' verbal protocols as they learnt how to operate a computer controlled robot. They confirmed that the discovery process seemed to fit the model of a search in the two spaces. However, they noted that prior knowledge of computer programming was a key determinant of whether subjects began their search in the hypothesis space or in the experiment space. If prior knowledge was available, then subjects formed hypotheses immediately through "analogical mapping, heuristic search, priming, reminders or conceptual combination" (Klahr, 2000, p. 33). Thus, a hypothesis was initially generated and tested in the experiment space. If hypotheses could not be evoked from memory, then a correct hypothesis was generalised based on the results of an experiment. The basic components of performance remained the same, however, regardless of whether the problem solver began their search in the hypothesis or the experiment space (Klahr & Dunbar, 1988). This model has found further support in subsequent studies both within the domain of scientific discovery (Dunbar, 1993; Klahr, Fay, & Dunbar, 1993), and in relation to knowledge acquisition in complex problems (Vollmeyer et al., 1996; Vollmeyer & Burns, 2002; van Jooligen & de Jong, 1997; Schoppek, 2002).

4.2.2 Implications of the SDDS model for individual differences in knowledge acquisition

In the context of novel complex problems for which prior knowledge is not available, the SDDS model implies that effective knowledge acquisition involves a) the design of experiments that are appropriate for abstracting general rules from the data or testing hypotheses, b) making predictions based on abstracted rules about the values of the outputs given certain inputs and c) the abstraction of general rules from the data. The latter two demands can be categorised as types of reasoning, and thus it can be inferred that reasoning ability represents one possible source of individual differences in the acquisition of structural knowledge. A number of other researchers have proposed that reasoning ability is the most important criterion for detecting systematic patterns among variables and to develop hypotheses about the causal structure of systems (Süß, Kersting, & Oberauer, 1991; Wittmann & Süß, 1999).

The SDDS model, in addition, predicts that the successful execution of these components is dependent on the problem solvers capacity to conduct appropriate experiments. In the context of scientific discovery learning, an appropriate experiment consists of a test that will control the influence of other variables in order to determine whether there is a causal relationship between two variables (Klahr & Dunbar, 1988). Similarly, in complex problem solving, the underlying structure of the system cannot be determined unless the input variables are changed systematically. Firstly, all variables should be held constant, in order to determine whether the outputs change independently. Then, each input should be varied individually, while the others are held constant, in order to determine how each input affects each output. This is referred to as the “Vary One Thing at a Time” (VOTAT) strategy (Tschirigi, 1980; Putz-Osterloh, 1993; Vollmeyer et al., 1996; Kröner et al., 2005). The empirical evidence supports the assumption that the abstraction of general rules with regard to the relationships between the variables in complex problems is causally dependent on the identification and use of the VOTAT strategy. Findings show that subjects who use the VOTAT strategy more frequently tend to acquire more structural knowledge than subjects who use it less frequently or consistently change all variables at the same time (Vollmeyer et al., 1996; Putz-Osterloh, 1993; Kröner et al., 2005). Most importantly, if subjects are instructed to use this strategy then they acquire more structural knowledge than subjects who do not receive this instruction (Vollmeyer et al., 1996). Thus, the SDDS model and empirical evidence converge on the hypothesis that a second significant source of individual differences in the acquisition of structural knowledge is likely to be the identification and use of the VOTAT strategy.

4.2.3 Conceptual and empirical overlap with traditional measures of intellectual ability

Kröner (Kröner, 2001; Kröner & Leutner, 2002; Kröner et al., 2005) argues that both reasoning ability and strategy identification are required in traditional measures of intellectual ability, and in particular, those that measure fluid intelligence such as the Raven’s Advanced Progressive Matrices (APM; Raven, 1958). However, while there is broad consensus that reasoning ability is central to performance on intelligence tests (Sternberg, 1986; Lohman, 2000), there seems to be little grounds for the claim that intelligence tests should also reflect individual

differences in strategy identification. Carpenter, Just and Shell's (1990) analysis of the requirements of the APM incorporated an extensive set of data derived from verbal protocols, eye-tracking and simulation models of performance. They found that "the processes that distinguish among individuals are primarily the ability to induce abstract relations and the ability to dynamically manage a large set of problem solving goals in working memory." (p. 404). Other studies have found further support for Carpenter et al.'s (1990) analysis of the processing requirements of the APM (e.g. DeShon, Chan & Weissbein, 1995; Embretson, 1998). These findings do not support a link between performance on traditional intelligence tests and strategy identification.

From a theoretical perspective, such a relation is not to be expected either. Klahr and Dunbar (1988) argue that the process that primarily distinguishes traditional rule induction tasks from scientific discovery is the need to perform experiments to generate data that can be used to test hypotheses. In traditional rule induction tasks, the entire set of instances that can be used to test hypotheses are given in the problem. In comparison, in scientific reasoning tasks, experiments must be designed and enacted by the problem solver. Similarly, in complex problems, the set of system states that are generated will be determined by how the user interacts with the system. This was the main motivation behind Klahr and Dunbar's (1988) extension of Simon and Lea's (1974) initial model of problem solving, as the original model did not specify a need to generate experiments to test hypotheses. These analyses suggest that while structural knowledge and performance on traditional tests of intelligence are likely to be strongly related due to the shared requirement to induce abstract relations, strategy identification may constitute an independent source of individual differences in knowledge acquisition.

Few studies, however, have examined the inter-relationships between strategy identification, structural knowledge and tests of fluid intelligence, and the results that are available are somewhat inconsistent. In two studies, Kröner (Kröner & Leutner, 2002; Kröner et al., 2005) found that structural knowledge was moderately correlated with tests of fluid intelligence, as assessed by the APM ($r = .43$) and the inductive reasoning scale of the Berlin Structure of Intellect Test ($r = .47$; BIS; Jäger, Süß & Beauducel, 1997). In contrast, Beckmann and Guthke (1995) found that structural knowledge and performance on the reasoning sub-tests of the Intelligence Structure

Test (IST; Amthauer, 1973) were only weakly related ($r = .11$). This is surprising considering that both studies operate within the linear structural equation framework, and presented the system with a novel cover story, in order to control for the influence of prior knowledge.

Similarly, the results regarding the relationship between strategy identification and performance on tests of fluid intelligence are also rather inconsistent. Kröner (2001) found no significant relationship between strategy identification and the APM ($r = .17$), while Kröner et al. (2005) found that strategy identification and the inductive reasoning scale of the BIS were moderately correlated ($r = .41$). Clearly, the inter-relationships between these constructs are in need of further investigation.

4.3 Individual differences in control performance

The situation with regard to identifying possible sources of individual differences in control performance is somewhat more complicated. Most theoretical analyses of the requirements of system control focus exclusively on the structural characteristics of complex problems (e.g. Dörner, Kreuzig, Reither & Stäudel, 1983; Putz-Osterloh, 1993b; Brehmer & Dörner, 1993; Rigas, Carling & Brehmer, 2002; Gonzalez, Thomas, & Vanyukov, 2005; Elg, 2005). These theorists argue that system control involves dealing with complex causal networks of inter-related variables, which must be uncovered by the problem solver while they deal with dynamic changes in the states of the system variables. An attempt is then made to map these demands onto those of traditional intelligence tests, and there is little agreement as to whether they are similar (Rigas, Carling & Brehmer, 2002; Gonzalez, Thomas, & Vanyukov, 2005) or dissimilar (Dörner et al., 1983; Putz-Osterloh, 1993b; Brehmer & Dörner, 1993; Elg, 2005).

The main problem with this approach is that, as Kluwe, Misiak and Haider (1991) argue, “subjects when operating complex systems elaborate highly different mental representations of the simulated complex environment, establish different goals for system control and apply different strategies” (p. 241). Thus, performance on the same complex problem does not necessarily reflect the same processes. This criticism can be generalised to comparisons of different studies, in which different constellations of task and structural characteristics determine whether structural knowledge will be acquired, and hence the strategies that are used to control the

outcomes of the system. This state of affairs is evident in the inconsistent results of previous research regarding the relationship between tests of intellectual abilities and control performance. While some studies report that the performance on the two types of tasks are highly related (e.g. Funke, 1985; Misiak & Kluwe, 1986; Krörner, 2001; Wagener, 2001; Krörner & Leutner, 2002; Körner et al., 2005; Bühner, et al., 2008), other studies report that they are not at all related (e.g. Dörner et al., 1983; Funke, 1983; Gediga, Schottke & Tuck-Bressler, 1984; Putz-Osterloh, 1981; Reichert & Dörner, 1988; Joslyn & Hunt, 1998). Clearly then, an approach simply based on an examination of the correlation between measures of intellectual ability and control performance across various studies is insufficient to identify consistent sources of individual differences in performance.

An alternative approach, which is the one adopted in this thesis, is to analyse the demands of control performance given certain conditions, in order to predict when we might reasonably expect intellectual ability to contribute to control performance. As discussed in Chapter 2, the use of ill-defined goals has significant consequences for the reliability and validity of control performance measures. Let us assume then that specific goals for the output variables are given. Secondly, in line with our discussion in Chapter 3, let us assume that the amount of structural knowledge that the problem solver acquires is a key determinant of the processes that will be used to control the outcomes of the system. Two extreme situations can be imagined with respect to the amount of structural knowledge that is acquired by the problem solver. In the first situation, the problem solver has been unable to acquire any structural knowledge. In the second situation, the problem solver is able to acquire or is instructed as to the complete underlying structure of the system. The following discussion will address the processing requirements of each of these situations in turn.

4.3.1 Individual differences in control performance under conditions of no knowledge

When no structural information, or relevant domain pre-knowledge, is available then problem solvers will not have any cognitive structures to apply to the task. Under these conditions, problem solvers may have to learn by doing, and control performance may be the result of ad hoc, or trial-and-error, processes. These

demands are not required in traditional psychometric tests of intellectual ability, and thus we should not expect to find any relationship between control performance and tests of intelligence under these conditions (Raaheim, 1985, 1989; Putz-Osterloh, 1993).

An alternative perspective is that if control behaviour is in fact random under conditions where no structural knowledge is acquired then it may not reflect any consistent source of individual differences. Again, control performance is not expected to correlate with measures of intelligence. In other words, zero correlations will not be informative as to whether knowledge acquisition took place (i.e. learning by doing) or not (i.e. random behaviour). One way to test this would be to compare the internal consistency of control performance scores under conditions where no knowledge has been acquired to conditions where complete structural knowledge has been acquired. The internal consistency of control performance scores should be low under conditions of no knowledge, if behaviour is in fact random. In comparison, the internal consistency of control performance scores under conditions of complete knowledge should be high, if problem solvers are repeatedly trying to apply the same rules to control the outcomes of the system.

The results of two studies provide preliminary support for the claim that the internal consistencies of control performance measures vary as function of the amount of structural knowledge that is acquired by problem solvers. Strohschneider (1986) found that control performance scores in MORO had test-retest coefficients in the range of .26 and .44, which indicates that the differences between individuals over time are rather unstable. In this study, subjects were required to reach specific goals from the start of the task and reported very little knowledge of the underlying structure of the system. In comparison, in Körner et al.'s (2005) study, subjects acquired at least half of the underlying structure of the system through a free exploration of the system variables prior to control performance, and in a later session were instructed as to the complete underlying structure of the system before controlling the task again. The internal consistency of the control performance scores was high in the two separate sessions ($\alpha = .90$ and $\alpha = .91$), and the correlation between control performance in the first and second session was also strong ($r = .65$). This indicates that the differences between people were highly stable within each testing session and between testing sessions. These results imply that under

conditions where no structural information is acquired or available, control performance may be random and is thus unlikely to reflect consistent sources of individual differences. In comparison, when structural knowledge is acquired, control performance appears to reflect stable sources of individual differences. However, further, more systematic research is required to adequately test this claim.

4.3.2 Individual differences in control performance under conditions of complete knowledge

In the alternative scenario, the problem solver has knowledge of the complete underlying structure of the system prior to control performance. Verbal protocols of successful problem solvers indicate that an efficient strategy for the application of structural knowledge entails a number of distinct steps. Firstly, problem solvers predict the next state of the output variables, under the assumption that the input variables are held constant. This accounts for the influence of autonomic changes in the output variables. Secondly, problem solvers then calculate the difference between the predicted states and goal states for each output variable. They then consider the dependency of each output on the inputs in turn and calculate the intervention required to bring each output to the desired state. Finally, they apply the interventions (Schoppek, 2002; 2004).

Schoppek (2004) argues that the consideration of the dependency of each output on the inputs during knowledge application requires a different perspective than that which is taken during knowledge acquisition. The acquisition of structural knowledge by systematically testing hypotheses encourages problem solvers to identify effects, given a particular cause. In contrast, knowledge application requires problem solvers to determine possible causes, given a particular effect (i.e. the desired goal state). In addition, the formulation of an intervention strategy to bring an output towards the desired goal state requires the problem solver to chunk the effect of multiple inputs together. Thus, structural knowledge must be translated into a different format before it becomes useful for system control (see also Beckmann, 1994, p. 82-86). In support of this hypothesis, Schoppek (2004) found that subjects who received instructions regarding the dependency of each output variable on the input variables had better control performance than those instructed as to the effect of each input variable on each output variable.

This suggests that there are three related sources of individual differences in control performance under conditions where complete structural information is acquired or available. Firstly, there may be differences in problem solvers capacity to transform effects-based knowledge into dependency-based knowledge. Secondly, problem solvers may have difficulty chunking such information into useable components. Finally and thirdly, problem solvers must co-ordinate this information in order to plan the correct sequence of interventions to bring about the desired goal state.

4.3.3 Conceptual and empirical overlap with traditional measures of intellectual ability

At a conceptual level, the processes involved in system control under conditions where complete structural is acquired have a clear overlap with the construct of working memory capacity. Oberauer, Süß, Schulze, Wilhelm, and Wittmann (2000) state that it is “generally assumed that working memory has a constrained capacity which acts as a limiting factor on performance in cognitive tasks, especially complex reasoning tasks” (p. 1018). They go on to argue that working memory has three main functional attributes. Firstly, working memory influences individuals’ capacity to manipulate and store information (Daneman & Carpenter, 1980). Secondly, it performs a supervisory function in monitoring task performance and inhibiting irrelevant responses (Baddeley, 1986; 1996). Thirdly, it coordinates and chunks information into structures (Halford, Wilson & Phillips, 1998; Oberauer, 1993). Thus, considering our task analysis, it seems likely that working memory capacity might contribute a significant source of individual differences in control performance when problem solvers attempt to apply their knowledge of the system to control its outcomes.

Studies indicate that performances on tasks designed to assess working memory capacity and control performance in complex problem solving are moderately correlated (Wittmann & Hatrup, 2004; Wittmann & Süß, 1999; Bühner, et al., 2008). The main problem with interpreting the results of these studies, however, is that subjects were required to independently acquire knowledge about the underlying structure of the tasks. As discussed, this results in the situation where subjects are rather heterogeneous in terms of the amount of knowledge that they

acquire. Therefore, this is likely to be an underestimation of the effect of working memory capacity on control performance under conditions where complete structural knowledge is acquired.

Further support for this claim can be found in studies that have investigated the relationship between fluid intelligence and control performance. Although there is an extensive and on-going debate as to whether working memory capacity and fluid intelligence are distinct constructs (e.g. Ackerman, Beier & Boyle, 2005; Oberauer, Schulze, Wilhelm & Süß, 2005), analyses of the processing requirements of tests of fluid intelligence, such as the APM, indicate that they place significant demands on working memory capacity (e.g. Carpenter, Just & Shell, 1990; DeShon, Chan & Weissbein, 1995; Embretson, 1998). Further, many studies have shown a strong relationship between measures of fluid intelligence and working memory capacity (e.g. Kyllonen & Christal, 1990; Kane, Hambrick, Tuholski, Wilhelm, Payne & Engle, 2004; Bühner, Krumm & Pick, 2005; Süß, Oberauer, Wittmann, Wilhelm & Schulze, 2002). Considering these results, and following our analyses of performance under different levels of structural knowledge, we would therefore expect that correlations between control performance and measures of fluid intelligence to be substantive when complete knowledge is acquired or instructed, and close to zero when no knowledge is acquired.

In line with this expectation, the results of a number of studies suggest that the relationship between control performance and measures of fluid intelligence varies according to the amount of structural information that is available to or was acquired by the problem solver. For example, the correlation between scores on the APM and control performance is moderate when subjects receive a structural diagram of the underlying structure of the system, yet in comparison, when no diagram is given, the reported correlation is close to zero (Putz-Osterloh, 1981; Putz-Osterloh & Lüer, 1981; Hörmann & Thomas, 1988). Similarly, Beckmann and Guthke (1995) found that the relationship between intellectual ability, as measured by the reasoning subtests of the IST (Amthauer, 1973) and control performance was moderate ($r = .36$) when a complex problem was presented with a novel cover story, and close to zero ($r = .03$) when the same complex problem was presented with a familiar cover story. Under conditions where the complex problem was presented with a novel cover story subjects acquired a significant amount structural knowledge, while no to little

knowledge was acquired when it was presented with a familiar cover story. These findings lend support to the claim that the processes used to control complex problems differ according to the amount of information that is available about the underlying structure of the system, and indicate that working memory capacity, as indicated by measures of fluid intelligence, may contribute significantly to individual differences in control performance under conditions where structural knowledge is available.

Thus far, we have discussed two extreme positions with respect to structural knowledge. One in which the problem solver fails to acquire any knowledge, and one in which the problem solver acquires complete knowledge of the underlying structure of the system. However, as discussed previously, problem solvers typically acquire only an incomplete representation of the underlying structure of the system. Under such conditions, it could be predicted that differences between people will still be attributable to differences in working memory capacity, if they are trying to apply their incomplete representations of the rules governing the systems' behaviour.

If different processes underlie control performance when different amounts of knowledge are acquired, this may explain why the relationship between control performance and traditional measures of intelligence is so inconsistent across different studies. The studies in this area can be broadly categorised in terms of whether subjects are given a specific goal from the outset of the task, or first allowed to acquire structural knowledge in the absence of a specific goal. As discussed in Chapter 3, if subjects receive a specific goal, then they are unlikely to acquire structural knowledge. In comparison, if subjects are required to explore the system in the absence of a specific goal, then they generally acquire at least some knowledge of the underlying structure of the system (Beckmann, 1994; Geddes & Stevenson, 1997; Vollmeyer et al., 1996; Burns & Vollmeyer, 2002). Typically, in studies where a specific goal is given from the outset of the task, no or only a weak relationship is found between control performance and traditional measures of intelligence (e.g. Dörner, Keuzig, Reither & Stäudel, 1983; Funke, 1983; Gediga, Schottke & Tuck-Bressler, 1984; Putz-Osterloh, 1985; Reichert & Dörner, 1988; Joslyn & Hunt, 1998). There are, however, a few studies that report moderate correlations between intellectual ability and control performance, even when subjects did not initially explore the complex problem (Rigas, 2000; Rigas, Carling & Brehmer, 2002; Süß,

1999; Wittmann & Süß, 1999; Wittmann & Hattrup, 2004). In comparison, in studies where subjects are first allowed to explore the system, a moderate to strong correlation is found between control performance and these same measures of intelligence. In addition, these studies report that subjects acquired at least some knowledge of the underlying structure of the system (Funke, 1985; Misiak & Kluwe, 1986; Krörner, 2001; Wagener, 2001; Kröner & Leutner, 2002; Krörner et al., 2005; Bühner, et al., 2008). Although this is to some extent an oversimplification of the differences between these studies, the overall pattern of results does suggest that the demands of traditional measures of intellectual ability and control performance are similar only when control performance is based on previously acquired knowledge.

This pattern of results must be interpreted cautiously, however, as the relationship between fluid intelligence and control performance may be mediated by the amount of knowledge that is acquired by the problem solver. As discussed, reasoning ability is likely to be a significant source of individual differences in the acquisition of structural knowledge through a goal-free exploration of the system structure. The quality of problem solvers' control performance then becomes an indirect function of the amount of knowledge that is acquired. Therefore, the question that remains is whether measures of intellectual capacity account for any additional variance in control performance, once the influence of structural knowledge has been taken into account.

4.4 Summary and conclusions

The application of the SDDS model to describe the acquisition of structural knowledge has identified two possible sources of individual differences in performance: Strategy identification and reasoning ability. A question is the extent to which the processes underlying traditional tests of fluid intelligence and strategy identification overlap. Theoretical analyses suggest that they should be distinct constructs, although the empirical evidence is inconclusive. Secondly, if the acquisition of structural knowledge is dependent on the identification and application of an efficient strategy, the question is then whether there is remaining variance in structural knowledge that can be accounted for by fluid intelligence. As yet, this question has not been empirically tested.

With regard to possible sources of individual differences in control performance, our task analyses suggest that the relationship between intelligence and control performance may be moderated by the amount of knowledge that is acquired by problem solvers. Under conditions where no knowledge is acquired, control performance may be random, or the result of ad hoc interventions. As yet it has not been established whether the differences between individuals under such conditions are consistent (i.e. if there is anything systematic to be explained), however, it is predicted that performance under such conditions should not be related to measures of intelligence. Alternatively, when complete structural knowledge is acquired, we expect that significant demands are made on working memory capacity, as problem solvers are required to simultaneously store and process information and coordinate between various sources of information and sub-goals. Therefore, it is predicted that under such conditions performance should be strongly related to measures of intelligence. Similarly, under conditions where only partial knowledge is acquired, we expect that differences between people will still be attributable to differences in intelligence, if they are trying to apply their incomplete representations of the rules governing the systems' behaviour.

As yet, these predictions have not been sufficiently tested for the following reasons. Firstly, in previous studies direct estimates of the impact of intelligence on control performance have not been obtained due to carry over effects of individual differences in knowledge acquisition. Secondly, few studies report the reliability of control performance measures. Therefore, as yet it is unclear whether the sources of individual differences in control performance differ according to the amount of knowledge that is available to problem solvers.

The empirical work reported in the subsequent chapters will address these challenges in the following ways. Firstly, in order to control for the effect of the individual differences in knowledge acquisition, the amount of knowledge available to problem solvers will be manipulated experimentally. This should allow us to determine whether the relationship between intelligence and control performance varies as a function of the amount of knowledge available to problem solvers. Secondly, the reliability of control performance measures will be examined across different trials and goal states under different levels of knowledge. This should allow us to determine whether the differences between individuals are systematic under

different levels of knowledge. Thirdly, and finally, we will examine whether there is still remaining variance in control performance that can be explained with respect to differences in intellectual ability, once differences in structural knowledge have been taken into account.

CHAPTER FIVE

STUDY 1: THE EFFECT OF STRUCTURAL INFORMATION ON THE CONTROL OF DYNAMIC SYSTEMS

5.1 Introduction

5.1.1 The effect of structural information on control performance

The acquisition of knowledge through an unguided exploration of a system and its interrelated variables can be characterised as discovery learning. In this approach, the learner is seen as an independent and active agent in the process of knowledge acquisition. In order to acquire a mental model of the underlying structure of the system they must develop hypotheses, design experiments to test them and appropriately interpret the data (see Chapter 4).

The problems that learners experience with discovery learning in hyper-media and computer simulations of conceptual domains are well documented, and there is no clear evidence that favours discovery learning over more traditional forms of learning such as expository instruction. Reviews of the findings of numerous studies show that learners need extensive guidance in order to facilitate the acquisition of deep conceptual knowledge (de Jong & van Joolingen, 1998; Mayer, 2004; de Jong, 2005; 2006; Kirschner, Sweller & Clark, 2006). Overall, unguided discovery learning is considered to be insufficient for the purposes of instruction and training, especially when it involves novice learners (Mayer, 2004).

Similarly, in research with complex problems, it has been found that most problem solvers are unable to acquire a complete or accurate representation of the underlying structure of a system through an unguided exploration of its variables (Funke & Müller, 1988, Müller, 1993; Beckmann, 1994; Beckmann & Guthke, 1995; Vollmeyer et al., 1996; Burns & Vollmeyer 2002; Kröner, 2001; Schoppek, 2002; Kröner et al., 2005; Kluge, 2008; Osman, 2008a). These studies also report a consistent positive relationship between the amount of structural knowledge that is acquired and the quality of problem solvers' control performance. It might be expected then, that the direct instruction of structural information should result in better control performance than an unguided exploration of the system variables,

because problem solvers will acquire more knowledge about the underlying structure of the system.

Against expectations, as discussed in Chapter 3, attempts to promote control performance through the direct instruction of structural information have proved unsuccessful (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996). Findings show that structural information only appears to benefit control performance after a period of goal-orientated practice (Putz-Osterloh, 1993; Preußler, 1998). This suggests that declarative knowledge about the underlying structure of a system can only be translated into procedural knowledge about how to control a system through practice.

However, Putz-Osterloh's (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993) and Preußler's (1996; 1998) studies have a number of instructional and methodological shortcomings that call this conclusion into question. From an instructional design point of view it is quite possible that subjects in Putz-Osterloh's (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993) studies may not have understood how a causal diagram depicted on paper related to input and output variables presented on a computer. Subjects in Preußler's (1996) study may not have been able to remember the information provided in the instructional phase when they were later required to control the outcomes of the system. In effect, problem solvers may not have understood the information that they were provided with, or have been able to remember it during the control task.

From a methodological point of view, one limitation of Preußler's (1998) study is that subjects who were provided with a diagram had more opportunities to interact with the system, and therefore independently induce the underlying structure of the system, than those who did not receive the diagram. All subjects completed the practice tasks. However, only those with the diagram were required to repeat these tasks until they found the correct answer. Clearly, this setting provided subjects with an opportunity to acquire structural knowledge and evaluate the accuracy of this knowledge. It could be speculated that they may not have used the diagram at all. Therefore, the results of Preußler's (1998) study do not clearly show that practice is necessary in order to effectively utilise structural information (see Chapter 3 for a detailed critical review of these issues).

With the current study the aim is to address some of the identified shortcomings in order to determine whether structural information can directly benefit control performance without practice, or provide an advantage over incomplete knowledge that is acquired through an unguided exploration of the system variables. In the proposed design, subjects will first explore a complex, dynamic system without guidance and try to acquire knowledge about its underlying structure. They will then try to control the system to reach specific goal values of the output variables. Subjects in an experimental condition will then watch an instructional video that explains the underlying structure of the system. Both the experimental and the control groups will then control the system again. If a period of practice is necessary in order to utilise structural information, then subjects who watch the instructional video should not show an improvement in their control performance, and should not be better at controlling the system than subjects who do not receive information.

In order to ensure the effectiveness of the structural information, the instructional material was designed in accordance with the principles of Cognitive Load Theory (CLT) (e.g. Sweller, 1994; 1999; for recent reviews see Beckmann, 2010; Sweller 2010). In the design of the video, the aim was to reduce the number of cognitive activities that subjects would have to undertake to translate the information provided into knowledge about the system. In particular, a number of studies have shown that cognitive effort is reduced, and learning is facilitated, when explanations of visual material are presented aurally, rather than as text (Kalyuga, Chandler & Sweller, 1999, 2000; Tindall-Ford, Chandler & Sweller, 1997; Tabbers, Martens & Van Merriënboer, 2004). This is known as the *modality effect* (Sweller, Van Merriënboer & Pass, 1998) or the *modality principle* (Mayer, 2001). In addition, findings show that information that is permanently available, or provided as it is needed, is much more effective in promoting the acquisition of knowledge than information that is given before the interaction with the task begins (Kotovsky, Hayes & Simon, 1985; Berry & Broadbent, 1987; Leutner, 1993; Hulshof & De Jong, 2006). In accordance with these findings, in the instructional video, visual material that shows how each input effects each output is explained by a narrator, and a causal diagram depicting this information will remain on screen during the control task as an external memory aid. These improvements should reduce the

extraneous cognitive effort that may be associated with understanding and remembering structural information.

5.1.2 Individual differences in the acquisition and application of knowledge

Even when the design of instructional material is optimised to reduce cognitive effort, the cognitive abilities of the problem solver may still influence whether such information can be effectively utilised. As discussed in Chapter 4, the application of structural information requires problem solvers to transform effects-based knowledge into dependency-based knowledge, chunk this knowledge into useable components, and co-ordinate this information in order to plan the correct sequence of interventions to bring about the desired goal state (Schoppek, 2002; 2004). At a conceptual level these processes have a clear overlap with the construct of fluid intelligence, and we would therefore expect that correlations between control performance and measures of fluid intelligence to be substantive when complete knowledge is acquired or instructed. In line with this expectation, previous findings show that when a causal diagram is provided, control performance is moderately to strongly correlated with scores on the APM, which is a key marker test of fluid intelligence (Putz-Osterloh, 1981; Putz-Osterloh & Lüer, 1981; Körner et al., 2005). This seems particularly relevant to studies that report that the provision of structural information has no effect on control performance (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996), as it could be speculated that the positive effect of structural information on control performance may be contingent on the subjects' level of fluid intelligence. That is, subjects who are more intelligent may be able to make use of structural information, while the provision of structural information may have little, or even a negative, impact on control performance for subjects who are less intelligent.

Therefore, in the current study it is predicted that under conditions where subjects receive structural information, the extent of improvements in control performance will be a function of their fluid intelligence. In comparison, the extent of improvements in control performance when subjects do not receive additional information should be due to practice applying their partial representations of the underlying structure, and therefore less strongly related to fluid intelligence.

There are also likely to be individual differences in the extent to which problem solvers can acquire knowledge through an unguided exploration of the system variables. As discussed in Chapter 4, Klahr and Dunbar's (1988) model of discovery learning implies that effective knowledge acquisition involves a) the design of experiments that are appropriate for abstracting general rules from the data or testing hypotheses, b) making predictions based on abstracted rules about the values of the outputs given certain inputs and c) the abstraction of general rules from the data. In the context of complex problems the pattern of interventions that are necessary to generate data that can be used to abstract rules or test hypotheses is known as the VOTAT strategy. As argued in Chapter 4, the capacity to identify and use the VOTAT strategy is unlikely to be reflected in tests of fluid intelligence. Alternatively, b) and c) can be categorised as types of reasoning, and hence are likely to be reflected by tests of fluid intelligence. Therefore, it is predicted that the use of the VOTAT strategy should be a significant predictor of the amount of structural knowledge that subjects are able to acquire through an unguided exploration of the system variables, in addition to fluid intelligence.

5.1.3 Aims and hypotheses

In summary, the aims of this study are to determine whether the provision of structural information confers any additional advantage in controlling a dynamic system over knowledge that is acquired through an unguided exploration of the system variables, and to investigate possible sources of individual differences in the acquisition and application of structural knowledge. Firstly, it is hypothesised that VOTAT strategy use should account for additional variance in the acquisition of structural knowledge over fluid intelligence (VOTAT Hypothesis). Secondly, it is hypothesised that subjects who acquire more structural knowledge through an unguided exploration of the variables should show better control performance than those who acquire less knowledge (Knowledge Hypothesis). Thirdly, subjects who receive structural information should improve their control performance more than those who receive no additional structural information (Information Hypothesis). Finally, under conditions where subjects receive structural information, the improvement in their control performance scores will be a function of their fluid intelligence. Under conditions where subjects do not receive structural information,

the improvement in their control performance scores will be a result of practice, and therefore less related to fluid intelligence (Intelligence Hypothesis).

5.2. Method

5.2.1 Subjects

Ninety-eight first year psychology students at the University of Sydney participated for course credit. Nine subjects failed to complete all tasks therefore their data were excluded from further analysis. A sample size of 100 would guarantee sufficient statistical power ($1 - \beta \geq .80$) in identifying at least medium effects ($d = .50$) at a significance level of $\alpha \leq .05$ (one-tailed) in the planned comparison between the partial and no information conditions.

5.2.2 Design

Subjects were randomly assigned to one of two conditions (45 subjects in the *Information* condition, 44 subjects in *No Information* condition). As subjects were required to control the system on two separate occasions this resulted in a (2) x 2 design. The within-subjects factor was control performance (phase 1 and phase 2). The between-subjects factor was whether or not they received structural information (Information and No Information). Subjects were assessed on their VOTAT strategy use, the amount of knowledge that they acquired through an unguided exploration of the system variables (structural knowledge), control performance for phase 1, control performance for phase 2 and performance in a test of fluid intelligence. Figure 5.1 displays the procedure of the experiment for each condition and indicates which performance measures were collected in each phase of the experiment.

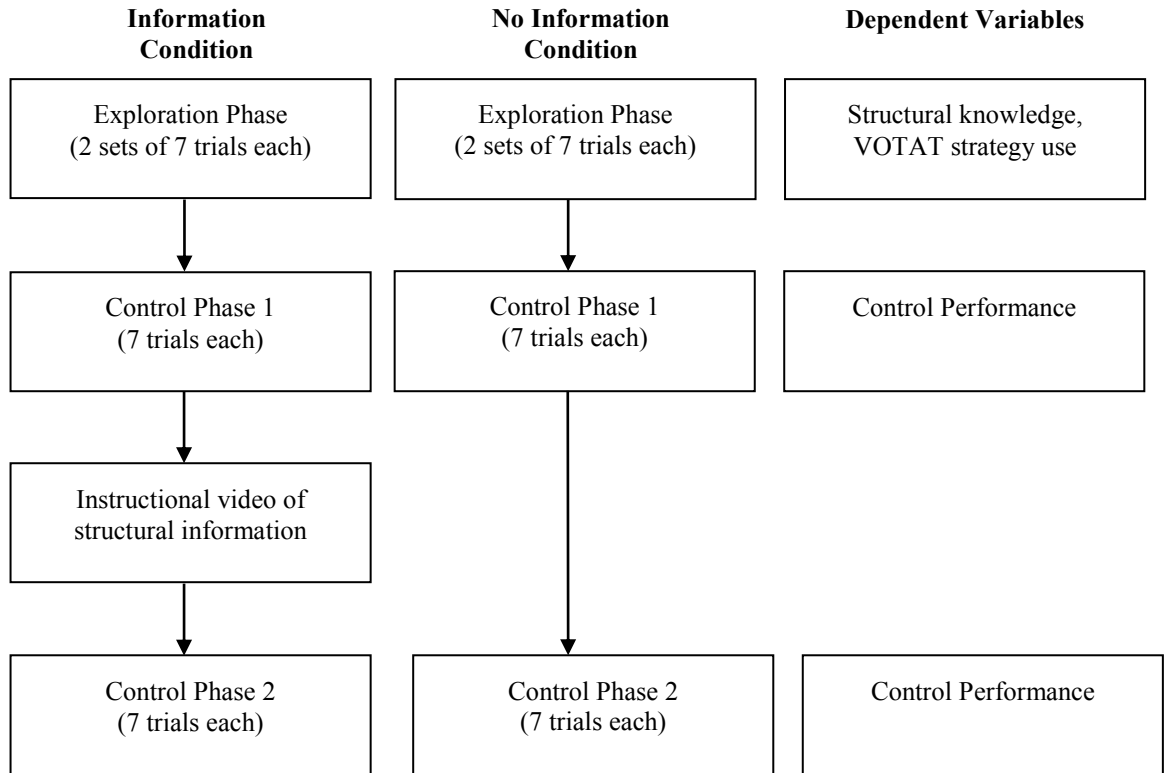


Figure 5.1: Diagram for the procedure of the experiment, illustrating the phases of the experiment by condition and indicating which performance measures were collected in each phase.

5.2.3 Dependent variables and individual differences measures

5.2.3.1 VOTAT strategy use

To determine whether subjects used the VOTAT strategy during the exploration phase, the number of trials on which one or less input at a time was varied was recorded. VOTAT scores range between 0 and 14, with the higher value implying a more frequent use of the VOTAT strategy.

5.2.3.2 Structural knowledge

Subjects' structural knowledge was assessed by asking them to create causal diagrams of the relationships between the input and output variables at the end of each trial during the Exploration Phase. The diagram that was generated on the final exploration trial (after 2 times 7 trials) was used to derive a structural knowledge score. Using a procedure introduced by Beckmann (1994), the operationalisation of structural knowledge is based on a threshold model for signal detection (Snodgrass & Corwin, 1988). Using the procedure, structural knowledge scores are corrected for

guessing by subtracting the false alarm rate from the hit rate. The hit rate is the number of correctly identified relations divided by six (the actual number of relations in the system). The false alarm rate is the number of incorrectly identified relations divided by six (the number relations that are not present in the system, but that could possibly exist given the number of variables). The final score has a theoretical range from -1 to 1, where a score below zero indicates inaccurate knowledge and a score above zero indicates accurate knowledge.

5.2.3.3 Control Performance

The scoring procedure used was based on Beckmann's (1994) scoring system (see Chapter 2, pp.33 - 35, for a discussion of the relative merits of different scoring procedures for control performance and for a detailed example of how Beckman's scoring procedure is calculated). Control performance was calculated by determining the Euclidean Distance between the actual and optimal values of the input variables. The ideal values for each input variable were calculated by using the values of the output variables on the previous trial and the goal output values to solve the set of linear equations underlying the system⁴. The theoretical range of this score is from 0 to 34, where a lower score indicates a smaller deviation from optimal control interventions and therefore better performance.

5.2.3.4 Fluid Intelligence

The Raven's Advanced Progressive Matrices (APM) was used as an indicator of fluid intelligence. Raw scores were transformed into percentage correct. This test has been extensively validated as an indicator of fluid intelligence for a university level population (Raven, Raven & Court, 1998).

5.2.4 Materials

The complex problem was programmed using Adobe Flash 8 and Captivate 3, and administered via a web browser on PCs.

As discussed in Chapter 2, previous research has found that the presence of a familiar context has an unpredictable effect on acquisition of structural knowledge

⁴ If the ideal values were within the range of possible input values (which was 10 to -10), then the ideal values were equal to the optimal values. In cases where the ideal value fell outside this range, then the optimal value was adjusted to the nearest possible value.

(Beckmann, 1994; Burns & Vollmeyer, 2002; Lazonder et al., 2008; Lazonder et al., 2009). Therefore, in order to ensure that the complex problem was relatively novel for all subjects and thus control for the influence of prior knowledge, the input and output variables are neutrally labelled with letters. As can be seen in Figure 5.2, the input variables are labelled A, B, and C, while the output variables are labelled X, Y and Z.

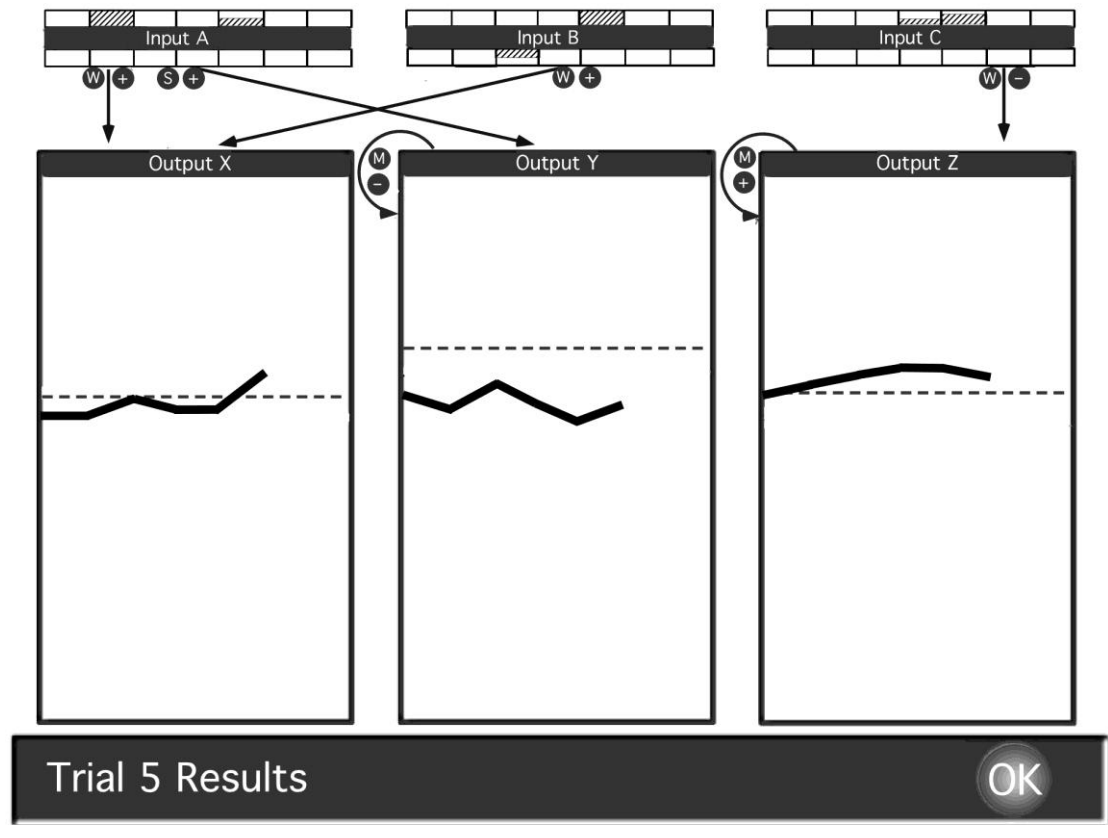


Figure 5.2: Screenshot of the task, as presented in the information condition after the instructional phase. The goals are indicated as dotted lines on the graphs for the output variables. The underlying structure of the system is represented on screen as a causal diagram, where the arrows represent the relationships between the variables, while the positive and negative signs denote the direction of the relationship, and the letters the relative strength. In this example, Input A, B and C were increased on Trial 5. As a result, Output X increased, Output Y increased, and Output Z decreased, as depicted in the output variable windows.

The user-interface is in a non-numerical graphical format, in order to encourage the formation of mental representations that are more aligned with the development of causal diagrams. In accordance with the principles of CLT, this should minimise the cognitive activities that are not directly relevant to the task. Figure 5.2 shows that the values of the input variables are displayed as bars in the

boxes on the input variables, where positive values are shown above the input label and negative values are shown below. Each box represents the value of the input variable on a single trial and in total seven trials can be conducted before the values are reset. Although the numerical values of the inputs are not available to subjects, the inputs can be varied in increments of one unit, within the range of -10 to 10.

The underlying structure of the system was originally developed by Beckmann (1994), and is based on the Linear Structural Equation approach developed by Funke (1985; 1993; 2001) (see Chapter 2). It consists of three input and three output variables that are connected by a set of linear equations:

$$X_{t+1} := 1.0 * X_t + 0.8 * A_t + 0.8 * B_t + 0.0 * C_t$$

$$Y_{t+1} := 0.8 * Y_t + 1.6 * A_t + 0.0 * B_t + 0.0 * C_t$$

$$Z_{t+1} := 1.2 * Z_t + 0.0 * A_t + 0.0 * B_t - 1.0 * C_t$$

X_t , Y_t , and Z_t denote the values of the output variables and A_t , B_t , and C_t denote the values of the input variables during the present trial while X_{t+1} , Y_{t+1} , Z_{t+1} denote the values of the output variables in the subsequent trial. The system is dynamic because the values of the output variables change as a result of the decisions made by the subject and independently on each trial.

5.2.5 Procedure

The complex problem and the APM were presented to subjects on PCs, over two separate sessions.

The complex problem began with a set of instructions that explained to subjects that they were required to perform three tasks. Firstly, they would have to explore the system to discover its underlying structure (Exploration Phase). In two subsequent tasks, they would have to control the system to reach certain goal values of the output variables (Control Phase 1 and 2). Further instructions explained how the values of the input and output variables were represented in the user-interface, how to change the values of the input variables and how to record what they learnt during the Exploration Phase using the causal diagram.

The Exploration Phase then began, in which subjects were prompted to explore the system for two cycles of 7 trials each by changing any of the input variables and observing the effect on the output variables displayed in the graphs. At the end of each trial, subjects had to record what they had learnt about the system using the causal diagram that was displayed on the screen.

The causal diagram could be altered using a set of twelve buttons (one for each possible relationship in the system) at the bottom of the screen. Each button referred to a particular relationship in the system. Using these buttons, subjects could record if they thought there was a relationship between two variables or if they thought the output variables changed independently. They could also specify the direction of the effect and its perceived strength. The buttons were designed to restrict subjects' search of the hypothesis space to the relationships that could potentially exist within the system (i.e. autonomic changes of the output variables and direct effects between input and output variables).

Subjects then had to control the system by manipulating the inputs to reach goal values of the outputs for seven trials (Control Phase 1). The goals were indicated as dotted lines on the output graphs. The causal diagram they had constructed during the Exploration Phase remained on screen, providing access to the structural knowledge that they had acquired.

In the information condition, subjects then watched the instructional video that explained the actual underlying structure of the system. The instructional video consisted of a recording of seven intervention trials with an accompanying narration, which explained what could be seen on screen during each trial. On the first trial the inputs were set at zero, so that the autonomic changes in the outputs could be detected. Subsequently, Input A was increased to maximum while the other inputs were set at zero, then on the next trial Input A was reduced to minimum while the other inputs were set at zero. This was repeated with Inputs B and C on subsequent trials so that the effect of each input on the outputs could be clearly observed. On each trial, the narrator explained how the inputs had been altered, how each of the outputs had changed and how this reflected the underlying structure of the system. A causal diagram was constructed on screen, to record this information, and it remained onscreen during Control Phase 2 (see APPENDIX A for a record of these

instructions). Figure 5.2 shows a screenshot of the complex problem in the information condition after the instructional video.

Subjects in the no information condition did not receive any additional information. In this condition, the causal diagram that subjects had constructed during the Exploration Phase remained on screen during Control Phase 2.

All subjects had to control the system again for seven trials, with different goals indicated on the output variables (Control Phase 2).

In a subsequent session, approximately one week later, subjects completed the APM.

5.3. Results

5.3.1 Internal consistencies

Firstly, internal consistency analyses were conducted to determine the variability in control performance scores across the trials and for different goal states as an estimate of the reliability of the dependent variables. Internal consistency was good across the first control phase ($\alpha_{\text{information}} = .83$, $\alpha_{\text{no information}} = .83$) and the second control phase ($\alpha_{\text{information}} = .92$, $\alpha_{\text{no information}} = .90$) (Cronbach, 1951). This indicates that subjects are rather consistent in their performance and justifies averaging the scores across each control phase.

An additional internal consistency analysis indicated that the reliability of the APM scores was acceptable across the 36 items ($\alpha = .75$) (Cronbach, 1951).

Correlations among the variables used in this study as well as their means and standard deviations are presented in Table 5.1.

Table 5.1: Descriptive statistics, and inter-correlations between the variables for each condition.

		<i>M (SD)</i>	2	3	4	5
No Information Condition <i>N</i> = 44	1.VOTAT strategy	5.68 (4.15)	.60**	-.45**	-.49**	.11
	2. Knowledge	.21 (.33)	...	-.38*	-.51**	.45**
	3. Phase 1	13.88 (4.20)56**	-.24
	4. Phase 2	13.21 (5.28)	-.18
	5. APM (% correct)	63.75 (16.32)
Information Condition <i>N</i> = 45	1.VOTAT strategy	5.76 (3.94)	.60**	-.20	-.30	.26
	2. Knowledge	.22 (.35)	...	-.31*	-.36*	.26
	3. Phase 1	13.92 (4.14)27	-.11
	4. Phase 2	10.24 (5.29)	-.52**
	5. APM (% correct)	58.00 (17.75)

Notes: * $p < .05$; ** $p < .01$

5.3.2 VOTAT hypothesis

Across the conditions, there was a significant strong correlation between VOTAT strategy use and structural knowledge; $r = .60$, $p < .01$. This indicates that subjects who used the VOTAT strategy more frequently acquired more knowledge of the underlying structure of the system during the Exploration Phase than those who used it less frequently. Further, there was a significant moderate correlation between scores on the APM and structural knowledge scores across the conditions; $r = .34$, $p < .01$. This indicates that subjects who were more intelligent were able to acquire more knowledge of the underlying structure of the system than those who were less intelligent. Finally, as predicted, VOTAT strategy use and scores on the APM were not significantly correlated; $r = .18$, $p = .09$. This indicates that the use of effective experiments in order to infer abstract rules is not a matter of fluid intelligence.

In order to determine whether VOTAT strategy use accounts for additional variance in structural knowledge over fluid intelligence, a hierarchical regression analysis was conducted. Mean corrected scores on the APM were entered in the first step, and VOTAT strategy use in a second step. Fluid intelligence accounted for 11.3% of the variance in structural knowledge scores, $F_{change}(1, 87) = 11.09$, $p < .01$. VOTAT strategy use accounted for an additional 30% of the variance in structural knowledge scores, $F_{change}(1, 86) = 43.90$, $p < .01$. In support of the VOTAT hypothesis, this indicates that although differences in fluid intelligence are important,

the largest source of individual differences in the acquisition of structural knowledge is the use of effective experiments.

5.3.3 The acquisition of structural knowledge during the exploration phase

For both information conditions (no information and information), the amount of structural knowledge that was acquired was significantly greater than zero; $M = .22$, $SD = .34$, $t(88) = 6.00$, $p < .01$. This indicates that on average, subjects had acquired some knowledge of the underlying structure of the system prior to the first control phase. However, as the histograms in Figure 5.3 show, the range of structural knowledge scores, $-.50$ to 1.00 , indicates that subjects differed widely in the amount of knowledge that they were able to acquire about the underlying structure of the system during the initial exploration phase. That is, while some subjects were able to acquire complete knowledge of the underlying structure of the system (one subject in the no information condition, and two subjects in the information condition), others acquired an incorrect representation of the underlying structure. Overall these results indicate that subjects found it difficult to acquire an accurate representation of the underlying structure of the system through an unguided exploration of the system variables, and that for the majority of subjects in the information condition the instructions given prior to the second control phase provided a significant source of new information about the underlying structure of the system.

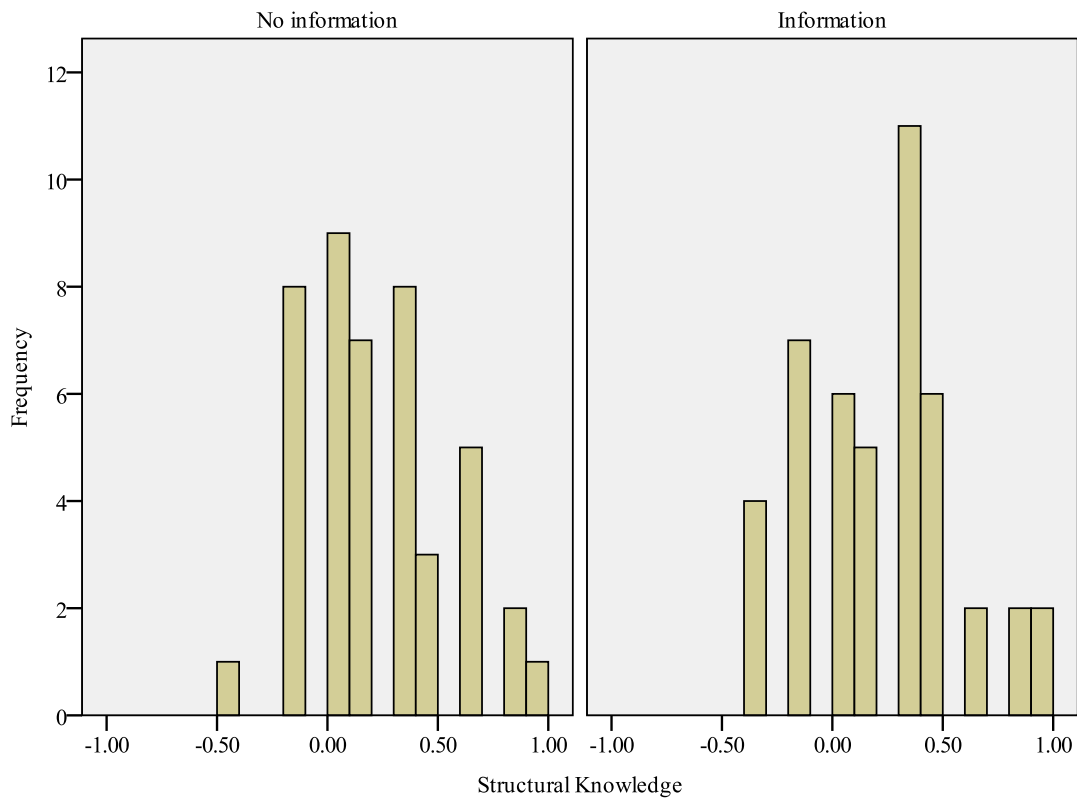


Figure 5.3: Histograms of structural knowledge test scores, by information condition

5.3.4 Knowledge hypothesis

In support of the knowledge hypothesis, across the conditions, there was a significant moderate negative relationship between structural knowledge scores and control performance in Phase 1 ($r = -.34, p < .01$). This indicates that subjects who acquired more knowledge about the underlying structure of the system through an unguided exploration of the system variables produced smaller deviations from optimal control interventions, and were therefore better at controlling the system. More knowledge is associated with better control performance.

5.3.5 Information and intelligence hypotheses

In order to determine whether the provision of structural information facilitates control performance (Information Hypothesis) and whether the extraction of knowledge from information in this context is determined by fluid intelligence (Intelligence Hypothesis) we conducted a series of hierarchical linear modelling analyses using the HLM software package (Raudenbush, Bryk, Cheong, & Congdon,

2000). This approach allowed us to model subject's change in performance from control phase 1 to control phase 2 as function of their fluid intelligence (see Raudenbush & Bryk, 2002). We used a two level model in which performance in control phase 1 and control phase 2 (level 1) was clustered within people (level 2).

Firstly, to examine the effect of condition (no information and information) and fluid intelligence on control performance it was necessary to check whether control performance under different conditions differed prior to the instructional phase. The amount of structural knowledge acquired by subjects during the exploration phase did not differ by condition; $t(87) = -.09, p = .93$, nor did their control performance scores in phase 1; $t(87) = -.05, p = .96$, or scores on the APM; $t(87) = 1.59, p = .12$. This demonstrates that the procedure used to randomly allocate subjects to the conditions was effective.

Secondly, a random coefficient regression analysis was conducted to assess whether control performance changed across the two control phases. At level 1, each subjects' performance was represented by an intercept term that denoted their mean performance across control phase 1 and control phase 2, and a slope that represented their change in performance from control phase 1 to 2. Control phase (1 or 2, effect coded as $-.5$ and $.5$, respectively) was entered as an independent variable at this level. The mean control performance scores and the change in control performance then became the outcome variables in a level-2 model, in which they were modelled as random effects. The results of this analysis are presented in the top section of Table 5.2. This analysis indicated that the mean control performance score was 12.81 across control phase 1 and 2 and on average, control performance scores improved by 2.19 points from control phase 1 to 2. The change in control performance was significantly different from zero; $t(88) = -3.86, p < .001$. There were also significant differences between problem solvers in terms of their mean control performance scores and the change in their control performance; $\chi^2 = 2867895259.55, df = 88, p < .001$ and $\chi^2 = 1274245149.43, df = 88, p < .001$, respectively. Variability in subjects' change in control performance from control phase 1 to 2 accounted for 64% of the total variability in control performance scores. These findings are an important prerequisite for the subsequent analyses, as they indicate that individuals show substantial variability in their mean control performance and the extent to which their control performance changed across the two phases.

Finally, in order to test the Information and Intelligence hypotheses an intercept- and slope-as-outcomes regression analysis was conducted in which mean control performance and the change in control performance from control phase 1 to 2 were modelled as a function of condition (as an effect coded variable indicating condition: -.5 = no information, .5 = information) and scores on the APM at level 2. The level 1 model was the same as in the random coefficients regression analysis. The results of this analysis are presented in the middle panel of Table 5.2, and are reported in relation to the Information and Intelligence hypotheses in the next sections.

5.3.5.1 Information hypothesis

The intercept- and slope-as-outcomes regression analysis indicated that information had a significant impact on average control performance scores as well as on the change in control performance from control phase 1 to 2, controlling for the effects of fluid intelligence; $t(86) = -2.33, p < .05, \Delta R^2 = 5\%$ and $t(86) = -3.13, p < .01, \Delta R^2 = 9\%$, respectively. In the second control phase, subjects in the information condition had an average control performance score 1.91 points better than those in the no information condition. Similarly, the change in control performance from Phase 1 to 2 for subjects in the information condition was 3.43 points higher than those in the no information condition. In support of the information hypothesis, as can be seen in Figure 5.4, these results indicate that subjects who received additional information with regard to the underlying structure of the system performed better in the second control phase, and improved at a greater rate from control phase 1 to control phase 2 than those who did not receive information.

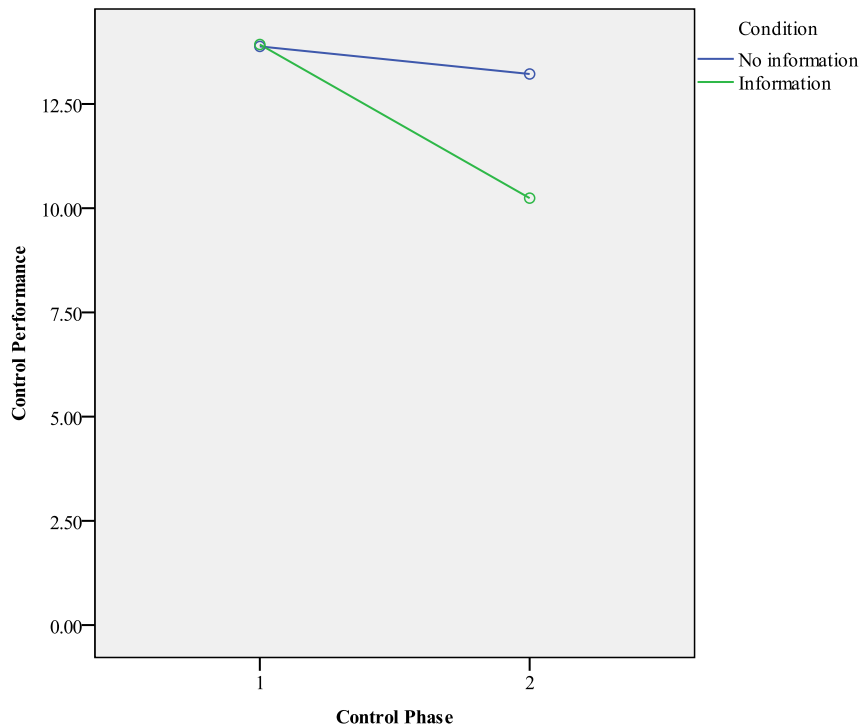


Figure 5.4: Change in control performance scores from Phase 1 to 2, by information condition

5.3.5.2 Intelligence hypothesis

The intercept- and slope-as-outcomes regression analysis also indicated that subjects scores on the APM had a significant impact on average control performance scores and their change in performance from control phase 1 to 2, controlling for the effects of condition; $t(86) = -3.42, p < .01, \Delta R^2 = 7\%$ and $t(86) = -2.34, p < .05, \Delta R^2 = 2\%$. On average, a one-point increase in scores on the APM was associated with a .07 increase in average control performance, and a .06-point increase in the change in performance scores from control phase 1 to 2. These results indicate that subjects with a higher score on the APM, tended on average to perform better overall, and improved more from control phase 1 to 2.

In order to determine whether the effect of fluid intelligence on control performance differed by condition a third analysis was conducted in which an interaction term (APM x Condition) was added to the main effects of the variables at level 2. The results are presented in the bottom panel of Table 5.2. There was no evidence that the effect of fluid intelligence (as measured via APM scores) on mean control performance scores varied by condition, as the interaction term was small and insignificant; $t(85) = -.03, p = .50, \Delta R^2 = 0\%$. However, the effect of fluid

intelligence on the *change in performance* from control phase 1 to control phase 2 did vary significantly by condition; $t(85) = -2.48, p < .05, \Delta R^2 = 3\%$. In support of the intelligence hypothesis, the results show that the change in performance scores for subjects who received information was more strongly related to fluid intelligence than for subjects who did not receive information.

Table 5.2 Results of the Random Coefficients Regression (RCR) Analysis and the Intercept- and Slope-As-Outcome Regression (ISAOR) Analyses

Variable	Parameter Estimate	SE	t	ΔR^2
RCR Analysis				
Mean control performance (β_{00})	12.81	.43	30.10**	
Mean change in control performance (β_{10})	-2.19	.57	-3.86**	
ISAOR Analysis 1				
Intercept-as-outcome				
Condition (β_{01})	-1.91	.82	-2.32*	5%
APM (β_{02})	-.08	.02	-3.21**	7%
Slope-as-outcome				
Condition (β_{11})	-3.42	1.07	-3.19**	9%
APM (β_{12})	-.07	.03	-2.17*	2%
ISAOR Analysis 2				
Intercept-as-outcome				
Condition (β_{01})	-1.89	.82	-2.33*	5%
APM (β_{02})	-.08	.02	-3.32**	7%
Condition x APM (β_{03})	-.03	.04	-.68	0%
Slope-as-outcome				
Condition (β_{11})	-3.38	1.03	-3.29**	9%
APM (β_{12})	-.06	.03	-2.37*	2%
Condition x APM (β_{13})	-.13	.05	-2.48*	3%

* $p < .05$; ** $p < .01$

Notes:

Level 1 model (for all analyses): $Y_{ti} = \pi_{0i} + \pi_{1i} \times (\text{Control Phase})$, where Y_{ti} is person i 's control performance score at time t , π_{0i} is their mean control performance score and π_{1i} is their change in control performance from control phase 1 to control phase 2.

Level 2 model for RCR Analysis:

$$\pi_{0i} = \beta_{00} + r_{0i} \text{ and } \pi_{1i} = \beta_{10} + r_{1i}$$

Level 2 model for ISAOR Analysis 1:

$$\pi_{0i} = \beta_{00} + \beta_{01} \times (\text{Condition}) + \beta_{02} \times (\text{APM}) + r_{0i} \text{ and } \pi_{1i} = \beta_{10} + \beta_{11} \times (\text{Condition}) + \beta_{12} \times (\text{APM}) + r_{1i}$$

Level 2 model for ISAOR Analysis 2:

$$\pi_{0i} = \beta_{00} + \beta_{01} \times (\text{Condition}) + \beta_{02} \times (\text{APM}) + \beta_{03} \times (\text{Condition} \times \text{APM}) + r_{0i}$$

and

$$\pi_{1i} = \beta_{10} + \beta_{11} \times (\text{Condition}) + \beta_{12} \times (\text{APM}) + \beta_{13} \times (\text{Condition} \times \text{APM}) + r_{1i}$$

When intercepts are outcomes, ΔR^2 is expressed as a percentage of the variability in mean control performance scores. When slopes are outcomes, ΔR^2 is expressed as a percentage of the variability in the change in control performance.

5.4. Discussion

In summary, the results supported each of the hypotheses: 1) VOTAT strategy use accounted for additional variance in structural knowledge over and above fluid intelligence (VOTAT hypothesis); 2) subjects who acquired more structural knowledge during the exploration phase had better control performance in phase 1 (Knowledge hypothesis); 3) subjects who received additional information improved their control performance more than those who received no additional information (Information hypothesis) and 4) when subjects received additional information, their change in control performance scores from control phase 1 to 2 was more strongly related to fluid intelligence than the change in control performance scores for subjects who did not receive information (Intelligence hypothesis). These results suggest that the provision of structural information does confer an advantage in controlling a dynamic system over knowledge that is acquired through an unguided exploration of the system variables and that subjects can translate such information into effective control actions without practice. However, benefiting from structural information depends on fluid intelligence; information needs to be translated into knowledge and this process involves intelligence-related capacities.

As in previous studies, it was found that the amount of structural knowledge acquired by problem solvers is strongly related to the quality of control performance (Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer et al., 1996, Burns & Vollmeyer, 2002; Osman, 2008a). In contrast to previous studies (Putz-Osterloh, 1993; Preußler, 1998), it was found that the provision of structural information provided an immediate advantage over incomplete knowledge that was acquired through an unguided exploration of the system variables. These findings suggest that the quality of problem solvers' control performance is *causally* dependent on the amount of knowledge that they are able to acquire.

The seemingly discrepant results obtained in previous studies (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996) can be explained by differences in the instructional methods used to provide structural information. The results of the present study show that if problem solvers receive a direct demonstration as to how each input effects each output and have access to this information in the form of a diagram during control performance, then they will be able to immediately and

effectively translate this information into the appropriate actions for control. This provides further support for the claim that information that is provided as needed is much more effective in promoting changes in performance than information that is provided before interaction with the task begins (Kotovsky et al., 1985; Berry & Broadbent, 1987; Leutner, 1993; Hulshof & De Jong, 2006).

On the other hand, the argument could be made that subjects' exploration of the system, prior to the instruction of knowledge, may have prepared them to effectively extract knowledge from the diagram. It remains an open question as to whether the same instruction might result in effective control performance without prior exposure to the complex problem. This issue will be addressed in Study 2 and 3. Nevertheless, these findings imply that it is not practice in applying information that is important for improving system control, but rather the efficacy of the instruction.

Indeed, as subjects in the no information condition showed little improvement across the control phases this suggests that practice at controlling the system does not have a significant impact upon the quality of problem solvers' control performance. The high level of internal consistency in control performance scores further suggests that problem solvers do not dramatically change their control behaviours through practice. Hence, improvements in control performance with practice are rather limited.

With regard to the relationship between fluid intelligence and control performance, these results are in line with previous findings that have shown that when a causal diagram is provided, control performance is moderately to strongly correlated with fluid intelligence. That is, more intellectually capable individuals are able to make use of structural information more effectively than individuals who are less so (Putz-Oterloh, 1981; Putz-Osterloh & Lüer, 1981; Kröner et al., 2005). This study extended on these previous findings, as it was also found that fluid intelligence had an impact on the acquisition of structural knowledge during the exploration phase, and subsequently in controlling the system when only incomplete knowledge was available. These results suggest that subjects who are more intelligent are at a double advantage in comparison to those who are less intelligent with regards to acquiring and utilising structural knowledge: They are able to acquire more

knowledge without assistance, and they also benefit more from direct instruction. These results support the view, frequently advanced by Snow (Snow & Yallow, 1982; Snow, 1986; Snow, 1989; Snow & Lohman, 1989), that individual differences among learners "...present a pervasive and profound problem to educators" (p.1029, Snow, 1989).

On the other hand, the findings with regard to the inter-relationships between VOTAT strategy use, fluid intelligence and knowledge acquisition indicates that problem solvers who lack in fluid intelligence might benefit from instruction in experimental design. The results found in the current study are in line with previous research that has found that the acquisition of structural knowledge is related to the use of the VOTAT strategy and that instruction of the VOTAT strategy also improves performance (Kröner et al., 2005; Vollmeyer et al., 1996; Putz-Osterloh, 1993). However, the results extend upon these previous findings, as they indicate that the use of this strategy is only weakly related to individual differences in fluid intelligence. This supports Dunbar and Klahr's (1988) claim that a key difference between the demands of traditional rule-induction tasks and scientific discovery learning is that the latter situation requires the design of appropriate experiments for effective rule induction. The results of the current study suggest that VOTAT strategy instruction may significantly reduce the problems that learners experience during discovery learning, regardless of their level of fluid intelligence.

The results with regard to the role of fluid intelligence also provide support for the claim that in previous studies (Putz-Osterloh & Lür, 1981; Putz-Osterloh, 1993; Preußler, 1996), the effect of structural information on control performance may have been masked by individual differences in the ability to understand and utilise the information. In addition, Preußler's (1998) finding that all subjects are able to effectively utilise information after a period of practice, may now be interpreted in a different light. It may be that practice *per se* is not the essential component, but rather that subjects who are less intellectually capable require more extensive instructions in order to be able to make sense of, and use, the information that is provided.

Overall, these results imply that the direct instruction of structural information has the potential to increase the amount of knowledge acquired by problem solvers

and improve the quality of their control performance. The crucial aspect of instruction, however, is that it is well designed. These findings are in line with those from other domains that show that learners experience many difficulties when they are required to acquire knowledge without guidance (de Jong & van Joolingen, 1998; Mayer, 2004; de Jong, 2005; 2006). There are also substantial individual differences in problem solvers' capacity to make use of such information, however, which seem to be attributable to differences in fluid intelligence. These results suggest that a combination of VOTAT strategy instruction, an exploration of the system variables, and direct instruction may be most appropriate for encouraging learners of differing levels of ability to develop complete and accurate models of systems for later control performance.

CHAPTER 6

STUDY 2: YOU NEED TO KNOW: THERE IS A CAUSAL RELATIONSHIP BETWEEN STRUCTURAL KNOWLEDGE AND CONTROL PERFORMANCE IN COMPLEX PROBLEM SOLVING TASKS

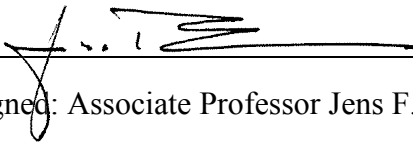
6.1 Permissions for published work

The following chapter consists of a published paper:

Goode, N. & Beckmann, J.F. (2010). You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence*, 38 (3), 345-352.

It has been amended in some sections in response to examiners comments.

The majority of the work in this paper is my own. The second author, Associate Professor Jens F. Beckmann, was a full collaborator in the publication of this paper.



Signed: Associate Professor Jens F. Beckmann

6.2 Introduction

The production of goods in a factory, managing an economy, and driving a car, could all be described as complex, dynamic systems of interdependent relationships between variables. Presumably, once the causal relationships between variables have been discovered, this information could be used to control the outcomes of the system, or change the system, although as yet this has not been established empirically. There might also be individual differences in how well problem solvers are able to understand or utilise such information. The aim of this study is to determine whether the amount of information problem solvers have about a system has a causal impact on how well they can control it, and whether their level of fluid intelligence co-determines the extent to which this information can be applied.

Several computer-based complex problem solving (CPS) tasks, sometimes referred to as simulations or micro-worlds, have been developed to represent the key features of dynamic systems (e.g. Dörner, 1986; Funke, 1992). They consist of a number of input and output variables that are represented in a computer program. The values of inputs can be changed, which affects the values of the output variables through a set of mathematical equations. These tasks are “dynamic” because the values of the outputs change in response to user input and independently over time. In order to study the role of structural knowledge in system control, often, problem solvers are first required to determine how the inputs affect the outputs (i.e. exploration phase). This is referred to as the underlying structure of the system. They then try to control the system through the input variables to reach and to maintain defined goal states of the output variables (i.e. control phase).

In complex problem solving research a common assumption is that problem solvers’ control of a task is dependent upon their knowledge of the underlying structure of the problem (Blech & Funke, 2005). Indeed, the available correlational evidence supports this assumption. Funke and Müller (1988) found that control performance and knowledge of the underlying structure of the task were significantly positively correlated ($r = .41$), as did Beckmann and Guthke ($r = .51$, 1995), Vollmeyer, Burns and Holyoak ($r = .57$ and $r = .65$, 1996), Kröner, Plass & Leutner ($r = .77$ and $r = .61$, 2005) and Kluge ($r = .82$, 2008).

These findings trigger two seemingly contradictory interpretations. One is that the quality of problems solvers' control performance is a function of the amount of structural knowledge acquired about the task. The other interpretation is that the acquisition of structural knowledge and the control of a complex, dynamic system depend on the same set of cognitive abilities.

In line with the latter explanation direct manipulations of the amount of structural information available to participants do not appear to improve control performance. For example, in an early study by Putz-Osterloh and Lürer (1981) an experimental group received a diagram of the underlying structure of the task, while a control group received no supporting information. The diagram was presented on paper, as a system of arrows linking the input and output variables. Although the groups differed in terms of their verbal protocols, they did not differ in the success of their control performance. Later, Putz-Osterloh (1993) replicated this design with a different task. Again, the experimental group performed no differently to the control group. Putz-Osterloh (1993) argues that these results show that successful control performance can be accomplished without the aid of specific knowledge, and that a simple strategy of trial-and-error may be as efficient in reaching the goal states as the application of the rules underlying the system.

However, in the following section we will argue that the method used to inform participants as to the underlying structure of the task in these studies has some limitations, and that these studies have not provided a strong enough test of the hypothesis that complex problems can be controlled without structural knowledge.

Firstly, the instructional method used in these studies presents two sources of difficulty to the participants. Firstly, the user interface of complex problem solving tasks can be quite complicated. Wagener (2001) has found that computer familiarity was a significant predictor of complex problem solving, even when reasoning ability was partialled out. Therefore, it may take time for the participant to understand how to apply the information that they have been given. This source of difficulty may initially diminish any advantage that problem solvers with structural knowledge hold with regard to controlling the system. A direct demonstration of how to alter the values of the input variables may be needed to familiarize participants with the user interface.

Secondly, participants are unlikely to have encountered causal diagrams as a method of representing relationships between variables in a computer program. Therefore, they may not have understood how the diagram presented on paper related to the input and output variables as they were presented in the task on the computer. In order to overcome this problem the structure of the task could be presented on screen as arrows linking the input and output variables. Given these improvements, it is predicted that problem solvers with supporting structural information should display better control performance than those without such information.

A question however that cannot be answered through the comparison of control performance under different amounts of information is whether a group's performance that received no information is as good as a group's performance that received full information or whether the full information group's performance was as poor as their no-information counterparts'. That is, whether there are two qualitatively different, but equally effective, methods that can be used to control a system: One that is knowledge based, and another that is ad hoc, or whether those that are given structural information fail to apply it.

This question is not adequately addressed in Putz-Osterloh's (Putz-Osterloh & Lüer, 1981; 1993) studies, as they do not provide a clear criterion for successful performance. Beckmann (1994) has suggested that control performance should be compared with what might have been achieved through the random action of the participant (see also Kluge, 2008). Unless this comparison is undertaken, there is no objective measure of successful performance. In the current study control performance under different conditions of structural information will be compared to scores resulting from simulated random control interventions. It is predicted that problem solvers who are given information about the structure of the system should perform better than random, while those without such information should perform no better than random.

We expect that even when provided with complete information control performance will not be perfect due to significant individual differences in problem solvers' capacity to apply what they have learnt. Fluid intelligence, as the capacity to reason abstractly, seems a likely source of individual differences in the application of

knowledge (Cattell, 1971). In particular, tests of fluid intelligence and control performance when structural information is given require problem solvers to manage sub-tasks related to the application of rules, and once a correct response to a situation has been produced, generalise these responses to new items, or in the case of complex problem solving tasks, trials with the same goal state. Therefore, we would expect that correlations between control performance under full information conditions and tests of fluid intelligence to be substantive. However, reported correlations between Raven's Advanced Progressive Matrices (APM) and control performance under such conditions are rather moderate (Putz-Osterloh, 1981; Putz-Osterloh & Lüer, 1981).

In comparison, no relationship has been found between fluid intelligence and control performance when participants are not given any structural information, which suggests that under such conditions the demands of the two tasks are relatively dissimilar (Putz-Osterloh, 1981; Putz-Osterloh & Lüer, 1981). Putz-Osterloh (1993) argues that when little structural information is available problem solvers must learn by doing. It could be argued that this skill is usually not required in traditional tests of intelligence.

However, if control performance is in fact not better than random in the first place, then we should not expect to find a relationship between intelligence and control performance. Another contributor to the lack of substantial correlations between control performance and indicators of intellectual capacity may be the poor reliability of control performance measures (Kluwe, Misiak & Haider, 1991). CPS researchers do not consistently report reliability indices and those reported tend to be low (Wenke & Frensch, 2003), with the notable exception of Kröner et al. (2005).

Therefore, in the present study we will address these issues in three ways: Firstly, we will report reliability indices for all performance measures. Secondly, control performance will be assessed across different goal states to test whether variability in performance is attributable to properties of the system, or the individual. Thirdly, control performance will be operationalised in a more meaningful way. Classically, knowledge application, or system control performance, has been operationalised by calculating the deviation of the current states of the output values from the goal states of the output values in terms of the root means

squares criterion (RMS). Thus, deviations become higher the further away the actual state of the system variables are from the goals (Funke, 1992; 1993). However, this approach does not take into account that the value ranges of the input variables are usually constrained in CPS tasks. If a poor decision is made on a given trial, a number of trials may be required to realign the output values with the goal states. Hence, the RMS criterion penalises participants although they may be making an optimal intervention towards the goal state given the current state the system is in. Therefore, Beckmann's (1994) approach to measuring the quality of knowledge application will be adopted in which the deviation from an optimal intervention (that would bring the system closest to the goal state) and the actual intervention made by the problem solver is the operationalisation of the quality of system control.

As the main goal of this paper is to determine whether there is a causal relationship between the amount of structural knowledge and the quality of system control we will also introduce a condition in the experimental design where problem solvers are provided with partial information. This should allow us to determine whether different amounts of structural information lead to quantitative changes in performance.

In summary, three hypotheses are put forward with regards to the effects of structural information and fluid intelligence on control performance in complex problem solving tasks. Firstly, it is hypothesised that the quality of participants control performance will be a function of the amount of structural information that is available to them (Information Hypothesis). Secondly, structural information is a precondition of better than random performance (Success Hypothesis). Thirdly, the magnitude of the relationship between fluid intelligence and control performance will increase as a function of the amount of structural information available to participants (Intelligence Hypothesis).

6.3. Method

6.3.1 Subjects

75 first year psychology students (32 male) at the University of Sydney participated for course credit. The sample size chosen guarantees sufficient statistical

power ($1 - \beta \geq .80$) in identifying at least medium effects ($f^2 \geq .15$) at a significance level of $\alpha \leq .05$ in the planned contrast analyses.

6.3.2 Design

A between-subjects design was used with three levels of structural information available to participants (complete, partial and no information). In the complete condition, participants were informed as to all the relationships between the variables, in the partial they were informed as to all the relationships but one, and in the no information condition they did not receive any information about the underlying structure of the task. A set of simulated data for twenty-five fictitious participants was also constructed based on randomly-generated control inputs. The simulated “random” condition served as a benchmark to assess the quality of the empirical data in the three experimental conditions. Three dependent variables were derived: control performance for phase 1, control performance for phase 2 and performance in a test of fluid intelligence.

6.3.3 Dependent Variables

6.3.3.1 Control performance

The system used in the current study consisted of three input variables and the participants were given seven trials to reach and maintain the goal values for three output variables during each control phase. Control Performance was operationalised as the distance between the intervention on the input variables made by the participant, and the optimal intervention necessary to reach the goal values for the output variables, averaged across the seven trials for each control phase. To calculate the distance between the intervention made by the participant and the optimal intervention for each trial, the previous values of the output variables and the goal state values were used to solve the set of linear equations underlying the system. This indicated the *ideal* values of the input variables. If the *ideal* values were between -10 and 10 , (which were the limits of the input variables in this system), then the ideal values were equal to the *optimal* values. However, if an ideal value was greater than 10 , then the optimal value was set at 10 . If it was less than -10 , then the optimal value was set at -10 . The Euclidean Distance (D_{Euclid}) between the actual and the optimal values of the input variables was calculated using the equation:

$$D_{\text{Euclid}} = \sqrt{\sum (X_{it}^{\text{actual}} - X_{it}^{\text{optimal}})^2} \quad (1)$$

where X_{it}^{actual} represents the actual value of input variable i at trial t , and X_{it}^{optimal} represents the optimal value of input variable i at trial t .

The distances calculated for the first set of seven trials and the second set of seven trials were then averaged to provide control performance scores for phase 1 and 2, respectively. The theoretical range of the score is from 0 to 34, where a lower score indicates that participants had better control.

6.3.3.2 Fluid intelligence

Raw scores on the Raven's Advanced Progressive Matrices (APM) were used as an indicator of fluid intelligence. The maximum score possible was 38. This test has been standardised for a university-level population and extensively validated as an indicator of fluid intelligence (Raven, Raven & Court, 1998).

6.3.4 Materials

The underlying causal structure of the complex problem employed in the study is identical to that used by Beckmann (1994), which is based on Funke's (1992) linear structural equation approach. The two main differences, in comparison to previous studies (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993), refer to the interface and the semantic embedment. With regard to the interface all inputs and outputs are non-numerical graphical (see Fig. 1). The main argument for this decision was that such an interface discourages low-level calculations and attempts to infer equations. Rather, the graphical interface is expected to encourage the development of mental representations that are more aligned with the causal diagrams that are used to inform participants of the underlying structure of the system than equation-like mental models would be. In order to avoid uncontrolled influences of prior knowledge the labels used in the complex system are context free. As can be seen in Figure 6.1, the input variables were labelled A, B and C, while the output variables were labelled X, Y and Z.

The complex problem consists of three input and three output variables that are connected by a set of linear equations:

$$X_{t+1} := 1.0 * X_t + 0.8 * A_t + 0.8 * B_t + 0.0 * C_t \quad (2)$$

$$Y_{t+1} := 0.8 * Y_t + 1.6 * A_t + 0.0 * B_t + 0.0 * C_t$$

$$Z_{t+1} := 1.2 * Z_t + 0.0 * A_t + 0.0 * B_t - 1.0 * C_t$$

X_t , Y_t , and Z_t denote the values of the output variables and A_t , B_t , and C_t denote the values of the input variables during the present trial while X_{t+1} , Y_{t+1} , Z_{t+1} denote the values of the output variables in the subsequent trial. Beckmann (1994; Beckmann & Guthke, 1995) has used this set of equations in previous studies to study complex problem solving in different semantic contexts. The system structure does not contain any dependencies between input variables nor between output variables. The structure is considered balanced, as out of a possible twelve relationships between the variables, six of these are active.

6.3.5 Procedure

The complex problem solving task was presented to participants on PCs. In contrast to previous studies, participants did not explore the system actively. Instead, a recording of seven intervention trials was shown to participants with an accompanying narration. As can be seen in Figure 6.1, the boxes on the input variables represent the values of the input variables over the seven trials, and for each trial positive values of the input variables are shown above the input line, while negative values are shown below. The strength of the inputs is indicated by the vertical dimension of the hatched field in the respective box per trial. After the values for each of the three input variables are set, the effects upon the output variables were displayed in graphs.

In the complete condition on the first trial the inputs were set at zero, so that the autonomic changes in the outputs could be detected. Subsequently, input A was increased to maximum while the other inputs were set at zero, then on the next trial input A was reduced to minimum while the other inputs were set at zero. This was repeated with inputs B and C on subsequent trials so that the effect of each input on the outputs could be clearly observed. After each trial, the narrator explained how each of the outputs had changed, and how this reflected the underlying structure of the system. A causal diagram was constructed on screen, to record this information.

Figure 6.1 shows a screen shot of the task, with the causal diagram as presented in the complete information condition. The arrows represent the relationships between the variables, while the positive and negative signs denote the direction of the relationship, and the letters the relative strength, where “W” represents a weak relationship, “M” represents a medium relationship and “S” represents a strong relationship. Rather than comparing the absolute value of the coefficients, the strength of each relationship was determined by comparing similar effects of similar quality. The direct effect of Inputs A and B on Output X, and Input C on Output Z, were considered to be weak effects in comparison to the strong effect of Input A on Output Y. The dynamic changes in the system were both considered medium effects, as the increase in Input Y is the same as the decrease in Output Z.

Participants in the partial information condition received similar instructions; however, the effect of input B on the outputs was omitted. On the first trial, the inputs were set at zero, so that the autonomic changes in the outputs could be detected. Subsequently, on the second trial input A was increased to maximum while the other inputs were set at zero, then on the third trial input A was reduced to minimum while the other inputs were set at zero. On the fourth trial all the inputs were increased to maximum and on the fifth trial all the inputs were decreased to minimum. This intervention was designed to disguise the effect of input B on the outputs. On the sixth and seventh trial, as in the complete information condition, the effect of input C on the outputs was demonstrated. After each of the seven trials, the narrator explained how each of the outputs had changed, and how this reflected the underlying structure of the system. A causal diagram was constructed on screen to record this information. The narration and diagram was identical to that given in the complete information condition except that they omitted information concerning the effect of Input B on the outputs.

In the no-information condition multiple inputs were varied on each intervention trial. This pattern of interventions mirrors what can be observed in problem solvers who fail to acquire knowledge when confronted with the task to explore the causal structure of a system, as if all inputs are varied at the same time inferences cannot be made concerning individual effects. In this condition, at the end of each trial the narrator explained how the outputs had changed, but did not make any inferences with regards to the structure of the system, and they did not receive

any additional supporting information. During this period participants were not informed about the goals that they would later have to reach.

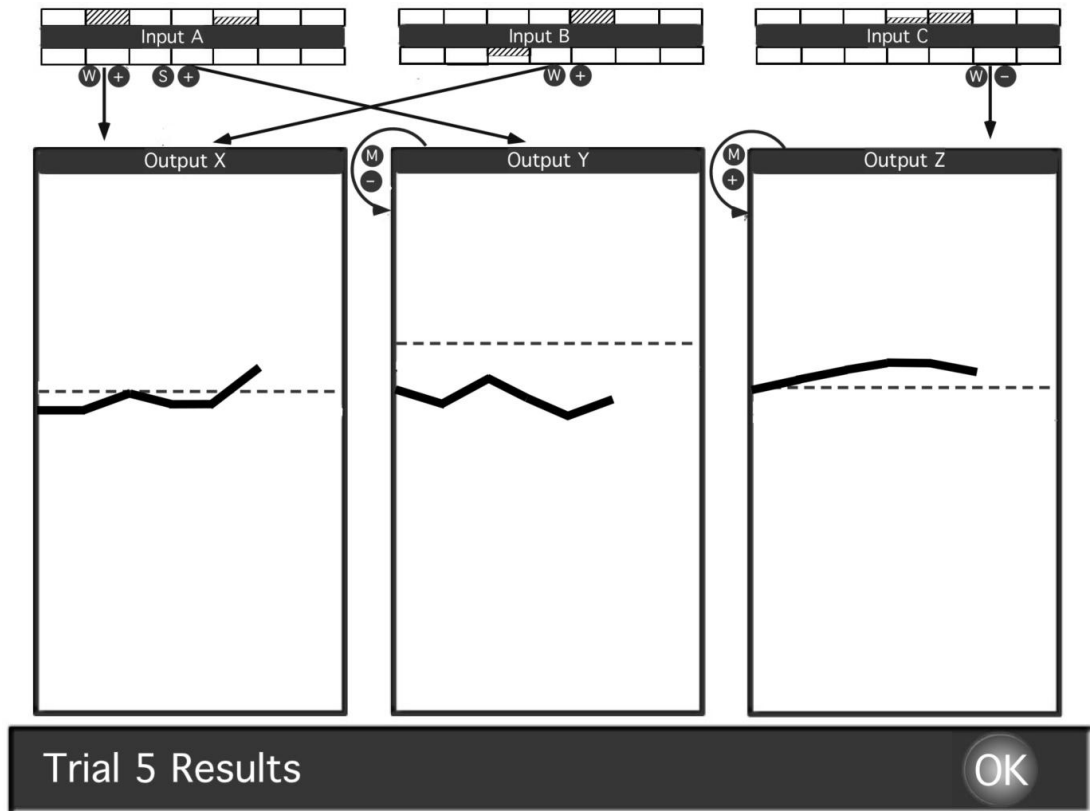


Figure 6.1: Screenshot of the task, as presented in the complete information condition, with the goals indicated as dotted lines on the graphs for the output variables. In this example, the fictitious problem solver set all inputs to zero in trial 1; in trial 2 only Input A was set to its maximum positive value while all other inputs were set to zero; in trial 3 Input B was set to a medium negative value; in trial 4 Input C was set to a medium positive value. In trial 5 all inputs were changed: Input A was set at a medium positive value, Input B at the maximum positive value and Input C was set to three quarters of its maximum positive value

In the second phase of the task, all participants had to control the system by manipulating the inputs to reach certain values of the outputs, which were indicated as lines on the output graphs. The structural information appropriate to each condition was available on screen. There were two control phases consisting of seven trials each. Each trial required three decisions from the participant. On each trial they had to decide whether they wanted to increase, decrease or set each input variable to zero. This was done step by step, such that after they decided what they wanted to do with input A, they then decided what they wanted to do with input B and ultimately input C. Previous decisions could not be altered. Although the numerical values of

the inputs were not available to participants, the inputs could be varied incrementally, within the range of -10 to 10. Each click of the mouse on either the “increase” or “decrease” button increases or decreases the value of the selected input variable by one unit. At the end of seven trials (control phase 1), the graphs were reset, and new goals were indicated onscreen (control phase 2).

Afterwards, participants completed the APM.

6.4. Results

6.4.1 Internal consistencies

Firstly, an internal consistency analysis was conducted to determine the level of variability in control performance across the trials and for different goals states. Reliability was consistently good in the complete, partial and no information conditions across the fourteen trials ($\alpha_{complete} = .89$, $\alpha_{partial} = .79$, $\alpha_{no} = .82$), and for the first seven trials ($\alpha_{complete} = .87$, $\alpha_{partial} = .72$, $\alpha_{no} = .73$), and the second seven trials ($\alpha_{complete} = .72$, $\alpha_{partial} = .63$, $\alpha_{no} = .70$). These results indicate that the variability in control performance within each individual is rather low, which also justifies averaging the scores across each set of seven trials.

In comparison, and as would be expected, the reliability of the randomly generated performance scores was poor across the fourteen trials ($\alpha_{random} = .34$), and for the first seven trials ($\alpha_{random} = .28$), and the second seven trials ($\alpha_{random} = .41$). This indicates that in comparison to randomly generated inputs, potential differences in control performance between the no, partial and complete information conditions is the result of systematic sources of variability.

6.4.2 The success hypothesis

To determine whether participants’ control performance was better than random, a multiple regression was conducted, using dummy coded variables to compare each experimental condition to the random performance scores. Mean control performance in the first phase was 11.03 ($SD = 4.10$) in the complete condition, 14.17 ($SD = 3.57$) in the partial condition and 16.39 ($SD = 4.44$) in the no information condition. In the second phase, mean control performance was 10.33 ($SD = 5.25$) in the complete condition, 13.82 ($SD = 3.96$) in the partial condition and

16.76 ($SD = 4.25$) in the no information condition. The mean random control performance was 16.77 ($SD = 2.18$) in the first phase, and 18.96 ($SD = 2.34$) in the second phase. In support of the success hypothesis, control performance in the complete condition was significantly better than random across both phases; $b = -5.74$, $t(96) = -5.52$, $p < .01$, $f^2 = .32$ and $b = -8.63$, $t(96) = -7.46$, $p < .01$, $f^2 = .58$ for the first and second phase respectively, as was control performance in the partial condition; $b = -2.60$, $t(96) = -2.50$, $p = .01$, $f^2 = .07$ and $b = -5.14$, $t(96) = -4.44$, $p < .01$, $f^2 = .21$. In contrast, control performance in the no information condition did not differ significantly from random in either of the control phases; $b = -.38$, $t(96) = -0.37$, $p = .72$, $f^2 = .001$ and $b = -2.19$, $t(96) = -1.90$, $p = .06$, $f^2 = .04$ respectively. This indicates that only participants who received structural information performed better than random.

6.4.3 The information hypothesis

To determine whether participants' control performance was dependent on the amount of structural information that was available to them, multiple regression analyses were conducted, using Cohen coded variables to conduct a contrast analysis. The model contained three predictors of control performance. Firstly, scores on the APM were included in order to control for any potential group differences in fluid intelligence, contrast 1 compared the conditions that received any information (complete and partial) to the no information condition and contrast 2 compared the complete and partial information conditions.

In support of the information hypothesis, in the first and second phases, control performance in the complete and partial information conditions was significantly better than those in the no information condition; $b = 3.22$, $t(71) = 3.30$, $p < .01$, $f^2 = .15$, and $b = 3.81$, $t(71) = 3.70$, $p < .01$, $f^2 = .19$, respectively. Control performance in the complete information condition was significantly better than in the partial information condition across both control phases; $b = -3.07$, $t(71) = -2.82$, $p < .01$, $f^2 = .11$ and $b = -3.81$, $t(71) = -2.94$, $p < .01$, $f^2 = .12$. These reported differences are independent of any potential group differences in fluid intelligence, although scores on the APM were a significant predictor of performance across both goal phases; $b = -.23$, $t(71) = -2.95$, $p < .01$, $f^2 = .12$ and $b = -.34$, $t(71) = -4.24$, $p < .01$, $f^2 = .25$, respectively. These results show that increasing the amount of structural information

available to the problem solver improves performance, and that even small changes in the amount of information available has a significant impact upon performance.

6.4.4 The intelligence hypothesis

We predicted that the relationship between fluid intelligence and control performance would increase as a function of the amount of information that problem solvers have about the structure of the task. As can be seen in Table 6.1, across both control phases the relationship between fluid intelligence and control performance is significant and strong in the complete information condition, while in the partial and no information conditions it is non-significant and weak to moderate. Overall there was a significant moderate correlation between control performance and fluid intelligence across the conditions in both control phases.

Table 6.1
The Relationship between APM scores and Control Performance in each Phase by Condition

	Complete information	Partial information	No information
Goal 1	-.53*	-.25	-.19
Goal 2	-.61*	-.29	-.37*

* $p < .05$, one tailed

In order to determine whether the strength of the relationship between fluid intelligence and control performance differs across the conditions two moderator analyses were conducted, which are presented in Table 6.2. The non-significant interaction terms indicate that the relationship between scores on the APM and control performance was not significantly moderated by condition in either control phase; $R^2_{change} = .01, p = .23, f^2 = .02$ and $R^2_{change} = .02, p = .17, f^2 = .03$, respectively. However, the effect sizes for the interaction terms are small and the sample size is not sufficient to detect such effects. Nevertheless, our findings do not directly support the intelligence hypothesis. Rather, it appears that control performance is related to fluid intelligence, to a certain degree, regardless of the amount of information available to the problem solver.

Table 6.2

Results of two moderator analyses examining the dependency of the relationship between control performance (phase 1 and phase 2, respectively) and fluid intelligence (APM score)

<i>Control Phase 1:</i>						
Step	Variables entered	β	t	p	R^2_{change}	F
1	APM	-.29	-2.90	< .01		
	Condition	-.43	-4.28	< .01		
					.312	16.36
2	APM	-.01	-0.05	.96		
	Condition	-.02	-0.06	.95		
	APM x Condition	-.55	-1.22	.23		
					.014	1.48
<i>Control Phase 2:</i>						
Step	Variables entered	β	t	p	R^2_{change}	F
1	APM	-.39	-4.22	< .01		
	Condition	-.43	-4.69	< .01		
					.406	24.64
2	APM	-.10	-0.44	.66		
	Condition	-.01	-0.02	.98		
	APM x Condition	-.58	-1.39	.17		
					.016	1.92

Note: The degrees of freedom for the F -test of R^2 change in the step 1 models are (2, 72) and for the step 2 models are (1, 71).

6.5 Discussion

In summary, support was found for the information hypothesis, as performance improved systematically as the amount of structural information available to participants increased. With regards to the success hypothesis, control performance was better than random only when partial or complete information was available to participants. This strengthens the conclusions that can be drawn from the differences between the conditions, demonstrating again that control performance when partial or complete information is available is better than when no information is available. Finally, the results did not support the intelligence hypothesis. Although it appears that the relationship between control performance and fluid intelligence differs between the different information conditions, the potential moderator effect was too small to be statistically identified under the given circumstances. This suggests that the quality of decision making in CPS tasks under different levels of information is associated with measures of intellectual capacity, although to various degrees.

The results presented contradict some previous findings (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993). This is likely because of the different method used to inform participants of the underlying structure of the problem in previous studies.

The results of this study suggest that if participants understand how to interact with the problem, and structural information is available on screen during the performance of the control task, then structural information does benefit control performance.

The study provides a causal explanation to the consistent positive correlation found between structural knowledge and control performance reported in other studies (Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer, Burns & Holyoak, 1996; Kröner, 2001; Kluge, 2008). The finding that control performance was better than random only when (any) structural information was available demonstrates that complex problems cannot successfully be controlled without structural knowledge, and contradicts Putz-Osterloh's (1993) claim that participants learn by doing when information about the structure of the system is not provided. Rather, the results indicate that the quality of control performance is a function of the amount of information available (and utilised). The design of the study presented also made it possible to establish a *causal* link between knowledge and control behaviour that previously was based on correlational evidence alone.

With regard to the generalisability of these findings to more complex tasks we can speculate that at a certain level of complexity there might be no difference in control performance under conditions where partial or complete information is available given that there is a limit on how many causal relationships problem solvers can consider at one time. Future studies could systematically vary the complexity of the system with the amount of information available. However, the fact that withholding information about only one out of a total of six relations in the current system results in a significant reduction in control performance suggests that the sheer count of the number of existing or non-existing relationships between variables within a complex system represents a rather poor proxy of its complexity.

With regard to the relationship between fluid intelligence and control performance, the results of this study contradict those reported by Putz-Osterloh (1981; Putz-Osterloh & Lüer, 1981) to some extent. Putz-Osterloh found that the correlation between fluid intelligence and control performance was moderate when complete structural information is available, and close to zero when no information is available. In the current study, it was found that the correlation between fluid

intelligence and control performance is strong when complete structural information is available, and weak to moderate when no information is available.

The present study extends upon previous findings, as it was found that the correlation between performance and fluid intelligence is weak to moderate even when partial structural information is provided. This indicates that when incomplete structural information is available to the problem solver, success in controlling a complex system cannot be sufficiently predicted by fluid intelligence. Given the complexity of the environment we are living in on one hand and the limitations of human information processing capacity on the other, it seems quite likely that people often operate using incomplete representations of the systems they attempt to control, a situation that appears to be reflected by control performance under incomplete information conditions. However, it remains an open question as to whether performance under such conditions has the potential for incremental validity in predicting real world problem solving over and above traditional intelligence tests.

It could be argued that the relationship between fluid intelligence and control performance might be attenuated because information about the system was *presented*, and problem solvers did not have a chance to actively explore the task in order to acquire information about the underlying structure of the system. If problem solvers were required to actively acquire knowledge, then it might be expected that intellectually more capable problem solvers would be able a) to create system states that are more informative with regard to the underlying structure, and b) to extract more knowledge about the system. Given that control performance is dependent upon the amount of structural knowledge acquired, they should then have an even greater advantage when they are required to control the system. From this perspective, the decision to homogenise the amount of information available to participants within each experimental group leads to a rather conservative testing of the intelligence hypothesis.

Findings in the current study may also indicate that control performance under different levels of knowledge has differential validity. That is, the processes used to control the task under different levels of information may differ. The strong correlation between control performance and scores on the APM suggest that when complete structural information is available the demands of the two tasks are highly

similar, as previously discussed. In comparison, the correlation between control performance and fluid intelligence is weak to moderate when partial or no information is available. This, in conjunction with the high levels of consistency in performance observed across trials and goal states, indicates that under these conditions control performance is systematically influenced by factors not sufficiently captured by traditional measures of intelligence. However, further studies are needed to investigate the determinants of control performance under different levels of information.

In conclusion, the results of this study indicate that successful problem solving in complex and dynamic situations requires task-specific knowledge, which is unlikely to be acquired while trying to control the system. Furthermore, individual differences in fluid intelligence appear to play a role in the utilisation of available information and the subsequent application of the acquired knowledge. Thus, complex problem solving performance, as measured by the current task, can be seen as a function of the amount of information available combined with the ability to utilise such information.

CHAPTER SEVEN

STUDY 3: THE EFFECT OF SYSTEM COMPLEXITY ON THE ACQUISITION OF STRUCTURAL KNOWLEDGE AND CONTROLLING THE OUTCOMES OF A DYNAMIC SYSTEM

7.1 Introduction

The research documented in this thesis indicates that the quality of system control is a function of the amount of structural information that is available to problem solvers (Study 1 and 2), and that dynamic systems cannot be successfully controlled without structural knowledge (Study 2). However, in previous research, the provision of structural information did not show the expected benefit to control performance (Putz-Osterloh & Lüer, 1981; Putz-Osterloh, 1993; Preußler, 1996). In Chapter 5 and 6 it was argued that these results were inconsistent due to differences in the method used to inform subjects about the underlying structure of the system. An alternative, or perhaps complementary, explanation is that the underlying structures of the systems, TAILORSHOP (Putz-Osterloh & Lüer, 1981) and LINAS (Putz-Osterloh, 1993; Preußler, 1996) that were used in previous studies, may be more complex than that used in Study 1 and 2.⁵ This may have made it more difficult for subjects to understand and utilise the information that they were given. Therefore, the aim of the study reported in this chapter is to determine to whether the complexity of the system has an impact on the translation of information into structural knowledge and effective control actions.

As a starting point then, we need to establish what makes one system more or less complex than another, and why a more complex system should be more difficult to acquire knowledge about and to control. However, although a number of researchers have commented on the need for a framework to compare the complexity

⁵ In comparison to the complex problem used in Study 1 and 2, TAILORSHOP and LINAS differ on a number of dimensions. LINAS was constructed using Funke's (1985; 1993; 2001) Linear Structural Equation approach, and so it is comparable to the complex problem used in Study 1 and 2. In both problems, the underlying structure of the system is described by a set of linear equations, specific goals were given for system control and abstract labels are given to the system variables to control for the influence of prior knowledge. In contrast, the construction of TAILORSHOP is more aligned with the Dynamic Decision Making approach (which is described in Chapter 2). The problem is embedded in a rich semantic context and ill-defined goals were given for system control. These factors may influence the difficulty of system control, independently of the complexity of the system. Therefore, comparisons in this chapter will be limited to LINAS and the complex problem used in Study 1 and 2.

of different systems (Quesada, Kintsch & Gomez, 2005; Gonzalez, Vanyukov & Martin, 2005; Osman, 2010), as yet this has not materialised. Moreover, in past research, system complexity has been only defined at an operational level. Due to a lack of proper theoretical conceptualisation, the *explanatory* link between system complexity and task performance is unclear.

Therefore, in the following sections, firstly system complexity will be defined at a conceptual level. This will a) establish an explanatory link between system complexity and task performance and b) identify a basis for the manipulation of system complexity that will be used in the empirical study reported in this chapter. Secondly, we will discuss the implications of this conceptualisation of system complexity for the effect of structural information on the acquisition of structural knowledge and system control, and the relationship between fluid intelligence and system control.

7.1.1 Definition of a construct: System complexity

In research concerned with dynamic systems, system complexity is often discussed as a potential moderator variable, although it is usually only defined at an operational level and there exist multiple views as to what makes one system more or less complex than another. For example, many researchers argue that the number of relations in a system is indicative of system complexity (e.g. Kluge, 2008; Gonzalez, Thomas & Vanyukov, 2005; Sterman, 2000). Other researchers argue that the characteristics of the entire system are important, such as the number of variables and the number and kind of relations that exist among those variables (e.g. Brehmer & Dörner, 1993; Kerstholt & Raaijmakers, 1997; Gonzalez, Vanyukov & Martin, 2005; Brehmer & Allard, 1991; Leutner, 1993). Funke (2001) argues that it is the connectivity of the system that is most important because it requires problem solvers to figure out the connections between variables. Connectivity is understood as the “dependency between two or more variables” (Funke, 2001, p. 73). The common theme across these perspectives is an emphasis on the number of certain system properties, where a higher number is taken as indicative of a higher level of system complexity, which in turn is thought to make it more difficult to acquire structural knowledge and to control the outcomes of the system (e.g. Dörner, 1987; Brehmer & Dörner, 1993; Kerstholt & Raaijmakers, 1997; Kluge, 2008).

However, any account of system complexity that defines the construct through the value of a particular system property is likely to be problematic for a number of reasons. Firstly, such an account does not explain why we might expect certain system properties to be indicative of system complexity, and not others. The explanation usually given, which is consistent with a predominantly data driven approach, is that these properties reliably predict the difficulty of knowledge acquisition and system control. However, a range of system properties, such as feedback delays (Brehmer & Allard, 1991; Brehmer, 1995), the consistency of the relationships between the inputs and the outputs (Ackerman & Cianciolo, 2002), the type of connectivity pattern (Howie & Vicente, 1998), and whether the relationships in the system are linear or non-linear (Dörner, 1987; Dörner & Scholkopf, 1991), have been shown to have significant impacts on the difficulty of knowledge acquisition and control performance. Hence, the justification for the selection of any particular system property over others is unclear.

Secondly, such an approach assumes that estimates of system complexity will be the same regardless of the task to be performed in relation to the system (i.e. acquiring knowledge on one hand, controlling the outcomes on the other). This is inconsistent with findings that show dissociations between the difficulty of knowledge acquisition and system control when certain system properties are manipulated (e.g. the number of variables; Preußler, 1997).

Thirdly, and perhaps most importantly, system properties cannot be used as a psychological explanation as to why one system is more or less complex than another, and thus more or less difficult to acquire knowledge about or to control. Such an explanation requires some reference to the impact that system properties have on the processes that must be executed by the problem solver. Overall, these issues highlight the need for a theoretically driven account of system complexity.

The problem of complexity is not confined to research with dynamic systems, and there exists an extensive literature on how the complexity of cognitive tasks might be conceptualised and measured. One perspective on this problem is that complexity can be conceptualised as a function of the information processing demands of a task (Wood, 1986; Campbell, 1988; Bonner, 1994; Halford, Wilson & Phillips, 1998a; Spilsbury, Stankov & Roberts, 1990; Stankov & Crawford, 1993; Stankov,

1999; Park & Jung, 2007). Although a number of different frameworks exist to measure these demands, they converge on the idea that complexity is a function of the number of cognitive processes that need to be executed in the performance of a task, and the dependencies among those processes. This is determined by the internal representation of the task-doer and the objective requirements of the task. Hence, decrements in performance related to complexity are explained by constraints on the human information processing system (see Appendix D for a review of the literature on task complexity).

Under this approach, system complexity could be conceptualised as one facet of task complexity. Task complexity represents the overall information processing demands of the task to be performed (i.e. acquiring knowledge in one case, controlling the outcomes of the system in the other). System complexity indicates the information processing demands of the task to be performed *engendered by the properties of the system*. This approach provides an explanatory link between the properties of the system and task performance, such that systems with more complex structures are more difficult to acquire knowledge about and to control because problem solvers must execute a larger number of processes in the performance of each of these tasks.

This approach has a number of implications for the operationalisation of system complexity. Firstly, it implies that estimates of system complexity may differ according to whether the task to be performed is to acquire knowledge or to control the system. Some system properties may have an impact on the number of processes that must be executed in order to acquire knowledge about the structure of a system, but not on system control. Conversely, some system properties may have a greater impact on the information processing demands of system control in comparison to knowledge acquisition. This accounts for dissociations between the difficulty of knowledge acquisition and system control. Secondly, under this view, any system property that has an impact on how the tasks of knowledge acquisition and system control are performed could be used to manipulate system complexity, or to compare different systems. These implications address some of the limitations associated with a purely operational approach to system complexity.

7.1.2 Indicators of system complexity

The aim of this section is to identify system properties that might be used to manipulate system complexity, and compare the complexity of LINAS to the system used in Study 1 and 2. As discussed above, system complexity can only be estimated in relation to a particular task to be performed. In this study we are primarily interested in how system complexity might influence the effect of structural information on control performance. This can be conceptualised as two separate tasks. Firstly, the problem solver must acquire structural knowledge through direct instruction. That is, convert information into knowledge (knowledge acquisition). Secondly, they must translate knowledge into effective and systematic control actions (system control). System properties may have a differential impact on the complexity of each of these tasks.

A complete review of the system properties that might influence the complexity of these tasks is far beyond the scope of this chapter. It is also unnecessary considering that we are primarily interested in whether the complexity of the underlying structure of LINAS differed to the system used in Study 1 and 2. These systems only differ on three dimensions: The number of variables, the number of relations and connectivity. Each of these properties may have a differential impact on the information processing demands of the task to be performed, although it should be noted that they are not entirely independent of each other, as connectivity is limited by the number of relationships in the system, which in turn is limited by the number of variables in the system. The following sections will examine how each of these properties might influence the information processing demands of knowledge acquisition and system control.

7.1.2.1 The number of variables

It could be argued that the number of variables in a system is likely to have a stronger impact on the complexity of knowledge acquisition and less impact on system control. With regard to knowledge acquisition, systems with more variables require problem solvers to incorporate more information into their mental model of the system, and this is likely to make it more difficult to acquire complete structural knowledge. In comparison, the total number of variables may not always provide an appropriate indicator of the information processing demands of system control. If the

problem solver is aware that some variables can be ignored while controlling the system (e.g. input variables that do not have any effect on other variables or output variables that do not have defined goal states) then they should have only a small effect on the information processing demands of the control task. At a conceptual level then, the number of variables is not a good candidate for estimating the complexity of both knowledge acquisition and system control.

In line with these expectations, findings show that the number of variables is related to the difficulty of knowledge acquisition, but not system control. Preußler (1997) presented subjects with a complex problem with abstract letters for the system variables. The underlying structure consisted of a set of linear equations. In one condition, subjects were presented with a system that contained four inputs and three output variables. In a second condition, the system contained four additional input variables that had no effect on the output variables. Subjects in both conditions tried to control the system to reach specific goals for two sets of 8 trials, after which they were given a test of structural knowledge. Subjects controlling the version of the system with fewer variables acquired significantly more structural knowledge, yet for control performance, the pattern was reversed; those who dealt with the system with more variables performed significantly better.

These results confirm our expectations that systems with more variables will be more complex in relation to knowledge acquisition but not necessarily in relation to system control. Further, they suggest that variables influence the information processing demands of system control via their relationship to other variables. It is not the variables *per se* that are important, but rather the relationships between them. This suggests that the number of relationships in a system may provide a more appropriate indicator of system complexity in relation to system control.

7.1.2.2 The number of relations

It has often been argued that the number of relations in a system is a key indicator of system complexity (e.g. Kluge, 2008; Gonzalez, Thomas & Vanyukov, 2005; Serman, 2000). However, again, it could be argued that the number of relations may be related to system complexity in relation to knowledge acquisition, but not necessarily to system control. With regard to knowledge acquisition, a larger number of relations will increase the amount of information that must be

incorporated into the problem solvers' mental model. Hence, it should be more difficult to acquire knowledge about systems with more relations.

With regard to system control, the goals set for the problem solver must be taken into account. If some relations can be ignored because they link variables that do not have defined goal states, then the number of relations may not be directly related to system complexity in relation to system control. Alternatively, in cases where all relations do impact upon desired goal states, then systems with more relations, but the same level of connectivity, will be more difficult to control because more information must be processed serially in order to make decisions about the goal states. However, this distinction between the number of relations in a system and the number of relations that impact upon goal states has not been made in previous research, so as yet there is insufficient empirical evidence to support this claim.

Nevertheless, a number of studies have examined the effect of the number of relations on knowledge acquisition and system control (Funke, 1985; 1992; Kluge, 2008). The experimental design of these studies is identical: Subjects were presented with a complex problem which consisted of three input and three output variables which were linked by a set of linear equations. Abstract labels were used for the system variables in order to control for the influence of prior knowledge. The number of relations in the system was varied across different experimental conditions, and all of the relations were linked to output variables with desired goal states. Subjects were first instructed to explore the system in order to acquire structural knowledge. Subsequently, they were instructed to control the system to reach specific goals for all of the output variables. Across these studies, findings indicate that systems with more relations, but the same number of variables, are more difficult to acquire knowledge about, and to control.

However, the conclusions that can be drawn from the results of Funke's (1985; 1992) studies with regard to the effect of the number of relations on system control are limited because the analysis did not take into account the controllability of different systems. Controllability reflects the control performance score that can be achieved through random interventions. If the underlying structure of the system changes, the controllability may change too. This may artificially inflate or deflate

control performance scores. This will confound comparisons of control performance across different system (Strauß, 1993; Kluge, 2008). Funke (1985; 1992) did not correct for controllability, and hence his findings may be attributable to differences in controllability, rather than differences in the information processing demands of the control task engendered by the system properties.

Kluge (2008), however, did correct for controllability in her comparisons of control performance across different systems. These results replicate those reported by Funke (1985; 1992), which makes it possible to conclude that the number of relations in a system appears to affect the difficulty of both knowledge acquisition and system control. Therefore, it appears that the number of relations in a system may provide an appropriate indicator of system complexity in relation to both of these tasks.

7.1.2.3 Connectivity

As discussed, Funke (2001) defines connectivity as the number of dependencies between two or more variables. It should be noted that connectivity and the number of relations in a system are not identical. A system may have many relations, but a low level of connectivity if they form single connections between many variables. Alternatively, systems with few relations may form highly interconnected structures if they exist between few variables. Systems may also have the same number of variables and relations, but different levels of connectivity, as Figure 7.1 shows. Thus, connectivity can be distinguished from the number of relations in a system.

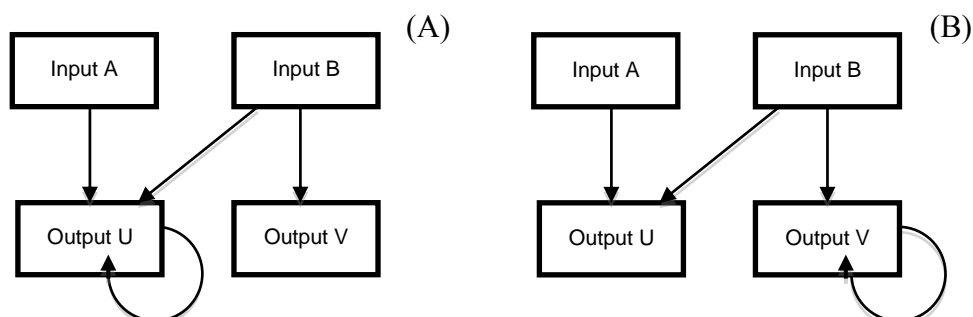


Figure 7.1: Diagram A depicts a system that has four relations between two input and two output variables. The value of Output U depends on three relations and the value of Output V depends on one relation. Diagram B depicts a system that has the same number of variables and relations. However, the connectivity is different; the value of Output U and V both depend on two relations.

Connectivity is likely to have a significant impact on the complexity of knowledge acquisition, as the number of dependencies will determine the amount of information that must be integrated into the problem solvers' mental model of the system. However, connectivity may have less of an impact on how system control is performed if these variables are not relevant to achieving the desired goal states.

Therefore, in the current chapter, we offer a revised definition of connectivity, based on our view of system complexity as a function of the information processing demands of the task to be performed engendered by the system properties. We propose that connectivity will be closely related to the complexity of system control when it is estimated in relation to output variables that have defined goal states. This can be undertaken by counting the number of relations that affect an output variable with a defined goal state. Under these circumstances, connectivity indicates the number of relations that must be processed in parallel in order to make a decision about a particular goal state. This will also impact upon the complexity of knowledge acquisition, as it reflects the number of relations that must be processed in parallel in order to understand how a particular output variable can be altered. Therefore, in the current chapter, all future references to connectivity define it in relation to the goal states.

Conceptually, the connectivity implied by the goal states has a close correspondence to the information processing demands of the task. The amount of information that can be processed in parallel has long been recognised as a critical constraint on human performance (e.g. Miller, 1956), and findings show that increases in the number of relations that must be processed in parallel in reasoning tasks consistently lead to increases in task difficulty (e.g. Halford et al., 1998a; Andrews & Halford, 1998; Birney & Halford, 2002; Halford, Baker, McCredden & Bain, 2005). In terms of the limit on human information processing, initial estimates placed it at seven relations, plus or minus two (Miller, 1956). However, this estimate has since been revised downwards in the light of further empirical evidence (e.g. Broadbent, 1975; Fisher, 1984; Halford, Maybery & Bain, 1986; Halford et al., 2005), which suggests that human information processing capacity is likely to be constrained to a soft limit of processing four relations in parallel (Halford et al., 1998a; Halford et al., 2005). This suggests that the connectivity implied by the goal states should be closely aligned with system complexity in relation to knowledge

acquisition and system control. Moreover, floor effects on performance should be expected when the connectivity implied by the goal states requires problem solvers to process four or more relations in parallel.

To the authors' knowledge, there have been no studies that have directly examined the effect of connectivity in general, or the connectivity implied by the goal states on the difficulty of system control or knowledge acquisition. However, in Funke's (1985; 1992) and Kluge's (2008) studies the manipulation of the number of relations also resulted in changes in the connectivity implied by the goal states. Thus, the results of these studies could also be interpreted as evidence that the connectivity implied by the goal states provides a good indicator of system complexity in relation to knowledge acquisition and system control. However, obviously, it cannot be determined whether these results are attributable to the number of relations, the connectivity implied by the goal states or both. The current study will attempt to determine whether each of these factors has an independent effect on system complexity.

In summary, it seems plausible that the number of relations and the connectivity implied by the goal states should impact upon the information processing demands of knowledge acquisition and system control, although for different reasons. Systems with larger numbers of relations, but similar levels of connectivity, require more information to be processed serially during knowledge acquisition and system control. Systems with higher levels of connectivity implied by the goal states, but the same number of relations, require problem solvers to process larger amounts of information in parallel during knowledge acquisition and system control. In the following section, these properties will be used to compare LINAS and the system used in study 1 and 2, to determine whether system complexity may account for the inconsistent effect of structural information on control performance.

7.1.3 The effect of system complexity on the effect of structural information on knowledge acquisition and system control

The question of interest in the current chapter is whether the inconsistent effect of structural information on system control is attributable to differences in the complexity of the systems used in different studies (i.e. Study 1 and 2; Putz-Osterloh,

1993; Preußler, 1996). We predict that the number of relations and the connectivity implied by the goal states can be used as an estimate of system complexity in relation to knowledge acquisition and system control. However, with regard to system control, the goals set for performance must be taken into account.

The underlying structure of LINAS entails more relations and a higher level of connectivity implied by the goal states than the complex problem used in Study 1 and 2. LINAS contains four input and seven output variables interconnected by fifteen relations. In comparison, the complex problem used in Study 1 and 2 contains three input and three output variables interconnected by six relations. In both of these systems, all relations were relevant to system control, as they all impacted upon output variables with defined goal states. In terms of the connectivity implied by the goal states, in LINAS, the maximum number of relations to be considered to reach a goal state is three. In comparison, in the system used in Study 1 and 2, the maximum number of relations to be considered to reach a goal state was two. This suggests that LINAS is at a higher level of system complexity in relation to knowledge acquisition and system control than the complex problem used in Study 1 and 2.

This may explain why we found in Study 1 and 2 that structural information benefited control performance, while Putz-Osterloh (1993) and Preußler (1996) did not. In effect, the information processing demands implied by LINAS may have been too high for subjects to effectively convert information into knowledge or translate knowledge into effective control actions (or at least the effect may have been too small to detect given the power implied by the sample size). It should be noted that this interpretation differs significantly from that given by Putz-Osterloh (1993), who argues that problem solvers can perform effectively without structural knowledge. In comparison, we argue that floor effects are to be expected on performance when the information processing demands implied by the system properties are too high for problem solvers to understand or utilise the information that has been provided through direct instruction.

Concomitantly, this implies that it should be easier to acquire knowledge about and to control systems that are less complex than LINAS and the complex problem used in Study 1 and 2. If the underlying structure of the system contains few relations or a low level of connectivity there will be less information for problem solvers to

incorporate into their mental model. Hence, it should be easier for problem solvers who are provided with structural information to acquire a complete representation of the system, which in turn should result in better control performance. When complete structural knowledge is acquired, systems with fewer relations or a lower level of connectivity implied by the goal states should also be easier to control because less cognitive steps are required to formulate effective control interventions.

However, the effect of structural information may not necessarily be stronger when the system is less complex as it could be speculated that it may be possible to learn to control such systems through trial-and-error. If this is the case, then problem solvers provided with structural information would not have an advantage over those who are required to control the system without such information. This implies that the provision of structural information may have a weak effect on control performance when the complexity of the system is low, although the overall quality of control performance will be high.

This suggests that the relationship between structural knowledge and control performance may be dependent on the complexity of the system. At high levels of system complexity, as realised in LINAS, there may be little relationship between structural knowledge and control performance because it may be too difficult for problem solvers to acquire knowledge or to control the system. In contrast, at moderate levels of system complexity, as in the complex problem used in Study 1 and 2, information can be successfully converted into knowledge. Subsequently, the effectiveness of control actions crucially depends on the amount of knowledge that can be acquired. At low levels of system complexity, there may be little relationship between structural knowledge and control performance because problem solvers may be able to adequately control the system either through trial-and-error or knowledge-based interventions. This may explain why the effect of structural information on control performance has been found to be inconsistent.

7.1.4 The relationship between fluid intelligence and control performance

Based on the insights gained in Study 1 and 2 in regard to the role of fluid intelligence in problem solving, we assume that more intelligent problem solvers benefit more from the provision of structural information and are better able to control the outcomes of dynamic systems than their less intelligent counter-parts.

Hence, in the current study we will test whether these findings generalise to systems of differing complexity.

Evidence from the psychometric literature suggests that the magnitude of the relationship between fluid intelligence and control performance may vary as a function of system complexity. Findings show that tasks that entail the execution of more processes (i.e. complexity) load more highly on broad ability factors, such as fluid intelligence. Snow, Kyllonen and Marshalek (1984) have proposed that even though they are derived using different mathematical techniques, the radex and hierarchical models of intelligence "...suggest a complexity continuum along which tasks can be arrayed" (Snow et al., 1984, p. 107). The most complex tasks in this framework, such as Raven's Advanced Progressive Matrices (APM), tend to load most highly on the central part of the radex model and on the general ability factor in hierarchical models. In comparison, more elementary cognitive tasks, such as Perceptual Speed, show only modest correlations with the general factor and lie on the periphery of the radex model. A similar conclusion can be derived from a series of studies conducted by Kyllonen (1985). Additive-factors logic was used to construct a set of information processing tasks that varied from tasks that entailed single processes, such as reaction time tasks, to tasks that entailed many processes, such as sentence verification tasks. Mirroring the findings observed in factor analytic studies, Kyllonen (1985) found that as the number of processes involved in task performance increased, so did the correlations with broad ability factors. This suggests that the association (correlation) between fluid intelligence and control performance should increase as system complexity increases because the control of more complex systems should require the execution of more processes. However, this effect may be attenuated in systems that are too complex to be effectively controlled; if there is a floor effect on control performance then there will not be any relationship between fluid intelligence and control performance. Therefore, we predict that the relationship between fluid intelligence and control performance should increase with system complexity, until there are floor effects on control performance.

In Study 1 and 2 we also found preliminary evidence that the relationship between fluid intelligence and control performance may be stronger under conditions where complete structural information is available, in comparison to when partial or

no information is available. This finding is consistent with previous research (Putz-Osterloh, 1981; Putz-Osterloh & Lürer, 1981; Krörner et al., 2005). It was argued that this indicates that the demands of system control and tests of fluid intelligence are more similar when control actions are knowledge-based, rather than the result of trial-and-error. However, the effect size was too small to draw strong conclusions. In this study, we will test whether this effect can be replicated, and whether this finding holds across systems of differing complexity.

7.1.5 Task manipulations and hypotheses

In the current study, a between-subjects design will be utilised with three levels of structural information available to participants (information condition: Complete, partial and no information) and four levels of system complexity. The first factor replicates the design of Study 2. In the complete information condition, subjects will be informed as to all the relationships between the variables, in the partial information condition they will be informed as to all the relationships but one, and in the no information condition they will not receive any information about the underlying structure of the system. Subsequently, they will be given a test of structural knowledge, and then they will be asked to control the system to reach specific goals.

With respect to system complexity, the underlying structure of the system used in Study 1 and 2 will be manipulated in order to create four systems of a comparatively increased and decreased system complexity. In two of the systems, we will manipulate the number of relations independently of the level of connectivity implied by the goal states to determine how each of these system properties impacts upon system complexity in relation to knowledge acquisition and system control.

The complexity of each system will be defined in terms of the number of input and output variables ($V(\text{number of inputs}) \times (\text{number of outputs})$), relations between these variables (R) and the connectivity implied by the goal states (C). Although this may seem a rather complicated way of labelling the levels of system complexity, it is designed to remind the reader of the structural characteristics of each system because the proposed levels of system complexity do not increase in a linear fashion.

All of the relations in the systems will be relevant to system control, as all output variables will have defined goal states. As in Study 1 and 2, in order to control for the influence of prior knowledge, the variables will be labelled with letters. Figure 7.2 shows a graphical representation of the underlying structure of the four systems (see Appendix E for equations).

The original system used in Study 1 and 2, which in this study is referred to as V3x3-R-6-C-2, contains six relations between three input variables and three output variables. The problem solver must consider two relations in parallel when making a decision about any of the goal states (i.e. connectivity = 2).

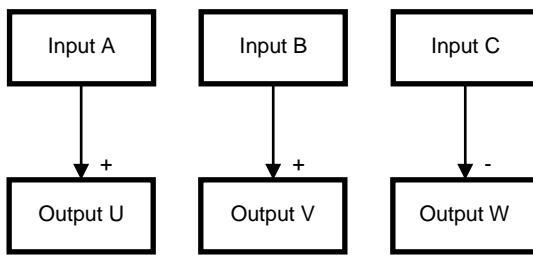
V3x3-R-3-C-1 is less complex than V3x3-R-6-C-2. It contains three relations between three inputs and three outputs. The problem solver needs only to consider the effect of a single relation on each goal state (i.e. connectivity = 1).

V6x6-R-12-C-2 is more complex than V3x3-R-6-C-2. The number of relations and input and output variables is double that of the V3x3-R-6-C-2, and consequently, problem solvers must reach six goals, rather than three. However, the connectivity is the same as V3x3-R-6-C-2 (i.e. connectivity = 2). V6x6-R-12-C-2 is similar to LINAS in terms of the number of relations in the system.

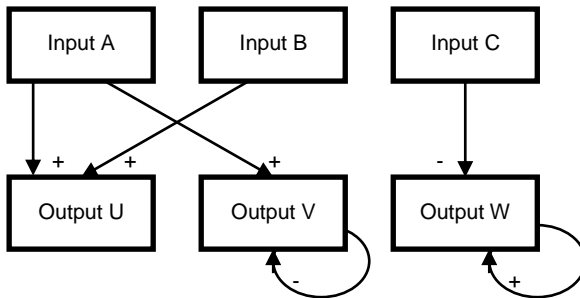
V3x3-R-7-C-3 is again more complex than V3x3-R-6-C-2, as it contains seven relations between six variables and a maximum of three relations must be considered in parallel when making a decision about a goal state (i.e. connectivity = 3). The connectivity of V3x3-R-7-C-3 is the same as LINAS.

To summarise, V3x3-R-3-C-1 is less complex than V3x3-R-6-C-2. V6x6-R-12-C-2 and V-6-R-7-C-3 are more complex than V3x3-R-6-C-2. At the present point in time, we do not have sufficient evidence to predict whether V6x6-R-12-C-2 is more complex than V-6-R-7-C-3, or whether the reverse is to be expected. V6x6-R-12-C-2 contains almost twice the number of relations of V-6-R-7-C-3, yet the connectivity of V-6-R-7-C-3 is higher than V6x6-R-12-C-2. This manipulation should allow us to determine whether the number of relations and the connectivity implied by the goal states have independent effects on the complexity of knowledge acquisition and system control.

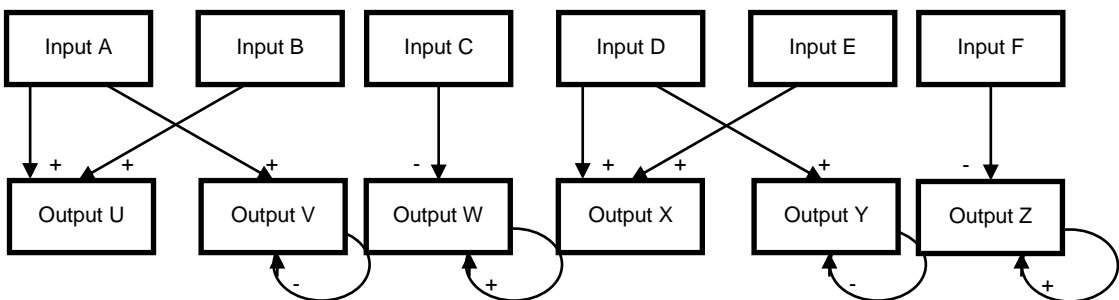
V3x3-R-3-C-1: 3 inputs, 3 outputs, 3 relations, connectivity 1



V3x3-R-6-C-2: 3 inputs, 3 outputs, 6 relations, connectivity 2



V6x6-R-12-C-2: 6 inputs, 6 outputs, 12 relations, connectivity 2



V3x3-R-7-C-3: 3 inputs, 3 outputs, 7 relations, connectivity 3

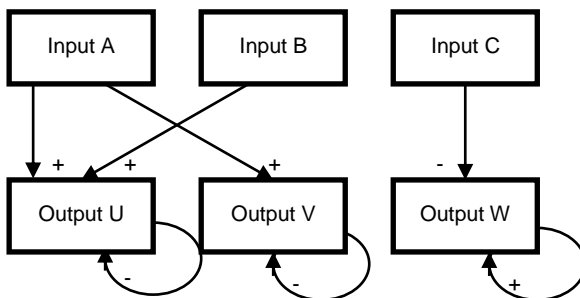


Figure 7.2: A graphical representation of the underlying structure of the four systems

In summary, the primary aim of this study is to examine the effect of structural information and system complexity on the acquisition of structural knowledge and system control. Firstly, it is hypothesised that the effect of structural information on structural knowledge will be moderated by system complexity, as it will be more difficult to convert information into knowledge in more complex systems. Secondly, system complexity will also moderate the effect of structural information on control performance, as the provision of structural information is not expected to confer an advantage to control performance at low and high levels of system complexity. Thirdly, it is hypothesised that the relationship between structural knowledge and control performance should be moderated by system complexity. The relationship should be curvilinear: With high correlations at medium levels of system complexity, and low correlations at low and high levels of system complexity, due to either floor or ceiling effects on control performance, respectively.

A secondary aim of this study is to consider the relationship between fluid intelligence and system control. Firstly, it is hypothesised that the magnitude of the relationship between fluid intelligence and control performance will increase as a function of the amount of structural information available to subjects, because the demands of the APM and the control task should be more similar when control behaviour is knowledge based rather the result of trial-and-error or random interventions. Secondly, it is hypothesised the magnitude of the relationship between fluid intelligence and control performance will increase with system complexity, because the number of processes involved in executing the control task should increase with system complexity.

The exploratory component of this investigation involves an examination of the contribution of the number of relations and connectivity implied by the goal states to system complexity. Overall, V3x3-R-3-C-1 is less complex than V3x3-R-6-C-2, which is less complex than V6x6-R-12-C-2 and V3x3-R-7-C-3. It is predicted that these differences should be reflected in structural knowledge and control performance scores and the relationship between fluid intelligence and control performance. However, at this stage, we cannot specify whether V6x6-R-12-C-2 is more or less complex than V3x3-R-7-C-3. It remains to be seen whether the number of relations, or the level of connectivity implied by the goal states both have an effect on system complexity.

7.2 Method

7.2.1 Subjects

Three hundred first-year psychology students at the University of Sydney participated for course credit ($M = 19.06$ years, $SD = 3.56$, range 17 – 55 years). Seven subjects failed to complete all of the tasks therefore their data was excluded from further analysis. A sample size of 286 would guarantee sufficient statistical power ($1 - \beta \geq .90$) in identifying at least medium effects ($f^2 = .25$) at a significance level of $\alpha \leq .05$ to detect the potential interaction effects.

7.2.2 Procedure

Subjects were randomly allocated to one of twelve experimental conditions, which determined the level of system complexity and structural information they would receive. All tasks were presented on PCs. The general procedure consisted of instructions relevant to the level of structural information (complete, partial and no information), a test of structural knowledge and two control phases. All subjects then completed the Raven's Advanced Progressive Matrices (APM).

7.2.2.1 Instructions in each condition

During the instructions subjects were not informed about the goals that they would later have to reach in the control phases. The instructional procedure used to implement the different levels of structural information was identical to that used in Studies 1 and 2. A recording of a number of intervention trials was shown to subjects with an accompanying narration, while a causal diagram was constructed on screen to record this information that remained on screen during the control phases. The number of intervention trials in the instructions differed according to the level of system complexity because the number of variables in each system differed and in the complete information condition the aim was to demonstrate the autonomic changes in the output variables and the independent effect of each input on the outputs. Thus, subjects who dealt with systems that contained six variables were shown seven intervention trials (V3x3-R-3-C-1, V3x3-R-6-C-1 and V3x3-R-7-C-1). Subjects who dealt with the system that contained twelve variables were shown thirteen intervention trials (V6x6-R-12-C-2).

In the instructions given in the complete information condition, on the first trial the inputs were set at zero, so that the autonomic changes in the outputs could be detected. Subsequently, Input A was increased to maximum while the other input variables were set at zero, and then on the next trial it was decreased to minimum while the other input variables were set at zero. This was repeated for each input variable, so that the effect of each input on the outputs could be clearly observed. After each trial, the narrator described how the inputs had been altered, how each of the outputs had changed, and how this reflected the underlying structure of the system. As in Study 1 and 2, a causal diagram was constructed concurrently on screen to record the information presented to subjects, which remained on screen during the control phases.

Subjects in the partial information condition received similar instructions. However, they were not shown the effect of input B on the outputs, nor did their causal diagram reflect this piece of information. Instead, on these trials all the inputs were increased to maximum, and then all the inputs were decreased to minimum. In the no information condition multiple inputs were varied on each trial. In this condition, at the end of each trial the narrator explained how the outputs had changed, but did not make any inferences with regards to the structure of the system, and a causal diagram was not developed on screen.

7.2.2.2 Test of structural knowledge

Subjects then had to complete a test of structural knowledge. The structural information (i.e. the causal diagram) was available on screen in the partial and complete information conditions. For each item of the structural knowledge test, subjects were shown the input variables in a particular configuration. They then had to predict whether each output variable would increase, decrease or stay the same as a result of this intervention. The items replicated the intervention pattern that was shown to subjects in the complete information condition, at the appropriate level of system complexity. Thus, subjects who dealt with systems that contained six variables completed seven items (V3x3-R-3-C-1, V3x3-R-6-C-2 and V3x3-R-7-C-3), while subjects who dealt with the system that contained twelve variables completed thirteen items (V6x6-R-12-C-2). This should allow us to determine whether problem solvers are aware of whether or not the output variables change

autonomously and how each input affects each output. This method has been used to assess structural knowledge in previous studies (e.g. Funke & Müller, 1988; Beckmann, 1994; Vollmeyer, Burns & Holyoak, 1996; Kröner et al., 2005; Bühner et al., 2008).

7.2.2.3 Control Phase 1 and 2

The third and fourth part of the task consisted of two control phases of seven trials each. Subjects had to manipulate the inputs to reach certain values of the outputs, which were indicated as lines on the output graphs. The goals for each control phase were identical to those used in Study 1 and 2. However, as V6x6-R-12-C-2 had six outputs, the goals set for outputs X, Y and Z were the same as those set for U, V and W.

During the control phases, the structural information appropriate to each condition was available on screen. On each trial subjects had to decide whether they wanted to increase, decrease or set each input variable to zero. This was done step by step, such that in V3x3-R-3-C-1, V3x3-R-6-C-2 and V3x3-R-7-C-3, subjects first set the value for Input A, then B and ultimately input C. In addition, in V6x6-R-12-C-2, subjects had to subsequently decide what to do with inputs D, E and F. Previous decisions could not be altered. Although the numerical values of the inputs were not available to subjects, the inputs could be varied incrementally, within the range of -10 to 10 units. At the end of seven trials (control phase 1), the graphs were reset, and the second set of goals were indicated on screen (control phase 2).

The key difference between the test of structural knowledge and the control phases was that the former primarily requires the reproduction of knowledge about the individual effect of each input variable on any of the output variables. In comparison, the control phases require the joint analysis of relations between multiple variables in the system and the application of the results of this analysis to manipulate the input variables in order to reach values of the output variables.

7.2.3 Dependent variables and individual differences measures

7.2.3.1 Structural knowledge

For each item, the predicted value of all output variables had to be in the correct direction in order to receive one point. Partial credit was not awarded. As the number of items completed was dependent on the number of input variables in the task, the maximum number points that could be gained was seven for V3x3-R-3-C-1, V3x3-R-6-C-2 and V3x3-R-7-C-3, and thirteen for V6x6-R-12-C-2. The scores were transformed into percentages, in order to compare the scores from the different systems.

7.2.3.2 Control performance

The scoring procedure used was the same as that used in Study 1 and 2 (see pp. 34 – 35 for a detailed example). Control performance was calculated by determining the Euclidean Distance between the actual and optimal values of the input variables. The ideal values for each input variable was calculated by using the values of the output variables on the previous trial and the goal output values to solve the set of linear equations underlying the appropriate system.⁶ A lower score indicates a smaller deviation from optimal control interventions and therefore better performance.

7.2.3.3 Fluid intelligence

Raw scores on the APM were used as an indicator of fluid intelligence. This test has been extensively validated as an indicator of fluid intelligence for a university-level population (Raven, Raven & Court, 1998).

7.3 Results

7.3.1 Internal consistencies

Firstly, internal consistency analyses were conducted to determine the variability in raw control performance scores across the trials and for different goal states as an estimate of the reliability of the dependent variables in each experimental

⁶ If the ideal values were within the range of possible input values (which was 10 to -10), then the ideal values were equal to the optimal values. In cases where the ideal value fell outside this range, then the optimal value was adjusted to the nearest possible value.

condition. Table 7.1 shows the raw control performance scores and the associated alpha coefficients, by condition.

Across the different systems, in the complete and partial information conditions, the internal consistency of the control performance scores was acceptable to excellent across the fourteen trials and individually for the first and second set of seven trials. Similarly, for V3x3-R-3-C-1 and V6x6-R-12-C-2, in the no information condition, the internal consistency was acceptable to good across the fourteen trials in the control phases and individually for the first and second set of seven trials (control phase 1 and 2, respectively). For V3x3-R-6-C-2, in the no information condition, internal consistency was poor across the fourteen trials and individually for the first and second set of seven trials. For V3x3-R-7-C-3, in the no information condition, internal consistency was poor across the fourteen trials and the first set of seven trials, but acceptable for the second set of seven trials.

The pattern of internal consistency scores suggests that the reliability of the control performance scores depends on the level of structural information and system complexity. If any information was available with regard to the underlying structure of the system or system complexity was low, then the differences between people were consistent across different trials and goal states. However, when no information was available with regard to the underlying structure of the system, and system complexity was high, the differences between people were much less consistent. This also implies that control behaviour tends to be random when the amount of information available mismatches the requirements in more complex systems. This claim will be tested further in relation to the effect of structural information on control performance.

In comparison, and as would be expected, the reliability of the random control performance scores was poor across the fourteen trials and individually for the first and second set of seven trials. This indicates that in comparison to randomly generated inputs, potential differences in control performance in the partial and complete information conditions are the result of systematic sources of variability.

An analysis of internal consistency of the APM scores indicated that the reliability of this scale was acceptable across the 36 items ($\alpha = .72$).

Table 7.1

Raw scores for control performance and alpha coefficients across the trials and for different goal states by condition

System	Information	Phase 1 (Trial 1 – 7)			Phase 2 (Trial 8 – 14)			All trials (Trials 1 – 14)		
		<i>M</i>	<i>SD</i>	α	<i>M</i>	<i>SD</i>	α	<i>M</i>	<i>SD</i>	α
V3x3-R-3-C-1	Complete n = 23	3.12	3.42	.83	2.26	1.67	.80	2.56	1.93	.86
	Partial n = 23	3.40	3.42	.92	2.26	1.67	.83	2.83	2.15	.89
	No n = 22	2.58	2.58	.79	2.50	1.52	.81	2.54	1.88	.86
	Random n = 25	12.45	1.58	.06	11.41	1.67	.28	11.93	1.03	.13
V3x3-R-6-C-2	Complete n = 25	13.32	3.95	.77	12.61	5.64	.91	12.97	4.64	.92
	Partial n = 23	14.25	3.52	.73	15.21	4.25	.76	14.73	3.47	.83
	No n = 24	16.38	3.03	.46	17.08	3.60	.67	16.73	3.05	.75
	Random n = 25	17.16	2.88	.57	17.20	2.87	.57	17.18	2.32	.66
V6x6-R-12-C-2	Complete n = 23	15.99	6.02	.90	14.34	7.21	.94	15.18	6.25	.95
	Partial n = 26	18.44	5.03	.85	19.27	6.34	.92	18.85	5.45	.94
	No n = 26	22.74	5.23	.82	22.36	6.55	.86	22.55	5.74	.92
	Random n = 25	24.44	2.31	.39	22.67	3.65	.62	23.55	2.43	.47
V3x3-R-7-C-3	Complete n = 26	12.72	3.43	.73	14.81	4.28	.83	13.77	3.41	.85
	Partial n = 26	10.02	3.31	.75	16.01	3.73	.85	13.01	2.68	.79
	No n = 26	13.43	2.37	.17	17.59	3.56	.73	15.51	2.49	.66
	Random n = 25	14.17	2.51	.43	16.08	2.87	.53	15.12	2.13	.53

7.3.2 Correcting for controllability

So that it is possible to compare control performance scores across the different systems, we need to correct for differences in the mean random control performance scores across the different systems (i.e. controllability). An analysis of variance (ANOVA) indicated that the mean random control performance scores in phase 1 and phase 2 differed significantly across the systems; $F(3, 96) = 124.73, p < .01$ and $F(3, 96) = 65.69, p < .01$. To correct this problem, subjects' control performance scores were then divided by the mean random control performance scores of the appropriate system. The corrected scores then reflect the success of subjects control performance relative to random interventions for the appropriate system.

Subsequent analyses refer to the corrected control performance scores. As before a score close to zero means that performance approximates the goal state. In addition, scores less than 1.00 indicate that performance is better than random, while scores greater than 1.00 indicate that performance is worse than random. Table 7.2 shows the means and standard deviations of the structural knowledge test scores, corrected control performance scores and APM scores by condition.

Table 7.2
Structural knowledge, corrected control performance and APM scores

System	Information	Structural Knowledge		Corrected Control Performance Phase 1		Corrected Control Performance Phase 2		APM	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
V3x3-R-3-C-1	Complete n = 23	98.18	4.32	0.23	0.21	0.16	0.14	66.81	14.10
	Partial n = 23	98.55	4.10	0.26	0.27	0.19	0.15	55.86	18.80
	No n = 22	99.24	2.45	0.19	0.20	0.20	0.14	64.34	15.90
	Random n = 25	-	-	1.00	0.13	1.00	0.16	-	-
V3x3-R-6-C-2	Complete n = 25	60.00	24.30	0.78	0.23	0.74	0.33	60.00	15.59
	Partial n = 23	54.71	16.63	0.83	0.21	0.89	0.25	64.78	13.69
	No n = 24	40.63	11.08	0.96	0.18	1.00	0.21	67.50	19.09
	Random n = 25	-	-	1.00	0.17	1.00	0.17	-	-
V-12-R-12-C-2	Complete n = 23	70.39	23.87	0.66	0.24	0.63	0.32	54.81	17.23
	Partial n = 26	65.65	24.40	0.75	0.20	0.85	0.28	58.65	15.52
	No n = 26	58.42	15.91	0.93	0.21	0.99	0.29	62.11	16.38
	Random n = 25	-	-	1.00	0.10	0.99	0.16	-	-
V3x3-R-12-C-3	Complete n = 26	49.67	22.78	0.95	0.27	0.93	0.27	54.81	17.23
	Partial n = 26	51.28	21.17	0.74	0.25	1.00	0.24	54.42	16.08
	No n = 26	48.72	17.11	0.98	0.17	1.10	0.23	62.11	16.38
	Random n = 25	-	-	1.00	0.18	1.00	0.18	-	-

7.3.3 The effect of structural information and system complexity on the acquisition of structural knowledge

In order to determine the effect of structural information and system complexity on the acquisition of structural knowledge a multiple regression analysis was conducted using effect-coded variables to conduct contrast analyses. The model contained 12 predictors of structural knowledge. Firstly, mean centred scores on the APM were included in order to control for any potential group differences in fluid intelligence, three contrasts examined the main effect of system complexity, two contrasts examined the main effect of structural information and six contrasts computed the interaction terms between these variables. It should be noted that in the construction of the system variants, system complexity was manipulated in relation to the system that was used in Study 1 and 2 (V3x3-R-6-C-2 in this study). Therefore, V3x3-R-6-C-2 is used as the baseline condition to evaluate the effect of system complexity on the outcome variables in this, and all subsequent, analyses.

The results of this analysis are reported in Table 7.3, and the interpretation of main effects and interactions are presented in the subsequent sections. These reported differences are independent of any potential group differences in fluid intelligence, although scores on the APM were a significant predictor of structural knowledge; $b = .27$, $t(280) = 4.53$, $p < .01$ $f^2 = .26$. Figure 7.3 shows structural knowledge test scores by information condition and level of system complexity.

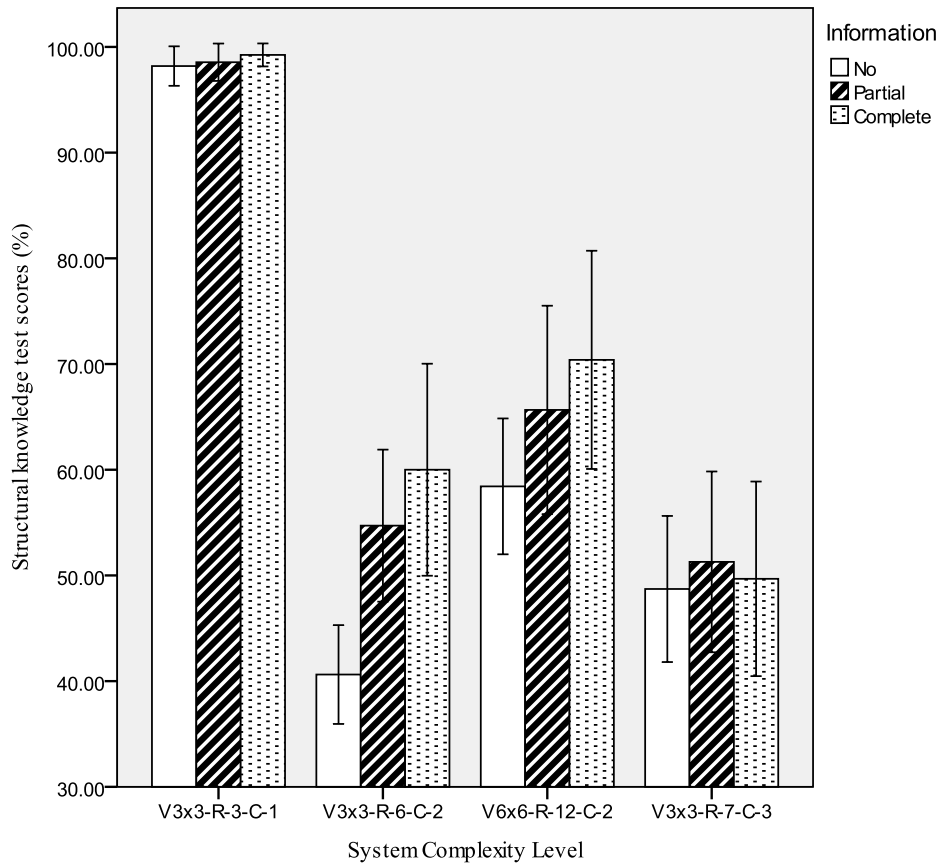


Figure 7.3: Structural knowledge test scores by information condition and system complexity (error bars represent 95% confidence intervals)

7.3.1 The effect of system complexity on the acquisition of structural knowledge

It was predicted that in comparison to V3x3-R-6-C-2, subjects would acquire more knowledge about V3x3-R-3-C-1, and less knowledge about V6x6-R-12-C-2 and V3x3-R-7-C-3. Across the levels of structural information, mean structural knowledge test scores were 98.65 ($SD = 3.71$) for V3x3-R-3-C-1, 51.58 ($SD = 19.84$) for V3x3-R-6-C-2, 64.60 ($SD = 21.91$) for V6x6-R-12-C-2 and 49.89 ($SD = 20.25$) for V3x3-R-7-C-3.

As predicted, on average, subjects who interacted with V3x3-R-3-C-1 acquired significantly more structural knowledge than subjects who interacted with V3x3-R-6-C-2; $b = 47.11$, $t(280) = 16.16$, $p < .01$, $f^2 = .70$. Against expectations, subjects who interacted with V6x6-R-12-C-3 acquired significantly more structural knowledge than subjects who interacted with V3x3-R-6-C-2; $b = 13.36$, $t(280) = 4.69$, $p < .01$, $f^2 = .27$. This is surprising considering that V6x6-R-12-C-3 contains

more than twice the number of relations in comparison to V3x3-R-6-C-2. As predicted, on average structural knowledge test scores were higher for subjects who interacted with V3x3-R-6-C-2 in comparison to V3x3-R-7-C-3, although this difference was not significant; $b = -.18$, $t(280) = -.06$, $p = .95$, $f^2 = .004$. These results show that the underlying structure of the system does have a significant impact on the amount of knowledge that subjects acquire through direct instruction. However, more relations per se does not seem to make the task of acquiring structural knowledge more difficult, rather, there is a trend that suggests that it is more difficult to acquire knowledge about systems with a higher level of connectivity implied by the goal states.

7.3.3.2 The effect of structural information on the acquisition of structural knowledge

It was predicted that the amount of structural knowledge acquired by subjects would be a function of the amount of structural information that was available to them. Across the levels of system complexity, mean knowledge test scores were 68.68 ($SD = 27.47$) in the complete information condition, 66.99 ($SD = 25.96$) in the partial information condition and 60.79 ($SD = 25.30$) in the no information condition.

On average, subjects who received any (complete or partial) structural information acquired more structural knowledge than those who received no information; $b = -6.05$, $t(280) = -4.20$, $p < .01$, $f^2 = .24$. Although the means were in the expected direction, the amount of knowledge acquired by subjects in the complete information condition did not differ significantly from the partial information condition; $b = -3.58$, $t(280) = -1.43$, $p = .15$, $f^2 = .09$. This indicates that those who received any structural information acquired more knowledge than those who did not receive such instructions. However, increments in the amount of information available to subjects did not have a significant impact on the amount of structural knowledge acquired. This is perhaps unsurprising, considering that the instructions in the partial information condition only omitted information about one relationship in the system.

7.3.3.3 Interaction between structural information and system complexity on the acquisition of structural knowledge

The main effects of system complexity and structural information on structural knowledge, however, are characterised by significant interaction effects, as depicted in Figure 7.3. The effect of structural information on the amount of knowledge that subjects acquired was much stronger in V3x3-R-6-C-2 in comparison to V3x3-R-3-C-1; $b = 5.53$, $t(280) = 2.67$, $p < .01$, $f^2 = .16$. An inspection of the means reveals that for subjects interacting with the least complex system (V3x3-R-3-C-1) structural knowledge test scores were close to ceiling level across the information conditions. This suggests that subjects interacting with this system who received only partial or no information were able to independently infer the complete underlying structure of the system.

Similarly, the effect of structural information on the amount of knowledge that subjects acquired was much stronger for V3x3-R-6-C-2 in comparison to V3x3-R-7-C-3; $b = 4.75$, $t(280) = 2.38$, $p < .01$, $f^2 = .14$. An inspection of the means reveals that for the system with the highest level of connectivity (V3x3-R-7-C-3), structural knowledge test scores were uniformly poor across the information conditions. This suggests that when the level of connectivity was high, subjects who were given information were unable to convert it into knowledge.

The effect of structural information on the amount of knowledge that subjects acquired was similar for V3x3-R-6-C-2 in comparison to V6x6-R-12-C-2. An inspection of the means reveals that under these conditions, the impact of information on the acquisition of structural knowledge is in the expected direction, with more structural information being associated with better structural knowledge test scores. As V6x6-R-12-C-2 has twice the number of relations of V3x3-R-6-C-2, this indicates that the number of relations in a system does not have the expected detrimental impact on knowledge acquisition.

In summary, we expected that the effect of structural information on the acquisition of structural knowledge would diminish as system complexity (as indicated by the number of relations or connectivity) increased. This expectation was partially supported as it was found that the connectivity implied by the goal states moderates the impact of structural information on the acquisition of structural

knowledge, and that this effect is curvilinear. The results show that the provision of structural information conferred little advantage to the acquisition of structural knowledge at low (V3x3-R-3-C-1) and high (V3x3-R-7-C-3) levels of connectivity. Information did confer an advantage to structural knowledge when the system was at a medium level of connectivity (V3x3-R-6-C-2 and V6x6-R-12-C-2). The number of relations in the system did not have a consistent impact on the acquisition of structural knowledge. This indicates that the number of relations in a system is not a good indicator of system complexity in relation to knowledge acquisition.

Table 7.3

Results of a regression analysis examining the effect of structural information and system complexity on structural knowledge test scores

Variables entered	β	t	p	f^2
APM	.283	4.53	< .01	.26
<i>Main effect: System complexity</i>				
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	47.11	16.16	< .01	.70
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	13.36	4.69	< .01	.27
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	-.18	-.06	.95	.004
<i>Main effect: Structural information</i>				
No vs. Any information	-6.05	-4.20	< .01	.24
Partial vs. Complete information	-3.58	-1.43	.15	.09
<i>Interaction effects</i>				
(No vs. Any information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	5.53	2.67	< .01	.16
(No vs. Any information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	2.07	1.04	.30	.06
(No vs. Any information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	4.75	2.38	< .05	.14
(Partial vs. Complete information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	4.78	1.32	.19	.08
(Partial vs. Complete information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	1.40	.40	.69	.02
(Partial vs. Complete information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	4.43	1.28	.20	.08

7.3.4 The effect of structural information and system complexity on the quality of system control

7.3.4.1 Comparisons to random control performance scores

Firstly, in order to appropriately interpret the effect of structural information on control performance, it is necessary to compare the control performance scores in each condition to random control performance scores. If control performance were in fact not better than random then it would be inappropriate to claim that efficient control performance is dependent on the amount of structural information that is available, regardless of whether subjects who receive structural information perform

better than those who do not receive such information. Therefore, to determine whether subjects' control performance was better than random, a series of multiple regressions were conducted, using dummy coded variables to compare performance in each information condition to the random performance scores for each system. Note that lower control scores indicate better performance.

Subjects controlling V3x3-R-3-C-1 in the complete, partial and no information conditions performed significantly better than random in both control phases; $b_{complete} = -.77$, $t(89) = -12.61$, $p < .01$, $f^2 = .80$, $b_{partial} = -.74$, $t(89) = -12.25$, $p < .01$, $f^2 = .79$ and $b_{no} = -.81$, $t(89) = -13.38$, $p < .01$, $f^2 = .82$, for the first phase, and $b_{complete} = -.84$, $t(89) = -19.26$, $p < .01$, $f^2 = .89$, $b_{partial} = -.81$, $t(89) = -19.26$, $p < .01$, $f^2 = .90$ and $b_{no} = -.80$, $t(89) = -18.85$, $p < .01$, $f^2 = .89$, for the second phase.

Subjects controlling V3x3-R-6-C-2 in the complete information condition performed significantly better than random across both phases; $b = -.22$, $t(93) = -4.01$, $p < .01$, $f^2 = .38$ and $b = -.27$, $t(93) = -3.78$, $p < .01$, $f^2 = .37$. Similarly, in the partial information condition, subjects performed significantly better than random in the first phase; $b = -.17$, $t(93) = -2.99$, $p < .01$, $f^2 = .29$. However, although the means are in the expected direction, subjects did not perform significantly better than random in the second phase; $b = -.11$, $t(93) = -1.57$, $p = .12$, $f^2 = .16$. Subjects in the no information condition did not perform better than random in either control phase; $b = -.04$, $t(93) = -.63$, $p = .53$, $f^2 = .07$ and $b = -.01$, $t(93) = -.07$, $p = .94$, $f^2 = .01$.

Subjects controlling V6x6-R-12-C-2 in the complete information condition performed significantly better than random across both phases; $b = -.34$, $t(96) = -6.13$, $p < .01$, $f^2 = .53$ and $b = -.36$, $t(96) = -4.63$, $p < .01$, $f^2 = .43$. Similarly, in the partial information condition, subjects performed significantly better than random in the first phase; $b = -.25$, $t(96) = -4.50$, $p < .01$, $f^2 = .41$. However, although the means are in the expected direction, performance was not significantly better than random in the second phase; $b = -.14$, $t(96) = -1.89$, $p = .06$, $f^2 = .19$. Subjects in the no information condition did not perform better than random in either control phase; $b = -.07$, $t(96) = -1.27$, $p = .21$, $f^2 = .13$ and $b = -.01$, $t(96) = -.07$, $p = .94$, $f^2 = .01$.

Subjects controlling V3x3-R-7-C-3 in the complete information condition did not perform better than random in either phase; $b = -.05$, $t(99) = -.81$, $p = .42$, $f^2 = .08$ and $b = -.07$, $t(99) = -1.10$, $p = .27$, $f^2 = .11$, respectively. Subjects in the partial

information condition performed significantly better than random in the first phase, $b = -.27$, $t(99) = -4.29$, $p < .01$, $f^2 = .40$, but not in the second phase; $b = .01$, $t(99) = -.05$, $p = .96$, $f^2 = .01$. Subjects in the no information condition did not perform better than random in either phase; $b = -.02$, $t(99) = -.37$, $p = .71$, $f^2 = .04$ and $b = -.10$, $t(99) = 1.60$, $p = .11$, $f^2 = .16$, respectively.

In summary, these results suggest that the effect of structural information is dependent on the complexity of the system, and in particular, on the connectivity of the system. When the connectivity was low (i.e. V3x3-R-3-C-1), successful control performance did not depend on access to structural information. Rather, subjects performed better than random regardless of the amount of information that was available. At medium levels of connectivity (V3x3-R-6-C-2 and V6x6-R-12-C-2), complete structural information was a pre-condition of better than random performance. However, at high levels of connectivity (i.e. V3x3-R-7-C-3), even those provided with complete or partial structural information did not consistently perform better than random. This suggests that any differences between the information conditions for subjects interacting with V3x3-R-7-C-3 should be interpreted cautiously; as subjects who received information did not on average perform better than random, the claim cannot be made that structural information confers an advantage to control performance under these conditions. These results support our hypotheses, and extend upon the findings reported earlier, which showed that the effect of structural information on the acquisition of structural knowledge was moderated by system complexity.

7.3.4.2 The effect of information condition on the quality of control performance

In order to determine whether structural information conferred an advantage to control performance, and whether this effect was moderated system complexity, multiple regression analyses were conducted using effect-coded variables to conduct contrast analyses. These analyses excluded the random control performance scores and the model contained 12 predictors of control performance. Mean centred scores on the APM were included in order to control for any potential group differences in fluid intelligence, two contrasts examined the main effect of structural information, three contrasts examined the main effect of system complexity and six contrasts computed the interaction terms between these variables. The results of these analyses

are reported in Tables 7.4 and 7.5, and the interpretation of the main effects and interactions are presented in the subsequent sections. These results are independent of any potential group differences in fluid intelligence, although scores on the APM were a significant predictor of performance across both control phases; $b = -.01$, $t(280) = -4.81$, $p < .01$, $f^2 = .28$ and $b = -.01$, $t(280) = -6.03$, $p < .01$, $f^2 = .34$, respectively. Figure 7.4 and 7.5 shows the corrected control performance scores, by information condition and level of system complexity.

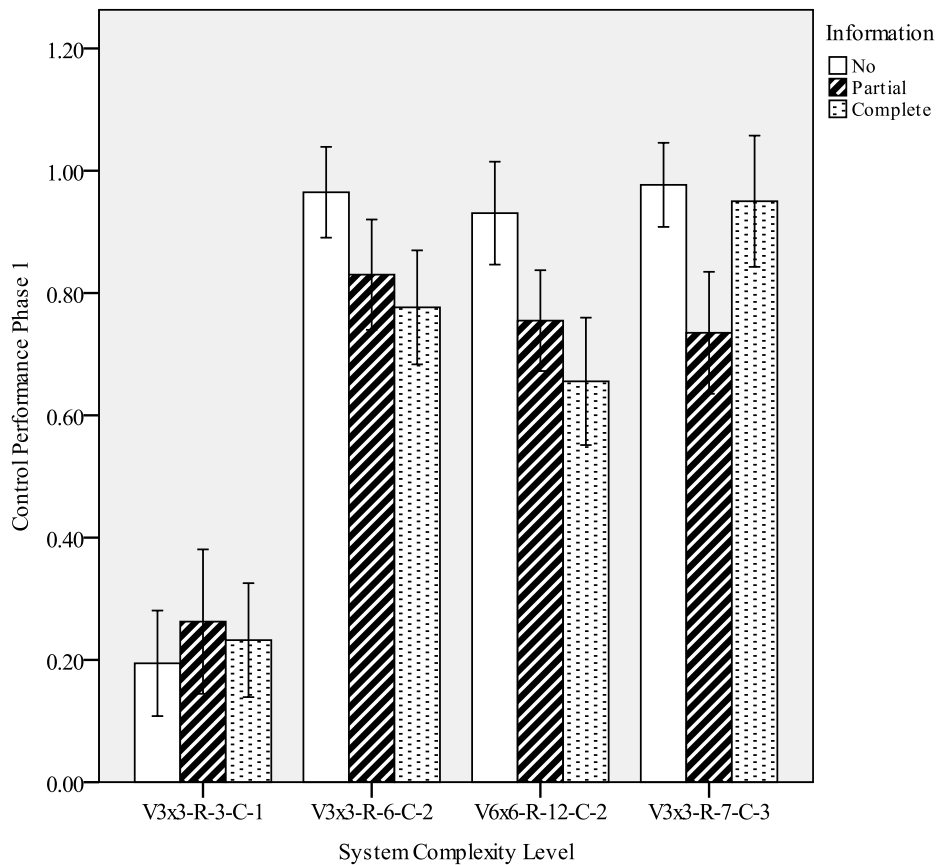


Figure 7.4: Corrected control performance scores in phase 1 by information condition and system complexity (error bars represent 95% confidence intervals)

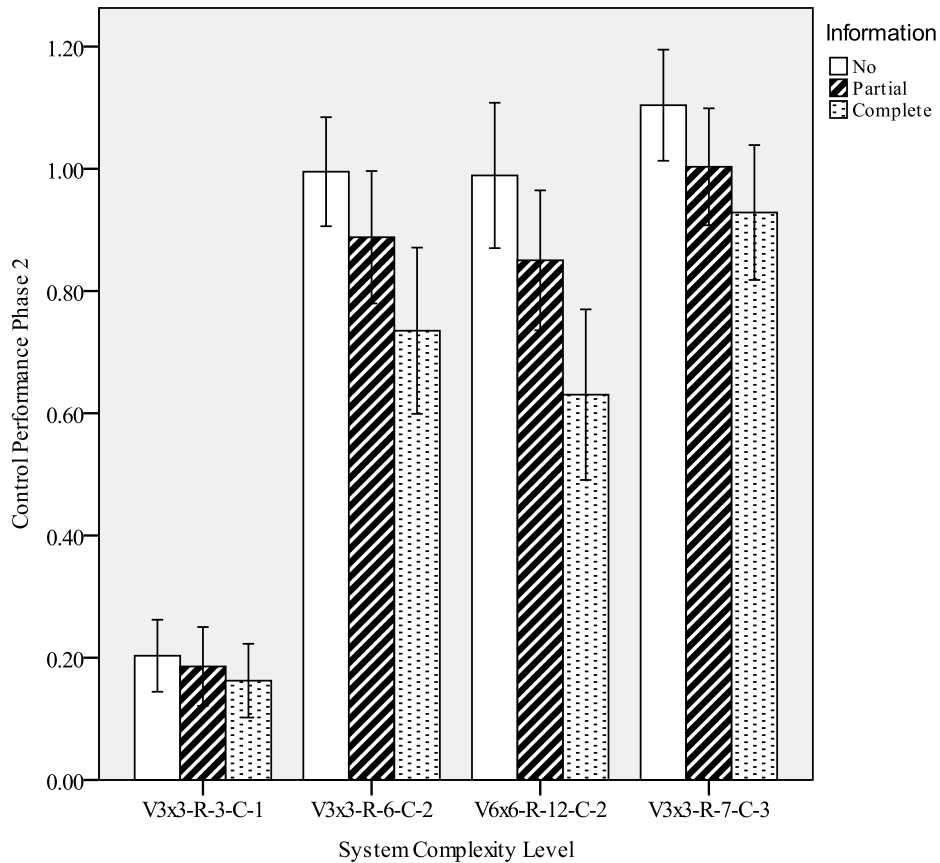


Figure 7.5: Corrected Control performance scores in phase 2 by information condition and system complexity (error bars represent 95% confidence intervals)

7.3.4.3 The effect of system complexity on system control

It was predicted that in comparison to V3x3-R-6-C-2, subjects would have better control over V3x3-R-3-C-1, and worse control over V6x6-R-12-C-2 and V3x3-R-7-C-3. For V3x3-R-3-C-1, the mean control performance scores across the levels of structural information were .23 ($SD = .23$) in the first phase and .18 ($SD = .14$) in the second phase. For V3x3-R-6-C-2, the mean control performance scores were .86 ($SD = .22$) in the first phase, and .87 ($SD = .29$) in the second phase. For V6x6-R-12-C-2, the mean control performance scores were .79 ($SD = .24$) and .83 ($SD = .33$). For V3x3-R-7-C-3, mean control performance scores were .89 ($SD = .25$) in the first phase and 1.01 ($SD = .25$) in the second phase.

In line with our predictions, subjects had significantly better control over V3x3-R-3-C-1, in comparison to V3x3-R-6-C-2 in phase 1 and 2; $b = -.63$, $t(280) = -17.49$, $p < .01$, $f^2 = .72$ and $b = -.69$, $t(280) = -17.51$, $p < .01$, $f^2 = .72$, respectively. Against expectations, subjects had significantly better control over V6x6-R-12-C-2,

in comparison to V3x3-R-6-C-2 in phase 1; $b = -.08$, $t(280) = -2.30$, $p < .05$, $f^2 = .14$. This result was particularly surprising, considering that V6x6-R-12-C-2 has twice the number of relations of V3x3-R-6-C-2. However, although it was in the same direction, this difference was not significant in phase 2; $b = -.06$, $t(280) = -1.43$, $p = .16$, $f^2 = .09$. As expected, subjects had better control over V3x3-R-6-C-2 in comparison to V3x3-R-7-C-3, although this difference was only significant in phase 2; $b = .01$, $t(280) = .22$, $p = .83$, $f^2 = .01$ and $b = .11$, $t(280) = 2.81$, $p < .01$, $f^2 = .17$, respectively. These results suggest that systems that contain more relations are not more difficult to control. Rather, there is a trend that suggests that systems that entail a higher level of connectivity are more difficult to control.

7.3.4.4 The effect of structural information on system control

It was predicted that the quality of subjects' control performance would be a function of the amount of structural information that was available to them, with more information being associated with better control performance. In the complete information condition, across the levels of system complexity, the mean control performance scores in the complete information condition were .67 ($SD = .35$) in the first phase and .63 ($SD = .39$) in the second phase. In the partial information condition, the mean control performance scores were .65 ($SD = .32$) in the first phase and .74 ($SD = .39$) in the second phase. In the no information conditions, the mean control performance scores were .78 ($SD = .37$) in the first phase, and .84 ($SD = .42$) in the second phase.

In line with our predictions, in the first and second phases, control performance when any information was available (complete and partial) was significantly better than when no information was available; $b = .06$, $t(280) = 3.37$, $p < .01$, $f^2 = .20$, and $b = .07$, $t(280) = 3.56$, $p < .01$, $f^2 = .21$, respectively. We predicted that control performance in the complete information condition should be better than in the partial information condition. In the first control phase, this difference was not significant, although the means are in the expected direction; $b = .04$, $t(280) = 1.26$, $p = .21$, $f^2 = .08$. However, in the second control phase, control performance in the complete information condition was significantly better than in the partial information condition; $b = .09$, $t(280) = 2.75$, $p < .01$, $f^2 = .16$. Overall, these results show that structural information shows an advantage to control performance, with a

trend that suggests that increasing levels of structural information results in better control performance.

7.3.4.5 Interaction between structural information and system complexity on the quality of system control

These main effects on control performance, however, are further qualified by significant interaction effects between the level of structural information and system complexity, as depicted in Figures 7.4 and 7.5. The results indicate that the difference in control performance for subjects who received any information in comparison to no information was much larger for V3x3-R-6-C-2 in comparison to V3x3-R-3-C-1 in phase 1 and 2; $b = -.07$, $t(280) = -2.90$, $p < .01$, $f^2 = .17$ and $b = -.06$, $t(280) = -1.96$, $p < .05$, $f^2 = .12$, respectively. Similarly, the difference in control performance for subjects who received complete information in comparison to partial information was much larger for V3x3-R-6-C-2 in comparison to V3x3-R-3-C-1 in phase 2; $b = -.11$, $t(280) = -2.23$, $p < .05$, $f^2 = .13$. These interactions support the findings observed in relation to the comparisons to the random control performance scores. For subjects controlling V3x3-R-3-C-1, across the information conditions, control performance was close to ceiling level, and all subjects performed better than random. In comparison, for subjects controlling V3x3-R-6-C-2, structural information was a pre-condition of better than random performance, and hence the difference between the information conditions in terms of control performance is much larger than for V3x3-R-3-C-1.

The effect of structural information on control performance was similar for V3x3-R-6-C-2 in comparison to V6x6-R-12-C-2. An inspection of the means reveals that under these conditions, the impact of information on control performance is in the expected direction: With more structural information being associated with better control performance. Again, this supports the findings reported in relation to random control performance scores, as subjects controlling these systems only performed better than random when structural information was provided. Thus, we can infer that under this level of system complexity, structural information does have a beneficial impact on control performance.

The difference in control performance for subjects who received complete information in comparison to partial information was much larger for V3x3-R-6-C-2

in comparison to V3x3-R-7-C-3 in phase 2; $b = -.15$, $t(280) = -3.45$, $p < .01$, $f^2 = .20$. Again, this supports the findings in relation to the comparisons to random performance scores, as control performance for subjects controlling V3x3-R-7-C-3 was uniformly poor across the levels of structural information. In comparison, subjects controlling V3x3-R-6-C-2 who received structural information performed better than random.

Overall, these results indicate that system complexity, as indicated by the connectivity implied by the goal states, to some extent determines whether structural information will be utilised. The number of relations does not seem to be a good indicator of system complexity in relation to system control. In contrast, the connectivity implied by the goal states appears to be systematically related to system complexity in relation to system control. At low and high levels of connectivity structural information does not show an advantage to control performance. In comparison, at medium levels of connectivity, structural information has the expected effect on control performance. These results support our hypothesis and further clarify the relative contribution of the number of relations and connectivity to system complexity.

Table 7.4
Results of a multiple regression analysis examining the effect of structural information, system complexity and the interaction between these variables on control performance in phase 1

Variables entered	β	t	p	f^2
APM	-.01	-4.81	< .01	.28
<i>Main effect: System complexity</i>				
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.63	-17.49	< .01	.72
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.08	-2.30	< .05	.14
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.01	.22	.83	.01
<i>Main effect: Structural information</i>				
No vs. Any information	.06	3.37	< .01	.20
Partial vs. Complete	.04	1.26	.21	.08
<i>Interaction effects</i>				
(No vs. Any information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	-.07	-2.90	< .01	.17
(No vs. Any information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.03	1.02	.31	.06
(No vs. Any information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	-.01	-.24	.81	.01
(Partial vs. Complete information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	-.04	-.99	.32	.06
(Partial vs. Complete information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.01	.19	.86	.01
(Partial vs. Complete information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	-.15	-3.45	< .01	.20

Table 7.5

Results of a multiple regression analysis examining the effect of structural information, system complexity and the interaction between these variables on control performance in phase 2

Variables entered	β	t	p	f^2
APM	-.01	-6.03	< .01	.34
<i>Main effect: System complexity</i>				
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.69	-17.51	< .01	.72
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.06	-1.43	.16	.09
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.11	2.81	< .01	.17
<i>Main effect: Structural information</i>				
No vs. Any information	.07	3.56	< .01	.21
Partial vs. Complete information	.09	2.75	< .01	.16
<i>Interaction effects</i>				
(No vs. Any information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	-.06	-1.96	< .05	.12
(No vs. Any information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.02	.10	.32	.06
(No vs. Any information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	-.01	-.40	.69	.02
(Partial vs. Complete information) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	-.11	-2.23	< .05	.13
(Partial vs. Complete information) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.01	.28	.78	.02
(Partial vs. Complete information) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	-.06	-1.21	.23	.07

7.3.5 The relationship between structural knowledge and system control

It was predicted that the relationship between structural knowledge and control performance would be weaker at lower and higher levels of system complexity in comparison to V3x3-R-6-C-2. In order to test this prediction, a moderator analysis was conducted for each control phase. The model included: The information condition to control for the effect of the experimental manipulation of knowledge, effect-coded variables to compare control performance in V3x3-R-6-C-2 to each of the other systems, mean centred structural knowledge test scores and the interactions between system complexity and structural knowledge test scores. The results of these analyses are reported in Table 7.6 below. Overall, the results indicate that controlling for the level of structural information and system complexity, subjects who acquired more structural knowledge had better control performance across both phases; $b = -.006$, $t(284) = -4.35$, $p < .01$, $f^2 = .25$ and $b = -.007$, $t(284) = -4.73$, $p < .01$, $f^2 = .27$, respectively.

In line with our expectations, in the first control phase, this result is qualified by a significant interaction effect that indicates that the strength of this relationship is stronger in V3x3-R-6-C-2 than in V3x3-R-7-C-3; $b = .005$, $t(284) = 2.82$, $p < .01$, f^2

= .17. The findings reported in the preceding sections may account for the weaker relationship between structural knowledge and control performance observed for V3x3-R-7-C-3. Subjects dealing with this system were unable to convert information into knowledge and subjects did not perform better than random regardless of the amount of information that they received. It is therefore not surprising that the relationship between structural knowledge and control performance should be weak: Control interventions cannot be knowledge-based when knowledge has not been acquired. However, these results should not be over interpreted, as this effect was not significant in the second control phase.

The relationship between structural knowledge and control performance was similar for V3x3-R-6-C-2 in comparison to V3x3-R-3-C-1 and V6x6-R-12-C-2 across both phases. Overall, these results do not support the hypothesis that the relationship between structural knowledge and control performance is moderated by system complexity. Rather, it appears that the amount of structural knowledge that subjects acquire is a significant predictor of performance regardless of the complexity of the system.

Table 7.6
Results of a moderator analysis examining the relationship between structural knowledge and control performance by system complexity for phase 1 and 2, respectively

Variables entered	β	t	p	f^2
<i>Control Phase 1</i>				
Information condition	-.036	-2.22	.03	.13
Structural knowledge	-.006	-4.35	< .01	.25
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.550	-2.23	< .05	.13
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	.003	.07	.94	< .01
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.099	2.22	< .05	.13
(Structural knowledge) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.006	.80	.43	.05
(Structural knowledge) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.001	.69	.49	.04
(Structural knowledge) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.005	2.82	< .01	.17
<i>Control Phase 2</i>				
Information condition	-.080	-4.68	< .01	.27
Structural knowledge	-.007	-4.73	< .01	.27
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.516	-2.01	< .05	.12
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	.041	.97	.33	.06
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.172	3.67	< .01	.21
(Structural knowledge) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.004	.54	.59	.03
(Structural knowledge) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.001	.32	.75	.02
(Structural knowledge) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.003	1.50	.13	.09

7.3.6 The relationship between fluid intelligence and system control by information condition

It was predicted that the magnitude of the relationship between fluid intelligence and control performance would increase with the amount of information available to subjects. As can be seen in Table 7.7, across both control phases the relationship between fluid intelligence and control performance is strong in the complete information condition, while it is weak in the partial and no information conditions.

Table 7.7
The relationship between APM scores and control performance in each phase, by information condition

	Complete information	Partial information	No information
Phase 1	-.42**	-.10	-.11
Phase 2	-.43**	-.16	-.17

** $p < .01$

In order to test whether the strength of the relationship between fluid intelligence and control performance varied across the information conditions two moderator analyses were conducted, which are presented in Table 7.8. The significant interaction terms indicate that the relationship between scores on the APM and control performance was significantly moderated by the amount of information available to subjects across both phases, $b = -.004$, $t(289) = -2.45$, $p < .05$, $f^2 = .14$, and $b = -.003$, $t(289) = -1.91$, $p < .05$, $f^2 = .11$. This indicates that the strength of the relationship between fluid intelligence and control performance increased as the amount of information available to subjects increased. Thus, subjects with a higher level of fluid intelligence were at a greater advantage in terms of control performance when more structural information was available.

Table 7.8

Results of two moderator analyses examining the dependency of the relationship between control performance and fluid intelligence (APM score) for phase 1 and 2.

Step	Variables entered	β	t	p	R^2_{change}	F
<i>Control Phase 1</i>						
1	Information condition	-.066	-2.66	< .01		
	APM	-.004	-3.27	< .01		
					.052	7.88
2	Information condition	-.062	-2.52	< .05		
	APM	.003	.96	.340		
	Information condition X APM	-.004	-2.45	< .05		
					.019	6.01
<i>Control Phase 2</i>						
1	Information condition	-.119	-4.25	< .01		
	APM	-.006	-4.28	< .01		
					.100	16.06
2	Information condition	-.116	-4.14	< .01		
	APM	.000	.07	.949		
	Information condition X APM	-.003	-1.91	< .05		
					.011	3.65

Note: The degrees of freedom for the F-test of R^2_{change} in the step 1 models are (2, 290) and for the step 2 models are (1, 289).

7.3.7 The relationship between fluid intelligence and system control by system complexity

It was predicted that the magnitude of the relationship between fluid intelligence and control performance would increase with system complexity, as long as there were no floor effects on control performance. As can be seen in Table 7.9, across both control phases, the relationship between fluid intelligence and control performance appears to be weak to moderate across the levels of system complexity.

Table 7.9

The relationship between APM scores and control performance in each phase, by system complexity

	V3x3-R-3-C-1	V3x3-R-6-C-2	V6x6-R-12-C-2	V3x3-R-7-C-3
Phase 1	-.23	-.33*	-.27*	.01
Phase 2	-.19	-.25*	-.29*	-.25*

* $p < .05$

However, as the relationship between fluid intelligence and control performance varies according to the amount of structural information available to subjects (see section 7.3.6) we must also take this into account. Therefore, in order to test whether the strength of the relationship between fluid intelligence and control

performance differed according to the level of system complexity, a moderator analysis was conducted for each control phase at each level of structural information. The model included: a) mean centred scores on the APM, b) dummy-coded variables to compare control performance in V3x3-R-6-C-2 to each of the other systems and c) the interaction between these variables to compare the differences in the strength of the relationship between the APM and control performance for each of the systems. The results are reported in Table 7.10 for the no information condition, in Table 7.11 for the partial information condition and in Table 7.12 for the complete information condition. Against expectations, the interaction terms are not significant, which indicates that controlling for the amount of structural information that subjects received, the strength of the relationship between fluid intelligence and control performance did not vary across the levels of system complexity.

Table 7.10
Results of moderator analyses examining the relationship between fluid intelligence (APM score) and control performance by system complexity for the no information condition

Variables entered	β	t	p	f^2
<i>Control Phase 1</i>				
APM	-.003	-1.71	.091	.18
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.782	-14.16	< .01	.83
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.022	-.40	.688	.04
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	-.007	-.13	.899	.01
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.001	.47	.638	.05
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	-.001	-.49	.628	.05
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.004	1.32	.191	.14
<i>Control Phase 2</i>				
APM	-.004	-1.63	.106	.17
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.803	-12.34	< .01	.79
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	.015	.23	.817	.02
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.091	1.46	.148	.15
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.001	.28	.781	.03
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	-.003	-.82	.416	.09
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.001	.24	.813	.03

Table 7.11

Results of moderator analyses examining the relationship between fluid intelligence (APM score) and control performance by system complexity for the partial information condition

Variables entered	β	t	p	f^2
<i>Control Phase 1</i>				
APM	-.007	-2.04	< .05	.21
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.612	-8.75	< .01	.68
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.114	-1.70	.092	.18
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	-.128	-1.85	.068	.19
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.004	.86	.393	.09
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.002	.48	.631	.05
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.006	1.40	.167	.15
<i>Control Phase 2</i>				
APM	-.009	-2.62	< .01	.27
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.742	-10.95	< .01	.76
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.086	-1.32	.189	.14
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.036	.54	.590	.06
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.008	1.85	.068	.19
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.003	.65	.519	.07
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.002	.50	.622	.05

Table 7.12

Results of moderator analyses examining the relationship between fluid intelligence (APM score) and control performance by system complexity for the complete information condition

Variables entered	β	t	p	f^2
<i>Control Phase 1</i>				
APM	-.008	-2.57	< .05	.26
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.499	-7.18	< .01	.61
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.105	-1.59	.115	.17
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.185	2.80	.006	.29
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.005	.99	.322	.11
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	.001	.21	.831	.02
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.006	1.52	.133	.16
<i>Control Phase 2</i>				
APM	-.008	-2.13	< .05	.22
V3x3-R-3-C-1 vs. V3x3-R-6-C-2	-.544	-6.72	< .01	.58
V6x6-R-12-C-2 vs. V3x3-R-6-C-2	-.092	-1.21	.231	.13
V3x3-R-7-C-3 vs. V3x3-R-6-C-2	.186	2.42	< .05	.25
(APM) X (V3x3-R-3-C-1 vs. V3x3-R-6-C-2)	.007	1.30	.197	.14
(APM) X (V6x6-R-12-C-2 vs. V3x3-R-6-C-2)	-.002	-.45	.653	.05
(APM) X (V3x3-R-7-C-3 vs. V3x3-R-6-C-2)	.003	.66	.510	.07

7.4 Discussion

In summary, the results indicate that the effect of structural information on the acquisition of structural knowledge and control performance is dependent on system complexity, as indicated by the connectivity implied by the goal states. It was more difficult for subjects to understand and utilise information in relation to systems with higher levels of connectivity. The effect of the number of relations on knowledge acquisition and system control was inconsistent. The strength of the relationship between fluid intelligence and control performance increased with the amount of information available to subjects, which suggests that the demands of system control most closely match those required in traditional measures of intelligence when control behaviour is knowledge-based. The strength of the relationship between fluid intelligence and control performance did not vary as a function of system complexity, which suggests that more intelligent problem solvers have a general advantage over their less intelligent counter-parts regardless of the complexity of the system.

The results elucidate important boundary conditions for the effect of structural information on control performance. Consistent with the results reported in Study 1 and 2, when the system consisted of 6 relations between 6 variables, and two relations had to be considered in parallel, only subjects provided with structural information performed better than random. This result was shown to generalise to a system that entailed the same level of connectivity, but twice the number of variables and relations. However, subjects who interacted with the system that had a lower level of connectivity and half the number of relations performed better than random regardless of the amount of information they received. In contrast, subjects who interacted with the system that required the consideration of three relations in parallel, and consisted of 7 relations between 6 variables had poor control performance regardless of the amount of information that they received. Overall, the results indicate that the beneficial effect of structural information on control performance is dependent on the complexity of the system.

This may explain why previous studies have found that structural information did not benefit control performance (Putz-Osterloh, 1993; Preußler, 1996). In both LINAS and V3x3-R-7-C-3 three relations must be considered in parallel in order to

make a decision about a goal state. The results of the current study show that at this level of system complexity, it is difficult for problem solvers to acquire structural knowledge through direct instruction or utilise structural information to control the outcomes of dynamic systems. This suggests that the level of system complexity entailed in LINAS may have been too high for problem solvers to effectively make use of structural information.

However, the criticism that the instructional method used in previous studies may have been inadequate remains valid. In the current study, subjects interacting with V3x3-R-7-C-3 failed to convert information into structural knowledge. In effect, this demonstrates that at this level of system complexity, even a direct demonstration of how each input affects each output, combined with a verbal explanation of the material and the presentation of the information on screen while the control task is performed, is insufficient to promote structural knowledge. The question that remains unanswered is whether any form of direct instruction would result in the acquisition of knowledge, and subsequently influence control performance, in systems with highly complex structures. It may be that under such conditions, problem solvers require an extended period of goal-orientated practice in order to acquire knowledge and chunk it into useable components. Effectively, this would reduce the information processing demands of the task (i.e. complexity) and hence the overall level of difficulty experienced by the problem solver.

In Study 2, the effect of structural knowledge on control performance could only be inferred indirectly through the effect information on control performance. However, in the current study the test of structural knowledge directly after the instructions allowed us to test whether efficient control performance is dependent on the acquisition of structural knowledge. It was found that better than random performance was consistently associated with the acquisition of a high level of structural knowledge prior to control performance, and under these conditions the relationship between structural knowledge and control performance was strong. Concomitantly, performance was not better than random when little knowledge was acquired prior to the control phase, and under these conditions the relationship between structural knowledge and control performance was weak. These results clarify the results observed in Study 2, and demonstrate that successful control performance is dependent on the acquisition of structural knowledge.

However, an alternative conclusion might be drawn from the control performance scores of subjects who interacted with the low complexity system (V3x3-R-3-C-1) and did not receive structural information. These subjects had almost perfect control over the system. This could be interpreted as evidence that systems at a low level of complexity can be controlled without structural knowledge, through a process of trial-and-error. However, it seems likely that subjects' efficient control interventions were knowledge-based under these conditions. Firstly, their structural knowledge test scores were almost perfect, indicating that they acquired almost complete structural knowledge prior to the control phase. Secondly, their control performance scores showed high levels of internal consistency, which would not be expected if they were learning to control the system through trial-and-error. Thirdly, the relationship between structural knowledge and control performance was similar to that observed in relation to V3x3-R-6-C-2, which suggests that the amount of knowledge that subjects acquired determined the quality of their control performance. Taken together, these findings suggest that at low levels of system complexity problem solvers can easily infer the underlying structure of the system without assistance, and that efficient control actions are the result of knowledge-based interventions.

Nevertheless, it was particularly surprising that subjects dealing with the low complexity system in the no information condition acquired almost perfect structural knowledge. In the instructional phase in the no information condition multiple variables were changed on each trial. This pattern of interventions was selected to prevent subjects from independently inferring the underlying structure of the system. This suggests that the relationship between VOTAT strategy use and the acquisition of structural knowledge is dependent on the complexity of the system. That is, a strong relationship between the two will only be found when the system is sufficiently complex to require a systematic testing strategy for the acquisition of knowledge.

With regard to the relationship between fluid intelligence and control performance, it was found that the strength of this relationship increased with the amount of structural information available to subjects. These results cross-validate those reported in Study 2, and support the hypothesis put forward in Chapter 4 that control performance under different levels of knowledge has differential validity.

That is, the processes used to control the systems under different levels of knowledge may differ.

The strong relationship between control performance and scores on the APM when complete structural information is available is consistent with a number of previous studies (Putz-Osterloh, 1981; Putz-Osterloh & Lürer, 1981; Hörmann & Thomas, 1988). The strong relationship suggests that when complete structural information is available there is a large overlap between the underlying processes of the two tasks. As discussed in Chapter 4, Carpenter et al.'s (1990) analysis of the requirements of the APM suggests that the main sources of individual differences in performance are the ability to induce abstract relations and dynamically manage information in working memory. One could argue, that the high correlation between the APM and control performance under conditions where complete information is available suggests that in both instances abstract relations have to be induced and dynamically managed.

The lower correlations under partial and no information conditions, however, would suggest that the aforementioned processes are less prevalent. Considering that subjects in the no information condition did not perform better than random this suggests that their control actions were the result of random interventions. There should be no systematic differences in the success of random interventions, and therefore we would not expect a relationship between fluid intelligence and control performance. In comparison, under conditions of partial information, control interventions may be a combination of knowledge-based interventions and trial-and-error. The presence of random interventions may introduce unsystematic sources of variability that may lower the relationship to fluid intelligence. Overall, these results suggest that a lower relationship between fluid intelligence and control performance is to be expected when problem solvers do not have access to complete knowledge of the underlying structure of the system, and thus are forced to act randomly to some extent.

These results may explain why the relationship between fluid intelligence and control performance is so inconsistent in previous research. Typically, when problem solvers are required to control the system without an initial exploration phase, the relationship between fluid intelligence and control performance has been found to be

low (e.g. Dörner et al., 1983; Funke, 1983; Gediga et al., 1984; Putz-Osterloh, 1985; Reichert & Dörner, 1988; Joslyn & Hunt, 1998). In contrast, when problem solvers first explore the system in order to acquire structural knowledge and then control the system, the relationship between fluid intelligence and control performance is moderate (Beckmann, 1994; Funke, 1985; Misiak & Kluwe, 1986; Krörner, 2001; Wagener, 2001; Kröner & Leutner, 2002; Krörner et al., 2005; Bühner et al., 2008). This pattern of results is consistent with the results of the current study, which suggests that if problem solvers do not get a chance to acquire knowledge then performance will be the result of random interventions, and thus will not reflect differences in fluid intelligence. Alternatively, when problem solvers acquire knowledge, performance is knowledge-based, and thus reflects differences in fluid intelligence.

Against expectations, the relationship between fluid intelligence and control performance did not vary as a function of system complexity. This is particularly surprising considering that it was more difficult to control the systems that were designed to be more complex. This would suggest that the control of more complex systems requires the execution of a greater number of processes. One explanation for this result may be that the change in the number of processes to be executed across the levels of system complexity was too small to be detected by changes in the relationship to fluid intelligence.

Nevertheless, the results of this chapter make an important contribution to our understanding of system complexity. In the introduction it was argued that system complexity could be conceptualised as a function of the information processing demands imposed by system properties in the performance of a particular task. In support of this claim, it was found that the connectivity implied by the goal states, rather than the number of relations, was a consistent predictor of the difficulty of knowledge acquisition and system control. This is interpretable from an information-processing point of view, as increases in connectivity increase the amount of information that must be processed simultaneously by the problem solver. In contrast, an increase in the number of relations in a system, without an increase in connectivity, simply increases the number of operations that must be performed sequentially. The results of the present study are in line with those from other domains that have shown that the amount of information that must be processed in

parallel is directly related to task complexity (e.g. Halford et al., 1998a; Andrews & Halford, 2002; Birney & Halford, 2002; Halford et al., 2005; Beckmann, 2010).

These results qualify those reported by Kluge (2008) and Funke (1985; 1992) in relation to the effect of the number of relations on the difficulty of knowledge acquisition and system control. In these studies, manipulations of the number of relations in systems with the same number of variables resulted in concomitant increases in the connectivity implied by the goal states. The results of the current study suggest that the connectivity implied by the goal states, rather than the number of relations, may be responsible for the decrements in knowledge acquisition and control performance observed in these studies.

In contrast to the thinking of many researchers in this domain, the results suggest that a system complexity framework based on counting the number of certain system properties will be insufficient for predicting the difficulty of knowledge acquisition or control performance. Rather, the framework must take into account a) the task to be performed by the problem solver and b) how system properties influence the amount of information that must be processed in parallel in the performance of this task. One approach to this problem might be to decompose task performance into a series of sub-tasks, and then count the number of system properties that influence performance on each sub-task. The sub-task with the highest processing load could be taken as indicative of overall system complexity in relation to performing that particular task. Halford and colleagues (Halford et al., 1998a; Halford et al., 2005) have developed and successfully used a similar approach to estimate the complexity of reasoning tasks (i.e. Relational Complexity Theory). The results suggest that such an approach to system complexity might prove useful for the comparisons of different systems and for the prediction of knowledge acquisition and control performance.

The results also point to the need to construct systems at the appropriate level of complexity for studying processes such as reasoning and decision-making, and to assess problem solving competencies. If the system is too complex, performance is likely to be random, and consequently there will be floor effects on performance. Under such conditions it would be inappropriate to derive inferences about cognitive processes or problem solving competencies from the data. The results of the current

study suggest that systems become “too complex” when problem solvers are required to integrate three sources of information in order to make a decision about a goal state. However, strong conclusions cannot be drawn from the results of a single study. Further research is required to establish the limit on human performance in dynamic systems, and establish guidelines for the construction of systems at appropriate levels of complexity.

In conclusion, the results of this study indicate that successful problem solving in complex and dynamic environments requires structural knowledge. The likelihood that problem solvers will be able to acquire such knowledge through direct instruction and use it to effectively control systems is dependent on fluid intelligence and system complexity. In particular, it was found that it is more difficult to convert information into knowledge and control systems that have a high level of connectivity. The number of relations in the system did not have a consistent impact on knowledge acquisition or system control. This indicates that the amount of information to be processed in parallel is an important determinant of performance on these tasks. This is congruent with our conceptualisation of system complexity as a function of the information processing demands of the task to be performed engendered by the properties of the system. Overall, the success of system control can be seen as a function of the interaction between the limitations imposed by the system in terms of the amount of information that must be processed in parallel and the knowledge and abilities of the problem solver.

CHAPTER EIGHT

CONCLUSIONS AND IMPLICATIONS

8.1 Introduction

The overall goals of this project were to determine the conditions that are required to learn how to effectively control dynamic systems, and the psychological processes that separate successful from less successful problem solvers in the performance of this task. The main emphasis of this investigation was to clarify the role of structural knowledge in the control of dynamic systems, and to identify sources of individual differences in problem solvers capacity to acquire such knowledge and apply it in a goal-orientated application. A combined experimental and differential approach was adopted to address these goals, which consisted of the experimental manipulation of the task and structural characteristics of complex problems combined with the use of process indicators and external psychometric tests. This allowed us to identify a causal relationship between structural knowledge and control performance across systems of differing complexities, and to account for individual differences in the strength of this relationship. This chapter will summarise the main findings from this project, and discuss the implications for the use of complex problems as a tool for the assessment of problem solving competencies and training. We will also identify some key issues in this domain that are in need of further research.

8.2 How does structural knowledge influence the control of dynamic systems?

At the outset of the project, a comprehensive review of the literature was conducted to integrate the findings from the three major research paradigms that have investigated how people learn to control dynamic systems (Dynamic Decision Making, Implicit Learning and Linear Structural Equations). This review revealed that the role of structural knowledge in the control of dynamic systems was unclear. While some studies had found that the amount of structural knowledge subjects acquire is positively associated with the quality of their control performance (e.g. Funke & Müller, 1988; Beckmann & Guthke, 1995; Vollmeyer et al., 1996; Kröner et al., 2005; Burns & Vollmeyer, 2002; Osman, 2008a), others had found evidence that subjects learn to control dynamic systems seemingly in the absence of structural

knowledge (e.g. Broadbent, 1977; Broadbent et al., 1986; Berry & Broadbent, 1984; 1988). This inconsistency had not been adequately addressed in the literature, as each research approach had largely operated independently of the others.

In order to reconcile this inconsistent pattern of results, a framework was constructed to describe complex problems in terms of their task, surface and structural characteristics. It was argued that each approach entails the application of a specific set of these characteristics to the design and use of complex problems, which in turn influences the type of knowledge that will be acquired and utilised. Based on this synthesis of the results of previous research, it was concluded that a positive association between structural knowledge and control performance is consistently observed when system states can be generated that are informative to the underlying structure of the system, the problem is presented in a novel context and problem solvers are given the opportunity to acquire knowledge *prior* to the instruction to control the system to achieve specific goals.

However, the dependency of efficient system control on the acquisition of structural knowledge had not been clearly established. The results of a number of studies had found that subjects provided with structural information do not show better control performance than subjects who perform without such information, or who are required to acquire structural knowledge through an unguided exploration of the system variables (Putz-Osterloh & Lürer, 1981; Putz-Osterloh, 1993; Preußler, 1996). This was particularly surprising considering that many studies had shown that the majority of subjects do not acquire complete knowledge of the underlying structure of systems through an unguided exploration of the system variables (Funke & Müller, 1988, Müller 1993; Beckmann, 1994; Beckmann & Guthke, 1995; Vollmeyer, et al., 1996; Burns & Vollmeyer 2002; Kröner, 2001; Schoppek, 2002; Kröner, et al., 2005; Kluge, 2008; Osman, 2008a). These findings had led some to argue that systems can be controlled successfully without the aid of specific knowledge, and that a simple strategy of trial-and-error may be as efficient in achieving goal states as the application of the rules underlying a system (Putz-Osterloh, 1993).

Our findings do not support this claim, but rather converge on the relevance of the acquisition of structural knowledge for successful system control. The main

findings were: 1) the amount of structural knowledge that problem solvers acquire has a systematic impact on the quality of their control performance, with increasing levels of structural knowledge resulting in better system control; 2) we found no evidence that problem solvers can efficiently control the outcomes of dynamic systems in the absence of structural knowledge. When structural knowledge was not acquired control performance was no better than random. When control performance was not better than random, individual differences in control performance were also rather inconsistent. This suggests that when problem solvers do not acquire structural knowledge their control performance is the result of trial-and-error, which is not a sufficient strategy for learning to control the outcomes of dynamic systems; and 3) these results were shown to generalise to systems at different levels of complexity. Overall, these results demonstrate that in order to successfully control the outcomes of dynamic systems, problem solvers must develop an accurate model of the underlying structure of the system.

These findings are particularly significant considering that in many studies that have investigated how people learn to control dynamic systems, subjects have not been given the opportunity to build up their knowledge base prior to the instruction to control the system (e.g. *The Dynamic Decision Making Approach*; Dörner, 1975, Brehmer, 1987; Sterman, 1994; Busemeyer, 2002; Dörner, 1980; Dörner & Wearing, 1995; Kerstholt & Raaijmakers, 1997). The conclusion that has been drawn from the results of these studies is that people perform poorly because they must process multiple sources of information in parallel, while co-ordinating between different goals and sub-tasks and acting within time constraints. It has been argued that these demands overwhelm the resources of the information processing system, which results in sub-optimal performance. However, the results of the current project suggest that problem solvers simply do not have the chance to acquire sufficient domain specific-knowledge to make effective decisions under such conditions.

Previously, the benefit of structural information had only been observed when subjects first practiced applying structural information to reach specific goals (Putz-Osterloh, 1993; Preußler, 1998). These results were interpreted as evidence that the application of structural information was a skill that had to be practiced in addition to learning the structure of the system. It was argued that structural knowledge alone is

insufficient to control the outcomes of dynamic systems (Schoppek, 2004; Preußler, 1996; Preußler, 1998).

The results of the current project lead to a significant revision of this view as well. We found that if problem solvers can convert information into structural knowledge, then it has a positive impact on the quality of control performance, without a period of goal-orientated practice or prior exposure to the system. Indeed, once structural knowledge was acquired, the quality of subjects' control performance was rather stable over a number of intervention trials and across different goal states. This indicates that practice does not have any significant impact on the effectiveness with which subjects are able to apply structural knowledge (at least within a limited time frame).

The differing pattern of results observed in comparison to previous research may be attributable to two factors. Firstly, the instructions used in the current project may have been more effective in promoting structural knowledge as they were designed in accordance with the principles of CLT. Previous studies employed either a diagram of the underlying structure of the system presented on paper (Putz-Osterloh & Lürer, 1981; Putz-Osterloh, 1993) or standardised examples as to how each input effected each output with text-based explanations (Preußler, 1996). In the current project, standardised examples with a concurrent audio explanation were utilised to try to minimise the number of cognitive activities that subjects would have to undertake to translate the information provided into knowledge about the system. Additionally, a diagram that depicted this information remained on screen during the control tasks as an external memory aid.

It should be noted, however, that one potential confound is that the method used to instruct structural knowledge also demonstrated a useful strategy for acquiring that knowledge (i.e. the VOTAT strategy). It may be argued that if the structural information was not understood, subjects may have used the VOTAT strategy to acquire additional knowledge during the control phase. Improvements in performance may then be due to the instruction of the VOTAT strategy, rather than the instruction of structural information. However, this seems unlikely considering that a) in order to successfully control the system subjects must vary multiple inputs at the same time which is incompatible with the use of the VOTAT strategy and b)

subjects in Study 1 showed significant improvements in control performance directly after they received structural information. It therefore seems unlikely that they were trying to acquire knowledge during the control phase. Similarly, in Studies 2 and 3, control performance in the complete information conditions was consistently superior to control performance in the no information conditions. This would not be the case if subjects in the complete information condition were attempting to acquire new knowledge while they controlled the system. Nevertheless, it must be acknowledged that the instruction of structural knowledge also demonstrated the best strategy for the acquisition of new knowledge.

To return to the differing pattern of results observed in comparison to previous research, secondly, the results of the current project suggest that the complexity of the system, as indicated by the connectivity implied by the goal states, to some extent determines whether structural information will be understood and utilised by problem solvers. In the study reported in Chapter 7, we constructed a complex problem that had an underlying structure that had the same level of connectivity implied by the goal states to that used by Putz-Osterloh (1993) and Preußler (1996). It was found that subjects interacting with this system failed to convert information into knowledge and subsequently did not perform better than random. This suggests that the benefit of practice proposed in previous studies may be attributable to the additional time needed to understand the information provided when the system is highly complex, rather than an effect of practice per se. Further research is required to determine the type of instructions that are required to promote structural knowledge when systems have highly complex structures.

8.3 What differentiates successful from less successful problem solvers in the control of dynamic systems?

Clearly, the results of this project establish the primary importance of the acquisition of structural knowledge for effective system control. Overall then, the quality of problem solvers' control performance can be seen as a product of their individual capacity to acquire knowledge and then use it to control the system. Few previous studies had considered the acquisition of structural knowledge and system control as capacities that might be determined by divergent sources of individual differences.

8.3.1 Individual differences in the acquisition of structural knowledge

The results show that the acquisition of structural knowledge through an unguided exploration of the system variables is largely determined by the identification and use of the VOTAT strategy and fluid intelligence (indicated by scores on the APM). Importantly, these variables were found to be relatively unrelated to one another. On the one hand, this suggests that training on how to design effective experiments to test hypotheses, or homogenising the system states that subjects can observe during the exploration phase, might significantly reduce the difficulties that problem solvers at all levels of ability experience when they try to acquire structural knowledge. On the other hand, more intelligent problem solvers will still be at an advantage as they are able to more effectively induce the rules from the data and test hypotheses.

In situations where problem solvers receive direct instruction with regard to the underlying structure of the system, the main requirement is that they effectively translate the information provided into knowledge. The results of the study reported in Chapter 7 showed that many subjects who received complete structural information had difficulty converting this information into knowledge. This was particularly surprising considering that the test of structural knowledge essentially required subjects to replicate system states that they had viewed in the instructional phase. It was also found that performance on the structural knowledge test was largely unrelated to fluid intelligence. Again, this suggests that training interventions designed to improve structural knowledge might have significant benefits to problem solvers at all levels of ability.

Overall these results indicate that more successful problem solvers are better able to: 1) design experiments that provide clear tests of hypotheses; 2) convert structural information into knowledge about the system and 3) have a higher level of fluid intelligence. Individual differences in structural knowledge are determined by a combination of acquired skills and cognitive ability.

8.3.2 Individual differences in control performance

The results of the current project indicate that individual differences in control performance can primarily be accounted for by the amount of knowledge that

problem solvers acquire. However, once the amount of structural knowledge is taken into account, control performance is strongly related to fluid intelligence. Importantly, the strength of this relationship did not depend on the complexity of the system or whether structural knowledge was acquired through direct instruction or an unguided exploration of the system variables. Indeed, it did not diminish even when subjects had practice on the control task prior to receiving structural information. Overall, this indicates that more successful problem solvers possess more structural knowledge and have a higher level of fluid intelligence.

However, it must be acknowledged that it cannot be ruled out that the consistent correlation between fluid intelligence and control performance is not due to some other common factor. A number of studies have found that non-cognitive factors such as motivation, metacognition and emotions have a significant impact on the cognitive processes involved in the acquisition of structural knowledge and system control (e.g. Vollmeyer, Rollett & Rheinberg, 1997, 1998; Vollmeyer & Rheinberg, 1999, 2000; Spering, Wagener & Funke, 2005; Barth & Funke, 2009). It is therefore possible that one of these unmeasured sources of individual differences could account for some of the variance in control performance currently attributed to fluid intelligence.

The results of the current thesis do, however, clarify the conditions that are required to observe a strong relationship between fluid intelligence and control performance: The strength of the relationship between fluid intelligence and control performance appears to depend, to a certain degree, on whether or not structural knowledge has been acquired. When complete structural knowledge was acquired, the quality of subjects' control performance was strongly related to fluid intelligence, which suggests that under these conditions the demands of the two tasks are similar. The main processes that distinguish between individuals on tests of fluid intelligence are the ability to induce abstract relations and working memory capacity (Carpenter et al., 1990). Since the requirements of system control precludes the acquisition of further structural knowledge (i.e. the abstraction of relations or rules), it can be inferred then, that the main source of individual differences in control performance is working memory capacity. These results support our task analysis that was presented in Chapter 4, and the results of previous research that has shown that control performance and working memory capacity are moderately correlated (Wittmann &

Süß, 1999; Wittmann & Hatrup, 2004; Bühner et al., 2008). These results extend upon these previous findings, as they show that the relationship holds when the amount of structural knowledge acquired by problem solvers has been taken into account. That is, that the effect of fluid intelligence on control performance is independent of the effect on fluid intelligence on the acquisition of structural knowledge.

When no structural knowledge was acquired, the relationship between fluid intelligence and control performance was somewhat weaker, although those who were more intelligent were still at a significant advantage compared to those who were less intelligent. However, as the reliability of control scores was low and performance was not better than random under these conditions, it seems likely that the differences between people may be the result of unsystematic sources of variance. Thus, it seems inappropriate to draw inferences about the processes that separate successful from less successful problem solvers when no structural knowledge has been acquired.

When only partial structural knowledge was acquired, the relationship between fluid intelligence and control performance was similar to that observed when no knowledge was acquired. This may be because for some subjects, the partial information instruction helped them gain insight into the structure, and for others it may have hindered learning. Thus, some subjects may have been using a knowledge-based strategy to control the system and others may have been acting through trial-and-error. This would weaken the relationship between fluid intelligence and control performance. This issue has significant implications for the assessment of problem solving competencies which will be discussed in more detail in the following section.

8.4 Implications for the use of complex problems as a tool for the assessment of problem solving competencies

Currently, no guidelines are available which describe how complex problems should be constructed and administered for the purposes of assessment and selection. Consequently, the task, structural and surface characteristics of the complex problems that are used in assessment settings are rather heterogeneous. Clearly, the results of the current project and previous research indicate that this has significant consequences in terms of what these so-called tests of complex problem solving are

actually measuring. However, it is usually claimed that complex problems measure the capacities to acquire and utilise structural knowledge (Hornke & Kersting, 2005; Kluge, 2008; Greiff & Funke, 2008; 2009). Therefore, in this section we will suggest some general guidelines for the construction and administration of complex problems that will ensure that performance does indeed reflect these capacities, based on the findings of the current project.

The results of the current project indicate that the acquisition of structural knowledge and system control impose different cognitive demands on the problem solver. However, system control is dependent on the acquisition of structural knowledge; structural knowledge cannot be acquired while problem solvers try to control the outcomes of a system. Thus, in concordance with Funke (1984; 1992; 2001), we argue that these tasks need to be separated experimentally in order to assess performance on each of these tasks independently of the other. This means that, firstly, test-takers should be given an opportunity to acquire structural knowledge through an unguided exploration of the system variables before they are required to control the outcomes of the system. This allows for an appropriate assessment of the problem solvers' capacity to design effective experiments, test hypotheses and induce rules. In addition, we suggest that before test-takers are required to control the outcomes they should be directly instructed as to the underlying structure of the system. This should allow for valid comparisons of control performance to be made between test-takers who may otherwise have acquired different levels of structural knowledge. If complex problems are administered in this way, the capacity to acquire structural knowledge can be assessed independently of the capacity to utilise knowledge.

There is some debate over whether it is appropriate to administer a single complex problem in order to assess problem-solving competencies, as is the typical practice in most assessment settings. The rationale for this practice is that complex problems are analogous to a work sample, in which the test-taker is required to deal with a complex and dynamic system (U. Funke, 1998; Kluge, 2008). Greiff and Funke (2008; 2009) argue that this practice is psychometrically questionable, and that multiple complex problems of differing complexity need to be administered in order to make valid inferences with regard to problem solving competencies. They propose that problems of differing complexity can be constructed by varying their

underlying structures. However, the results of Study 3 show that even small changes in the underlying structure of a system can result in ceiling and floor effects on performance. This suggests that the administration of multiple complex problems with different underlying structures may not provide any more information about test-takers than the administration of a single complex problem. The important point is that the complex problem must be at an *appropriate* level of complexity to differentiate between test takers, and that extensive pilot testing is required to ensure that performance on any individual complex problem does indeed reflect the capacities to acquire and utilise structural knowledge.

The current project clearly demonstrates the necessity of comparing test-takers' control performance to control performance that has been produced through random interventions, in order to establish a clear standard for successful performance. While this is not a new idea (see Beckmann, 1994), it has rarely been employed (with the notable exceptions of Beckmann, 1994; Kluge, 2008). The results of Study 3 clearly demonstrate that if the controllability of the system is not taken into account then "good" performance may be attributable to "good" decision making on the part of the test-taker or, alternatively, structural characteristics that result in "good" control performance no matter what interventions are undertaken.

The importance of structural knowledge for system control suggests that complex problems should be presented in a novel context. Currently, most of the complex problems that are used for personnel selection and assessment purposes are presented with familiar cover stories and labels for the system variables (U. Funke, 1998; J. Funke, 1998; Greiff & Funke, 2009). While the situations are usually selected so that test-takers do not have direct experience with them in their everyday life, it is most likely that will be familiar with the variables in question, which may allow them to form assumptions about the underlying structure of the system. These assumptions may not be consistent across test-takers, and as such, the familiarity of the context represents a source of uncontrolled variance that is likely have a significant impact on the acquisition of structural knowledge and subsequent control performance. The presentation of problems in entirely novel contexts (i.e. with abstract labels for the system variables) will ensure that it is possible to make valid comparisons between test-takers.

All of the issues discussed so far, however, have only been concerned with how best to measure the capacities to acquire structural knowledge and use it to control the outcomes of dynamic systems. Clearly, a critical consideration in personnel selection and educational assessment must also be whether the scores that are produced are in fact meaningful predictors of job or academic performance. In addition, it is crucial to show that they have incremental validity in the prediction of job and academic performance over and above traditional measures of intelligence, which have been shown to be one of the best predictors of these criteria (Neisser et al., 1996; Hunter & Schmidt, 1997). Although the current project did not address the first issue, with regard to the second, the results indicate that the acquisition of structural knowledge has some demands that are divergent from those that are captured in traditional tests of intelligence, such as the APM. In particular, problem solvers must design experiments that will yield data that can be used to test hypotheses in order to acquire structural knowledge. This skill is likely to be particularly relevant in jobs that require research and development, and educational settings that foster a constructivist perspective on learning. This suggests that the capacity to acquire structural knowledge through an unguided exploration of the system variables may show incremental validity in the prediction of success in certain types of jobs and academic settings over and above traditional measures of intelligence.

8.5 Directions for future research

It remains an important question as to whether structural information is useful when problem solvers already have a partial mental model of the system that is based on their real world experiences. The results of Study 1 suggest that problem solvers can incorporate new information into an existing mental model. However, this knowledge had only been acquired moments before, in the same context. If that knowledge was acquired over a more extensive period of time, in multiple contexts in the real world, then it may be more difficult to revise. To the authors' knowledge, there has not been any research that has addressed this question, although a number of studies have investigated how familiarity and domain expertise influences the acquisition of knowledge through an exploration of the system variables, and the results are rather inconsistent (e.g. Reither, 1981; Beckmann, 1994; Beckmann & Guthke, 1995; Burns & Vollmeyer, 2002; Rolo & Diaz-Caberera, 2005; Lazonder et

al., 2008; 2009). In order to address this issue, future research could examine the effect of structural information on system control with complex problems that are embedded in a familiar context.

In other domains, it has been shown that practice influences the organisation of knowledge and strategies for performance in complex tasks. Findings show that increasing amounts of practice are associated with better performance (Holyoak, 1991; Ericsson, 2003), and that the influence of intelligence declines as performance becomes more automatised (Ackerman, 1988; 1990; 1992; Kanfer & Ackerman, 1989). In the current project, we did not find that practice had a significant impact on the quality of system control, or the relationship between control performance and fluid intelligence. However, our investigations were rather time-limited. Therefore, the need is apparent to determine whether and how control performance changes with practice over longer periods of time, and whether the sources of individual differences in control performance change as a result of practice.

The current project primarily focussed on the role of cognitive factors in complex problem solving. Other research has shown that non-cognitive factors, such as motivation and emotions, also have a significant impact on performance in complex and dynamic environments (Vollmeyer et al., 1997, 1998; Vollmeyer & Rheinberg, 1999, 2000; Spering et al., 2005; Barth & Funke, 2010). Moving forward, one of the biggest challenges facing researchers in this field is to determine how cognitive and non-cognitive factors might interact and influence performance. Until then, our model of performance in complex and dynamic environments remains incomplete.

8.6 Conclusion

The conditions that are required to learn how to effectively control the outcomes of dynamic systems are now clear. The results of the current project demonstrate that the successful control of complex systems requires domain-specific knowledge. That is, knowledge of the relationships between the variables in the system. Without such knowledge, control behavior is rather unsystematic, and as a result, performance is no better than what might be achieved through random interventions on the system variables. The acquisition of structural knowledge can be achieved either through an unguided exploration of the system variables, or direct

instruction. Direct instruction is preferable if the aim is to provide the problem solver with maximal amounts of structural knowledge, however, it may be ineffective if the underlying structure of the system is highly complex. Individual differences in the capacity to design effective experiments in order to test hypotheses plays a crucial role in the acquisition of knowledge, as does fluid intelligence, which also influences the subsequent application of the acquired knowledge. The success of control performance can then be seen as a function of the amount of knowledge that is acquired, combined with the intelligence of the problem solver.

REFERENCES

- Ackerman, P.L., Beier, M.E. & Boyle, M.O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, 131, 30–60.
- Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, 117, 288-318.
- Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, 117, 288-318.
- Ackerman, P.L. (1996) A Theory of Adult Intellectual Development: Process, Personality, Interests, and Knowledge. *Intelligence*, 22, 227-257
- Ackerman, P.L. & Cianciolo, A.T. (2002). Ability and task constraint determinants of complex task performance. *Journal of Experimental Psychology: Applied*, 8, 194-208.
- Ackermann, P.L. (1990). A correlational analysis of skill specificity: Learning, abilities, and individual differences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 883-901.
- Ackermann, P.L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 598-614.
- Ackermann, P.L., Kanfer, R. & Goff, M. (1995). Cognitive and Noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*, 1, 270-304.
- Amthauer, R. (1973). *IST-70. Intelligenz-Struktur-Test [Test of the structure of intelligence]*. Göttingen: Hogrefe.
- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, Massachusetts: Harvard University Press.
- Anderson, J.R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J.R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Andrews, G. & Halford, G.S. (1998). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. *Cognitive*

- Development*, 13, 470–513.
- Andrews, G. & Halford, G.S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, 45, 153–219.
- Anzai, Y. & Simon, H.A. (1979). The theory of learning by doing. *Psychological Review*, 86, 124–140.
- Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, 49A, 5-28.
- Baddeley, A.D. (1986). Working memory. Oxford: Clarendon Press
- Barth, C. & Funke, J. (2010). Negative affective environments improve complex solving performance. *Cognition and Emotion*, 24(7), 1259-1268
- Beckmann, J.F. (1994). *Lernen und komplexes Problemlösen: Ein Beitrag zur Konstruktvalidierung von Lerntests* [Learning and complex problem solving: A contribution to the construct validation of tests of learning potential]. Bonn, Germany: Holos.
- Beckmann, J.F. (2010). Taming a beast of burden – On some issues with the conceptualisation and operationalisation of cognitive load. *Learning and Instruction*, 20, 250-264.
- Beckmann, J.F. & Guthke, J. (1995). Complex problem solving, intelligence and learning ability. In P.A. Frensch & J. Funke (Eds.) *Complex Problem Solving: The European Perspective* (pp. 177 – 200). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berry, D. (1984). *Implicit and explicit knowledge in the control of complex systems*. Unpublished doctoral thesis, University of Oxford, United Kingdom.
- Berry, D. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology*, 43, 881–906.
- Berry, D. & Broadbent, D. (1995). Implicit Learning in the Control of Complex Systems. In P.A. Frensch & J. Funke (Eds.) *Complex Problem Solving: The European Perspective* (pp. 131 – 150). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit–explicit distinction. *British Journal of Psychology*, 79, 251–272.
- Berry, D.C. & Broadbent, D.E. (1987) Explanation and verbalization in a computer-assisted search task. *The Quarterly Journal of Experimental Psychology*

Section A, 39, 585-609.

- Berry, D.C. & Broadbent, D.E. (1984) On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology Section A*, 36, 209-231
- Bhaskar, R. & Simon, H.A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science*, 1, 193-215.
- Birney, D.P. & Halford, G.S. (2002). Cognitive complexity of suppositional reasoning: An application of the relational complexity metric to the knight-knave task. *Thinking and Reasoning*, 8, 109-134.
- Blech, C., & Funke, J. (2005). Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology. Heidelberg University: Unpublished manuscript. Retrieved 15th July 2006:
<http://www.psychologie.uni-heidelberg.de/ae/allg/forschun/dynamis/approach.html>.
- Blech, C. & Funke, J. (2006). Zur Reaktivität von Kausaldiagramm-Analysen beim komplexen Problemlösen [On the reactivity of complex problem solving to causal diagrams]. *Zeitschrift für Psychologie*, 117, 185-195.
- Bonner, S.E. (1994). A model of the effects of audit task complexity. *Accounting, Organizations and Society*, 19, 213-234.
- Boring, E.G. (1923). Intelligence as the tests test it. *New Republic*, 36, 35-37.
- Brattico, P. (2008). Shallow reductionism and the problem of complexity in psychology. *Theory Psychology*, 18, 483-504.
- Brehmer, B. (1987). System design and the psychology of complex systems. In J. Rasmussen & P. Zuhde (Eds.), *Empirical foundations of information and systems science* (pp. 21-32). New York: Plenum Press.
- Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. Hogarth (Ed.), *Insights in decision making* (pp. 262-279). Chicago: University of Chicago Press.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81, 211-241.
- Brehmer, B. (1995). Feedback delays in dynamic decision making. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European Perspective* (pp.

- 103-130). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brehmer, B. (2005). Micro-worlds and the circular relation between people and their environment. *Theoretical Issues in Ergonomics Science*, 6, 73 - 93.
- Brehmer, B. & Dörner, D. (1993) Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9, 171-184
- Brehmer, B. & Allard, R. (1991). Dynamic decision making: The effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision making: Cognitive models of cooperative work*. Chichester: Wiley.
- Broadbent, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.) *Studies in long-term memory*. UK: Wiley.
- Broadbent, D.E. (1977). Levels, hierarchies and the locus of control. *Quarterly Journal of Experimental Psychology*, 32, 109–118.
- 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.
- Bruce, A.C. & Johnson, J.E.V. (1996). Decision-making under risk: effect of complexity on performance. *Psychological Reports*, 79, 67-76.
- Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In Frensch, P.A. and Funke, J. (Eds.). *Complex problem solving: the European perspective* (pp.27 – 63). Lawrence Erlbaum Associates, Hillsdale, NJ.
- Buchner, A. & Funke, J. (1993). Finite state automata: Dynamic task environments in problem solving research. *Quarterly Journal of Experimental Psychology*, 46A, 83–118.
- Buchner, A., Funke, J. & Berry, D. (1995). Negative correlations between control performance and verbalizable knowledge: Indicators for implicit learning in process control tasks. *Quarterly Journal of Experimental Psychology*, 48, 166–187.
- Bühner, M., Kröner, S. & Ziegler, M. (2008). Working memory, visual-spatial-intelligence and their relationship to problem solving. *Intelligence*, 36, 672-680.

- Bühner, M., Krumm, S. & Pick, M. (2005). Reasoning is equal to working memory not equal to attention. *Intelligence*, 33, 251–272.
- Burns, B. & Vollmeyer, R. (2002) Goal specificity effects on hypothesis testing in problem solving. *The Quarterly Journal of Experimental Psychology Section A*, 55, 241-261.
- Burns, B.D. & Vollmeyer, R. (2000). Problem solving: phenomena in search for a thesis. In *Proceedings of the Cognitive Science Society Meeting* (pp. 627–632), 13–15 August 2000, Pittsburgh, USA.
- Busemeyer, J. R. (2002). Dynamic decision making. In N.J. Smelser and P.B. Bates (Eds.) *International Encyclopaedia of the Social and Behavioural Sciences: Methodology, Mathematics and Computer Science*. (pp. 3903–3908). Oxford: Elsevier.
- Byström, K. & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management*, 31, 191-213.
- Campbell, D.J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, 13, 40-52.
- Cañas, J.J. & Waern, Y. (2005). Cognitive research with microworlds. *Theoretical Issues in Ergonomics Science*, 6, 1–3.
- Carpenter, P.A., Just, M.A. & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the raven progressive matrices test. *Psychological Review*, 97, 404–431.
- Carroll, J.B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Cattell, R.B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.
- Chase W.G. & Simon H.A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Chi, M.T.H., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.) *Advances in the Psychology of Human Intelligence*, Hillsdale, NJ Erlbaum, Vol. 1, pp. 1-75.
- Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg

- (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, NJ: Erlbaum.
- Chomsky, N. (1957). *Syntactic structures*. The Hague, the Netherlands: Mouton.
- Christoffersen, K., Hunter, C.N. & Vicente, K.J. (1996). A longitudinal study of the effects of ecological interface design on skill acquisition. *Human Factors*, 38, 523– 541.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297 – 334.
- Cronbach, L.J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30: 116–127.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19, 450–466.
- de Freitas, S. & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education*, 46, 249-264.
- de Groot, A. (1978). *Thought and Choice in Chess*. The Hague: Mouton. (Original work published 1946).
- de Jong T. & van Joolingen W.R. (1997). An extended dual search space model of scientific discovery learning. *Instructional Science*, 25, 307-346.
- de Jong T. & van Joolingen W.R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179–202.
- de Jong, T. (2005). The guided discovery principle in multimedia learning. In R.E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 215 – 229). Cambridge, UK: Cambridge University Press.
- de Jong, T. (2006). Scaffolds for Computer Simulating Based Scientific Discovery Learning. In Jan Elen & Richard E. Clark (Eds.), *Handling Complexity in Learning Environments: Theory and Research (Advances in Learning and Instruction)* (pp. 107-128). Oxford, UK; Boston: Elsevier
- DeShon, R. P., Chan, D. & Weissbein, D. A. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional

- performance determinants. *Intelligence*, 21, 135-155
- Diehl, E. & Serman, J. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behaviour and Human Decision Processes*, 62, 198-215.
- Dienes, Z. & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 848-862.
- DiFonzo, N., Hantula, D. A. & Bordia, P. (1998). Microworlds for experimental research: Having your (control and collection) cake, and realism too. *Behaviour Research Methods, Instruments, & Computers*, 30, 278-286.
- Dörner, D. (1975). Wie Menschen eine Welt verbessern wollten [How people wanted to improve the world]. *Bild der Wissenschaft*, 12, 48-53.
- Dörner, D. (1980). On the difficulty people have in dealing with complexity. *Simulation & Games*, 11, 87-106.
- Dörner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.) *New technology and human error* (pp. 97-109). Chichester, UK: Wiley
- Dörner, D. & Scholkopf, J. (1991). Controlling complex systems; or, expertise as “grandmother’s know-how”. In K.A. Ericsson & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits*. New York, NY: Cambridge University Press.
- Dörner, D. & Wearing, A.J. (1995). Complex problem solving: Toward a (computer simulated) theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65-99). Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Inc.
- Dörner, D., Kreuzig, H.W., Reither, F. & Stäudel, T. (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität* (Lohhausen. On dealing with uncertainty and complexity), Hans Huber, Bern, Switzerland.
- Dörner, D., Stäudel T. & Strohschneider, S. (1986) Moro-Programmdokumentation. *Memorandum No 23 Lehrstuhl Psychologie II*, Bamberg
- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17, 397-434
- Duncker, K. (1945). On Problem Solving. *Psychological Monographs*, 58, American

Psychological Association.

- Earley, P. C. (1985). Influence of information, choice and task complexity upon goal acceptance, performance, and personal goals. *Journal of Applied Psychology*, 481-491.
- Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, 4, 59-73.
- Elg, F. (2005). Leveraging intelligence for high performance in complex dynamic systems requires balanced goals. *Theoretical Issues in Ergonomics Science*, 6, 63-72.
- Embretson, S.E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3, 300-396.
- Ericsson, A.K. & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49, 725-747.
- Ericsson, A.K., Prietula, M.J. & Cokley, E.T. (2007). The making of an expert. *Harvard Business Review*, retrieved September 2010, from: <http://hbr.org/2007/07/the-making-of-an-expert/ar/1>.
- Ericsson, K. A. (2003). The development of elite performance and deliberate practice: An update from the perspective of the expert-performance approach. In J. Starkes and K. A. Ericsson (Eds.), *Expert performance in sport: Recent advances in research on sport expertise* (pp. 49-81). Champaign, IL: Human Kinetics
- Fisher, D. L. (1984). Central capacity limits in consistent mapping, visual search tasks: Four channels or more? *Cognitive Psychology* 16, 449-84
- Frensch, P. & Funke, J. (1995). Definitions, Traditions, and a General Framework for Understanding Complex Problem Solving. In P. Frensch and J. Funke (Eds.) *Complex Problem Solving: The European Perspective* (pp. 3-24), UK: Lawrence Erlbaum Associates.
- Funke, J. (1991). Solving complex problems: exploration and control of complex systems. In R.J. Sternberg and P.A. Frensch (Eds.) *Complex problem solving: principles and mechanisms* (pp 185-222), Lawrence Erlbaum, Hillsdale.
- Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Praediktator? (Some remarks on problems of problem solving research, or: is test intelligence nevertheless a predictor?).

- Diagnostica*, 29, 283–302.
- Funke, J. (1984). Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen [Assessment of West German problem solving research]. *Sprache & Kognition*, 3, 159–172.
- Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle [Control of dynamic systems by building up and using subjective causal models]. *Zeitschrift für Psychologie*, 193, 435-457
- Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R.J. Sternberg & P.A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185–222). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Funke, J. (1992). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *German Journal of Psychology*, 16, 24-43.
- Funke, J. (1993). Microworlds based on linear equation systems: A new approach to complex problem solving and experimental results. In G. Strube & K.F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 313-330). Amsterdam: Elsevier Science Publishers.
- Funke, J. (1998). Computer-based testing and training with scenarios from complex problem-solving research: Advantages and disadvantages. *International Journal of Selection and Assessment*, 6, 90-96.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgment. *Thinking and Reasoning*, 7, 69–89.
- Funke, J. & Frensch, P.A. (2007). Complex problem solving: The European perspective – 10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25-47). New York/London: Lawrence Erlbaum Associates.
- Funke, J. & Müller, H. (1988). Eingreifen und Prognostizieren als Determinanten von Systemidentifikation und Systemsteuerung [Intervention and prediction as determinants of system identification and system control]. *Sprache & Kognition*, 7, 176-186.
- Gardner, P. H., & Berry, D. C. (1995). The effect of different forms of advice on the control of a simulated complex system. *Applied Cognitive Psychology*, 9 (Special Issue), S55-S79.

- Gary, M.R. & D.S. Johnson (1979). *Computers and Intractability—A Guide to the Theory of NP-completeness*. New York: W. H. Freeman.
- Gary, M.S. & Wood, R.E. (2007). Testing the effects of a system dynamics decision aid on mental model accuracy and performance on dynamic decision making tasks. *Conference Proceedings: The 2007 International Conference of the System Dynamics Society and 50th Anniversary Celebration, July 29 – August 2, 2007, Boston, Massachusetts, USA*. Retrieved June 2008: <http://www.systemdynamics.org/conferences/2007/proceed/index.htm>
- Gebauer, G.F. & Mackintosh, N.J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33, 34-54.
- Geddes, B.W. & Stevenson, R. J. (1997). Explicit learning of a dynamic system with a nonsalient pattern. *The Quarterly Journal of Experimental Psychology*, 50A, 742-765.
- Gediga, G., Schöttke, H. & Tücke, M. (1983). Problemlösen in einer komplexen Situation [Problem solving in a complex situation]. *Archiv für Psychologie*, 135, 325-339.
- Gibson, F.P., Fichman, M. & Plaut, D.C. (1997). Learning dynamic decision tasks: computational model and empirical evidence. *Organizational Behaviour & Human Decision Processes*, 71, 1–35.
- Gold, S. (2003). The design of a business simulation using a systems-dynamics-based approach. *Developments in Business Simulation and Experiential Learning*, 30. (Reprinted in *The Bernie Keys Library*, 2nd ed., by H. M. Cannon, Ed.).
- Goldman, S.R. (2009). Explorations of relationships among learners, tasks and learning. *Learning and Instruction*, 19, 451-454.
- Goldstein, W. M. & Weber, E. U. (1995). Content and discontent: indications and implications of domain specificity in preferential decision making. In J. Busmeyer, D. L. Medin, & R. Hastie (Eds.), *Decision making from a cognitive perspective* (Vol. 32, pp. 83–126). San Diego: Academic Press.
- Goldstone, R.L. & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology*, 46, 414-466.
- Gonzalez, C. (1997). Improving tools to create learning laboratories for complex problem solving. *Computers in Human Behaviour*, 13, 559-569.
- Gonzalez, C., Lerch, J. F. & Lebiere, C. (2003). Instance-based learning in dynamic

- decision making. *Cognitive Science*, 27(4), 591 – 635.
- Gonzalez, C., Vanyukov, P. & Martin, M. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behaviour*, 21, 273-286.
- Gonzalez, C., Thomas, R.P. & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence*, 33, 169-186.
- Goode, N. & Beckmann, J.F. (2010). You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence*, 38, 345-352.
- Goode, N. & Beckmann, J.F. (under revision). The Benefit of Being Naïve and Knowing it: The Impact of Semantic Embedment on Knowledge Acquisition in Complex Problem Solving Tasks. *Computers and Human Behaviour*.
- Goosen, K., Jensen, R. & Wells, R. (2001). Purpose and learning benefits of simulations: A design and development perspective. *Simulation & Gaming: An Interdisciplinary Journal*, 32, 21-39.
- Gray, W. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly*, 2, 205-227.
- Greiff, S. & Funke, J. (2008). What makes a problem complex? Factors determining difficulty in dynamic situations and implications for diagnosing complex problem solving competence. In J. Zumbach, N. Schwartz, T. Seufert & L. Kester (Eds.), *Beyond knowledge: The legacy of competence. Meaningful computer-based learning environments* (pp. 199-200). Heidelberg: Springer.
- Greiff, S. & Funke, J. (2009). Measuring complex problem solving - The MicroDYN approach. In F. Scheuermann (Ed.), *The transition to computer-based assessment - Lessons learned from large-scale surveys and implications for testing* (pp.157-163). Luxembourg: Office for Official Publications of the European Communities.
- Guastello, S.J. (2004). Coping with complexity and uncertainty. In Lowell Douglas Kiel (Ed.) *Knowledge Management, Organizational Intelligence and Learning, and Complexity*, in Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford,UK.
- Hackman, R.J. (1969). Toward understanding the role of tasks in behavioural research. *Acta Psychologica*, 31, 97-128.

- Haerem, T. & Rau, D. (2007). The influence of degree of expertise and objective task complexity on perceived task complexity and performance. *Journal of Applied Psychology*, 92, 1320-1331.
- Halford, G.S., Baker, R., McCredden, J.E. & Bain, J.D. (2005). How Many Variables Can Humans Process? *Psychological Science*, 16(1), 70-76.
- Halford, G.S., Maybery, M.T. & Bain, J.D. (1986). Capacity limitations in children's reasoning tasks: A dual task approach. *Child Development*, 57, 616-627.
- Halford, G.S., Wilson, W.H., & Phillips, S. (1998a). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioural & Brain Sciences*, 21, 803–831.
- Halford, G.S., Wilson, W.H., & Phillips, S. (1998b). Relational complexity metric is effective when assessments are based on actual cognitive processes. *Behavioural & Brain Sciences*, 21, 848-864.
- Hayes, J. R. & Simon, H. A. (1974). Understanding written problem instructions. In L. W. Gregg (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Erlbaum.
- Hesse, F.W. (1982). Effect of semantic context on the solution of complex problems. *Zeitschrift fur Experimentelle und Angewandte Psychologie*, 29, 62-91.
- Hogarth R.M. (1975). Decision time as a function of task complexity, In D. Wendt and C. Cvlek (Eds.), *Utility, Probability, and Decision Making* (pp. 321–338), Dordrecht, Holland: D. Reidel.
- Holyoak, K.J. (1991). Symbolic connectionism: toward third-generation theories of expertise. In Karl Anders Ericsson and Jacqui Smith (Eds.) *Toward a general theory of expertise: prospects and limits* (pp. 301-347), UK: Cambridge University Press.
- Horn, J.L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75, 242-259.
- Hornke, L. F. & Kersting, M. (2006). Optimizing quality in the use of Web-based and computer-based testing for personnel selection. In D. Bartram & R.K. Hambleton (Eds.), *Computer-based testing and the Internet: Issues and advances* (pp. 149-162). New York: John Wiley.
- Howie, D.E. & Vicente, K.J. (1998). Measures of operator performance in complex, dynamic microworlds: advancing the state of the art. *Ergonomics*, 41, 85–150.
- Hu, J., Huhmann, B.A. & Hyman, M.R. (2007). The relationship between task

- complexity and information search: the role of self-efficacy. *Psychology & Marketing*, 24, 253-270.
- Hulshof, C.D. & De Jong, T. (2006). Using just-in-time information to support scientific discovery learning in a computer-based simulation. *Interactive Learning Environments*, 14, 79-94.
- J. Schaeffer, J., Burch, N., Björnsson, Y., Kishimoto, A., Müller, M., Lake, R., Lu, P. & Sutphen, S. (2007). Checkers is solved. *Science*, 317, 1518-1522.
- Jäger, A.O., Süß, H.M. & Beauducel, A. (1997). *Berliner Intelligenzstruktur - Test. Form 4*. Göttingen: Hogrefe.
- Jobidon, M.E., Rousseau, R. & Breton, R. (2005). The effect of variability in temporal information on the control of a dynamic task. *Theoretical Issues in Ergonomics Science*, 6, 49-62.
- Jonassen, D.H. & Hung, W. (2008). All problems are not equal: Implications for problem-based learning. *The interdisciplinary journal of problem-based learning*, 2, 6-28.
- Joslyn, S. & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. *Journal of Experimental Psychology. Applied*, 4, 16-43.
- Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.
- Kalyuga, S., Chandler, P. & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 126-136.
- Kane, M.J., Hambrick, D.Z., Tuholski, S. W., Wilhelm, O., Payne, T.W. & Engle, R.W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189-217.
- Kanfer, R. & Ackerman, P.L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657-690.
- Kerstholt, J.H. & Raaijmakers, J.G.W. (1997). Decision making in dynamic task environments. In R. Ranyard, R. W. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations*. Norwood, NJ: Ablex.

- Kerstholt, J.H. (1992). Information search and choice accuracy as a function of task complexity and task structure. *Acta Psychologica*, 80, 185-197.
- Keys, J. & Biggs, W. (1990). A review of business games. In J. Gentry (Ed.), *Guide to business gaming and experiential learning* (pp. 48-73). London: Kogan Page.
- Kirschner, P.A., Sweller, J. & Clark, R.E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist*, 41, 75-86.
- Klahr, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT Press
- Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-55.
- Klahr, D., Fay, A.L. & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, 24, 111-146.
- Kleinmuntz, D.N. & Thomas, J. (1987). The value of action and inference in dynamic decision making. *Organizational Behaviour and Human Decision Processes*, 39, 341– 364.
- Kluge, A. (2008). Performance assessments with microworlds and their difficulty. *Applied Psychological Measurement*, 32, 156-180.
- Kluwe, R. H., Haider, H. & Misiak, C. (1990). Learning by doing in the control of a complex system. In H. Mandl, E. de Corte, N. Bennet, & H.F. Friedrich (Eds.), *Learning and Instruction*. Oxford: Pergamon Press
- Kluwe, R. H., Misiak, C. & Haider, H. (1991). The control of complex systems and performance in intelligence tests. In H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 227–244). Hillsdale, NJ: Erlbaum.
- Kluwe, R.H. (1993). Knowledge and performance in complex problem solving. In G. Strube and K. Wender (Eds.) *The cognitive psychology of knowledge* (pp. 401-423). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.
- Kluwe, R.H. (1995). Single case studies and models of complex problem solving. In P.A. Frensch & J. Funke (Eds.) *Complex Problem Solving: The European Perspective* (pp. 269-294). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kotovsky, K., Hayes, J.R. & Simon, H.A. (1985). Why are some problems hard?

- Evidence from Tower of Hanoi. *Cognitive Psychology*, 17, 248-294.
- Kröner, S. (2001). Intelligenzdiagnostik per Computersimulation [Intelligence assessment via computer simulation]. Münster: Waxmann.
- Kröner, S. & Leutner, D. (2002). MultiFlux—Pilotstudie für die Entwicklung eines Verfahrens zur simulationsbasierten Intelligenzdiagnostik [MultiFlux—Pilot study for the development of a simulation-based method for intelligence assessment]. *Zeitschrift für Arbeits und Organisationspsychologie*, 46, 84–88.
- Kröner, S., Leutner, D. & Plass, J. L. (2000). Expanding the concept of intelligence testing: The use of domain-general computer simulations. Poster presented at the APA Conference Models of intelligence for the next millennium. New Haven Yale University (June 14-18).
- Kröner, S., Plass, J.L. & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, 33, 347 – 368.
- Kyllonen, P. C. (1985). *Dimensions of information processing speed* (Rep.No. AFHRL-TP-84-56). Brooks Air Force Base, TX: Air Force Systems Command.
- Kyllonen, P. C. (1994). Aptitude testing inspired by information processing: A test of the four-sources model. *Journal of General Psychology*, 120, 375–405.
- Kyllonen, P. C. & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14, 389–433.
- Latzina, M. (1990). Problem solving: Using various dynamic model presentations. *Computers in Human Behaviour*, 6, 97-113.
- Lazonder, A.W., Wilhelm, P. & Hagemans, M.G. (2008). The influence of domain knowledge on strategy use during simulation-based inquiry learning. *Learning and Instruction*, 18, 580-592.
- Lazonder, A.W., Wilhelm, P. & van Lieburg, E. (2009). Unraveling the influence of domain knowledge during simulation-based inquiry learning. *Instructional Science*, 37, 437-451.
- Lebiere, C., Wallach, D. & Taatgen, N. A. (1998). Implicit and explicit learning in ACT-R. In F. Ritter & R. Young (Eds.), *Cognitive Modelling II* (pp. 183-193). Nottingham: Nottingham University Press.
- Leutner, D. (1993). Guided discovery learning with computer-based simulation games: effects of adaptive and non-adaptive instructional support. *Learning*

- and Instruction*, 3, 113-132.
- Leutner, D. (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behaviour*, 18, 685-697.
- Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492-527.
- Logan, G.D. (1990). Repetitive priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, 22, 1-35.
- Lohman, D.F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp.285-340). New York: Cambridge University.
- Mackinnon & Wearing, A.J. (1985). Systems analysis and dynamic decision making. *Acta Psychologica*, 58, 159-172.
- Marescaux, P.J., Luc, F. & Karnas, G. (1989). Modes d'apprentissage selectif et nonselectif et connaissances acquies au control d'un processus: Evaluation d'un modele simule [Selective and nonselective learning modes and acquiring knowledge of process control: Evaluation of a simulation model]. *Cahiers de Psychologie Cognitive*, 9, 239-264.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R.E. (2004) Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59, 14-19.
- McCarthy, J. & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In D. Michie, & B. Meltzer, *Machine intelligence* (Vol. 4) (pp. 463-502). Edinburgh University Press, 463-502.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Misiak, C. & Kluwe, R.H. (1986). *Complex system control*. Research report No. 13. Institut fuer Kognitionsforschung, Universitaet der Bundeswehr Hamburg.
- Mone, A.A. & Shalley, C.E. (1995). Effects of task complexity and goal specificity on change in strategy and performance over time. *Human Performance*, 8, 243-262.
- Müller, H. (1993). Komplexes Problemlösen: Reliabilität und Wissen [Complex problem solving: Reliability and knowledge]. Bonn: HoloS.
- Mumford, M.D., Baughman, W.A., Uhlman, C.E., Costanza & Threlfall, V. (1993).

- Personality variables and skill acquisition: Performance while practicing a complex task. *Human Performance*, 6, 345-381.
- Naylor, G.F. (1968). Perception times and rates as a function of the qualitative and quantitative structure of the stimulus. *Australian Journal of Psychology*, 20, 165-172.
- Neisser, U. (1976). General, academic, and artificial intelligence. In: Resnick, L. (Ed.), *The nature of intelligence* (pp.135-144), Lawrence Erlbaum Associates, Hillsdale, NJ.
- Neisser, U., Boodoo, G., Bouchard Jr., T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J. & Urbina, S. (1996). Intelligence: knowns and unknowns. *American Psychologist*, 51, 77–101.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood, NJ: Prentice-Hall.
- Norris, D. (1986). External validity of business games. *Developments in Business Simulation and Experiential Learning*, 13, 126-129.
- Oberauer, K. (1993). Die Koordination kognitiver Operationen. Eine Studie zum Zusammenhang von 'working memory' und Intelligenz [The coordination of cognitive operations. A study on the relationship of working memory and intelligence]. *Zeitschrift für Psychologie*, 201, 57-84
- Oberauer, K., Schulze, R., Wilhelm, O. & Süß, H.M. (2005). Working memory and intelligence – their correlation and their relation: A comment on Ackerman, Beier, and Boyle. *Psychological Bulletin*, 131, 61-65.
- Oberauer, K., Süß, H.M., Schulze, R., Wilhelm, O. & Wittmann, W.W. (2000). Working memory capacity - Facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017–1045.
- OECD Program for International Student Assessment, Retrieved June 2010, from http://www.pisa.oecd.org/pages/0,3417,en_32252351_32235731_1_1_1_1_1_1_1_0.html
- Olshavsky, R.W. (1979). Task complexity and contingent processing in decision making: A replication and extension. *Organizational Behaviour and Human Decision Processes*, 24, 300–316.
- Omodei, M.M. & Wearing, A.J. (1995). The fire chief micro-world generating program: an illustration of computer-simulated microworlds as an

- experimental paradigm for studying complex decision-making behaviour. *Behaviour Research Methods, Instruments & Computers*, 27, 303–316.
- Osman, M. (2008a). Observation can be as effective as action in problem solving. *Cognitive Science: A Multi-disciplinary Journal*, 32(1), 162-183.
- Osman, M. (2008b). Seeing is as good as doing. *The Journal of Problem Solving*, 2, 29-40.
- Osman, M. (2010). Controlling uncertainty: a review of human behavior in complex dynamic environments. *Psychological Bulletin*, 136(1), 65-86.
- Paquette, L. & Kida, T. (1988). The effect of decision strategy and task complexity on decision performance. *Organizational Behaviour and Human Decision Processes*, 41, 128-142.
- Park J, Jung W, Kim J, Ha J, Shin Y. (2001). The step complexity measure for emergency operating procedures—comparing with simulation data. *Reliability Engineering and System Safety*, 74, 63–74.
- Park, J. & Jung, W. (2007). A study on the development of a task complexity measure for emergency operating procedures of nuclear power plants. *Reliability Engineering and System Safety*, 92, 1102-1116.
- Park, J., Jung, W., & Ha, J. (2001). Development of the step complexity measure for emergency operating procedures using entropy concepts. *Reliability Engineering and System Safety*, 71, 115-130.
- Park, J., Jung, W., Kim, J. & Ha, J. (2003). Step complexity measure for emergency operating procedures—determining weighting factors. *Nuclear Technology*, 143, 290–308.
- Park, J., Jung, W., Kim, J. & Ha, J. (2003). The step complexity measure—its meaning and applications. *Journal of Korean Nuclear Society*, 35, 80-90.
- Payne, J.W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organisational Behaviour and Human Performance*, 16, 366-387.
- Pizlo, Z., Stefanov, E., Saalweachter, J., Li, Z., Haxhimusa, Y. & Kropatsch, W.G. (2006). Travelling salesman problem: A foveating pyramid model. *The Journal of Problem Solving*, 1, 83-101.
- Preußler, W. (1996). Zur Rolle expliziten und impliziten Wissens bei der Steuerung dynamischer Systeme [On the role of explicit and implicit knowledge in

- controlling dynamic systems]. *Zeitschrift für Experimentelle Psychologie*, 43, 399-434.
- Preußler, W. (1997). Effekte des Kontexts auf den Wissenserwerb bei der Steuerung eines dynamischen Systems [Effects of context on the acquisition of knowledge in controlling a dynamic system]. *Sprache & Kognition*, 16, 48-59.
- Preußler, W. (1998). Strukturwissen als Voraussetzung für die Steuerung komplexer dynamischer Systeme [Structural knowledge as a precondition of controlling complex dynamic systems]. *Zeitschrift für Experimentelle Psychologie*, 45, 218-240.
- Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg [On the relationship between test intelligence and success in problem solving]. *Zeitschrift für Psychologie*, 189, 79-100.
- Putz-Osterloh, W. (1993). Strategies for knowledge acquisition and transfer of knowledge in dynamic tasks. In G. Strube and K. Wender (Eds.) *The cognitive psychology of knowledge* (pp. 331-350). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.
- Putz-Osterloh, W. (1993b). Unterschiede im Erwerb und in der Reichweite des Wissens bei der Steuerung eines dynamischen Systems [Differences in the acquisition and extent of knowledge during system control]. *Zeitschrift für Experimentelle und angewandte Psychologie*, 40, 386-410.
- Putz-Osterloh, W. & Lüer, G. (1981). Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest [On the predictability of complex problem solving performance by intelligence test scores]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 28, 309-334.
- Quesada, J., Kintsch, W. & Gomez, E. (2005). Complex problem-solving: a field in search of a definition? *Theoretical Issues in Ergonomics Science*, 6, 5-33.
- Raaheim, K. (1985). Task novelty and intelligence. *Scandinavian Journal of Psychology*, 26(1), 35-41.
- Raaheim, K. (1988). Intelligence and task novelty. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, vol 4 (pp. 73-97). Hillsdale: Erlbaum.

- Raaheim, K. (1989). The dilemma of the intelligent problem solver. In I. Bjorgen (Ed.) *Basic Issues in Psychology*. Bergen: Sigma.
- Raaheim, K. (1991). Is the high IQ person really in trouble? Why? In H.A.H. Rowe (Ed.), *Intelligence, reconceptualization and measurement* (pp. 35-46). Hillsdale, NJ: Erlbaum.
- Rasmussen, J. & M. Lind (1981). Coping with Complexity. Risø-M-2293, presented at European Annual Conference on Human Decision and Manual Control, Delft.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis method (2nd ed.)*. Newbury Park, CA: Sage.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y. & Congdon, R. T. (2000). *HLM 6: Hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.
- Raven, J., Raven, J.C. & Court, J. H. (1980). *Advanced progressive matrices*. Oxford, England: Oxford Psychologists Press.
- Raven, J., Raven, J.C. & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. San Antonio, TX: Harcourt Assessment.
- Raven, J.C. (1958). *Standard Progressive Matrices. Sets A, B, C, D & E, H. K.* Lewis, London.
- Reichert, U. & Dörner, D. (1988). Heurismen beim Umgang mit einem "einfachen" dynamischen System [Heuristics in the control of a "simple" dynamic system]. *Sprache Kognition*, 7, 12–24.
- Reif, F. & Heller, J. I. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist*, 17, 102- 127.
- Reither, F. (1981). About thinking and acting of experts in complex situations. *Simulation & Games*, 12, 125-140
- Rigas, G. (2000). *On the Relationship Between Psychometric Intelligence and Decision Making in Dynamic Systems*. 98 (Uppsala, Sweden: Uppsala University).
- Rigas, G. & Brehmer, B. (1999). Mental processes in intelligence tests and dynamic decision making tasks. In P. Juslin & H. Montgomery (Eds.), *Judgement and decision making: Neo-Brunswikean and process-tracing approaches* (pp. 45-

- 65). Hillsdale, NJ: Erlbaum
- Rigas, G., Carling, E. & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, 30, 436-480.
- Rolo, G. & Diaz-Caberera, D. (2005). Decision-making processes evaluation using two methodologies: field and simulation techniques. *Theoretical Issues in Ergonomics Science*, 6, 35-48.
- Rouse, W.B. & Rouse, S.H. (1979). Measures of complexity and fault diagnosis tasks. *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 720-727.
- Rowe, H.A.H. (1985). Problem solving and intelligence. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Sanderson, P.M. (1989). Verbalizable knowledge and skilled task performance: Association, dissociation, and mental models. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 15, 729-747.
- Sanderson, P.M. & Vicente, K.J. (1986). *Verbalizable knowledge and skilled task performance: Explaining association and dissociation* (Tech. Rep. No. EPL-86-04). Urbana, IL: University of Illinois at Urbana-Champaign, Engineering Psychology Research Laboratory.
- Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schoppek, W. (1998). Toward a theory of the control of dynamic systems. (Paper presented at the 2. European Conference on Cognitive Modelling 1998 in Nottingham, UK).
- Schoppek, W. (2002). Examples, rules and strategies in the control of dynamic systems. *Cognitive Science Quarterly*, 2, 63-92.
- Schoppek, W. (2004). Teaching structural knowledge in the control of dynamic systems: Direction of causality makes a difference. In K. D. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp.1219-1224). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schroder, H., Driver, M., & Streufert, S. (1967). *Human information processing*. New York: Holt, Rinehart, and Winston.
- Shanahan, M. & Baars, B. (2005). Applying global workspace theory to the frame

- problem. *Cognition*, 98, 157-176.
- Shanks, D.R. & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioural & Brain Sciences*, 17, 367-447.
- Shute, V.J. (1990). Individual differences in learning from an intelligent discovery world: Smithtown (Report No. AFHRL-TP-89-57). Air Force Systems Command, Brooks Air Force Base, TX.
- Simon, H.A. and Lea, G. (1974). Problem solving and rule induction: A unified view. In Lee W. Gregg (Ed.), *Knowledge and Cognition*. Erlbaum, Hillsdale, N.J.
- Simonson, I. & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29, 281-295.
- Snodgrass, J. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.
- Snow, R.E. & Yallow, E. (1982). Education and intelligence. In R. J. Sternbert (Ed.), *Handbook of Human Intelligence* (pp. 493-586). London: Cambridge University Press.
- Snow, R.E., Kyllonen, P.C. & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Eds.). *Advances in the Psychology of Human Intelligence*, vol. 2 (pp. 47-103). Hillsdale, NJ7 Erlbaum.
- Snow, R.E. (1986). Individual differences and the design of education programs. *American Psychologist*, 41, 1029-1039.
- Snow, R.E. (1989). Towards assessment of cognitive and conative structures in learning. *Educational Researcher*, 18, 8-14.
- Snow, R.E. & Lohman, D.F. (1989). Implication of cognitive psychology for education measurement. In R.L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 263-331). New York: Macmillan.
- Speier, C. (2006). The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies*, 64, 1115-1131.
- Spring, M., Wagener, D. & Funke, J. (2005). The role of emotions in complex problem-solving. *Cognition and Emotion*, 19, 1252-1261
- Spilisbury, G., Stankov, L. & Roberts, R. (1990). The effect of test's difficulty on its

- correlation with intelligence. *Personality and Individual Differences*, 11, 1069-1077.
- Stahl, E., Pieschl, S. & Bromme, R. (2006). Task complexity, epistemological beliefs and metacognitive calibration: An exploratory study. *Journal of educational computing research*, 35, 319-338.
- Stainton, A.J., Johnson, J.E. & Borodzicz, E.P. (2010). Educational Validity of Business Gaming Simulation: A Research Methodology Framework. *Simulation & Gaming*, 41, 705-723.
- Stankov, L. (1999). Complexity, metacognition and fluid intelligence. *Intelligence*, 27, 1–23.
- Stankov, L. & Crawford, J.D. (1993). Ingredients of complexity in fluid intelligence. *Learning & Individual Differences*, 5, 73-111.
- Stanley, W. B., Mathews, R. C., Buss, R. R., & Kotler-Cope, S. (1989). Insight without awareness: On the interaction of verbalization, instruction, and practice in a simulated process control task. *Quarterly Journal of Experimental Psychology*, 41, 553–577.
- Sterman, J.D. (1989). Misperceptions of feedback in dynamic decision making. *Organisational Behaviour and Human Decision Processes*, 43, 301-335.
- Sterman, J.D. (1989a). Modelling managerial behaviour: misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35, 321–339.
- Sterman, J.D. (1994). Learning in and about complex systems. System dynamics review. *Systems Thinkers, Systems Thinking*, 10, 291–330.
- Sterman, J.D. (2000). Learning in and about complex systems. *Reflections: The SoL Journal*, 1, 24–51.
- Sternberg, R. J. (1986). *Intelligence applied*. Orlando, FL: Harcourt Brace College.
- Sternberg, R. J. (1988). Intelligence. In R. J. Sternberg, & E. E. Smith (Eds.), *The psychology of human thought* (pp. 267-308). Cambridge, UK:Cambridge University Press.
- Sternberg, R. J. & Frensch, P. A. (Eds.) (1991). *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Erlbaum.
- Sternberg, R.J. (1977). *Intelligence, Information Processing, and Analogical Reasoning*. Hillsdale, NJ : Erlbaum.
- Sternberg, R.J. (1982). Reasoning, problem solving, and intelligence. In R.J.

- Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). Cambridge: Cambridge University Press.
- Stock, D. & Watson, C. (1984). Human Judgment Accuracy, Multidimensional Graphics, and Humans versus Models. *Journal of Accounting Research*, 192-206.
- Strauß, B. (1993). *Konfundierungen beim komplexen Problemlösen* [Confundations in complex problem-solving: The influence of the ratio of correct solutions on problem-solving in complex situations]. Bonn: Holos.
- Strohschneider, S. (1986). Zur Stabilität und Validität von Handeln in komplexen Realitätsbereichen (On the stability and reliability of complex problem solving behaviour). *Sprache and Kognition*, 5, 42-48.
- Süß, H.M. (1996). *Intelligenz, Wissen und Problemlösen* [Intelligence, Knowledge, and Problem Solving], Hogrefe, Göttingen.
- Süß, H.M. (1999). Intelligenz und komplexes Problemlösen: Perspektiven für eine Kooperation zwischen differentiell-psychometrischer und kognitionspsychologischer Forschung [Intelligence and complex problem solving: Perspectives for a cooperation between differential–Psychometric and cognitive research], *Psychologische Rundschau*, 50, 220–228.
- Süß, H.M., Kersting, M. & Oberauer, K. (1991). Intelligenz und Wissen als Prädiktoren fuer Leistungen bei computersimulierten komplexen Problemen [Intelligence and knowledge as predictors for achievements with computer-simulated complex problems]. *Diagnostica*, 37, 334–352.
- Süß, H.M., Oberauer, K., Wittmann, W. W., Wilhelm, O. & Schulze, R. (2002). Working memory capacity explains reasoning ability—and a little bit more. *Intelligence*, 30, 261–288.
- Sweller, J. (1976). The effect of task complexity and sequence on rule learning and problem solving. *British Journal of Psychology*, 67, 553-558.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12, 257-285.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295-312.
- Sweller, J. (1998). Can we measure working memory without contamination from knowledge held in long-term memory? *Behavioural & Brain Science*, 21, 803-

- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Research Review*, 22, 123-138.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition & Instruction*, 12, 185–233.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176–192.
- Sweller, J., Van Merriënboerr, J.J.G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296.
- Taatgen, N.A. & Wallach, D. (2002). Whether Skill Acquisition is Rule or Instance Based is determined by the Structure of the Task. *Cognitive Science Quarterly*, 2, 1-42.
- Tabbers, H.K., Martens, R.L. & van Merriënboer, J.J.G. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology*, 74, 71 - 81.
- Tindall-Ford, S., Chandler, P. & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3, 257–287.
- Tschirigi, J.E. (1980). Sensible Reasoning: A Hypothesis About Hypotheses. *Child Development*, 51, 1-10.
- U. Funke (1995). Using Complex Problem Solving Tasks in Personnel Selection and Training. In P.A. Frensch & J. Funke (Eds.) *Complex Problem Solving: The European Perspective* (pp. 219 – 242). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vakkari, P. (1999). Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing and Management*, 35, 819-837.
- Vernon, P.E. (1965). Ability factors and environmental influences. *American Psychologist*, 20, 723–733.
- Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management*, 27, 103-119.

- Vollmeyer, R. & Burns, B. (1995). Does hypothesis instruction improve learning? In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp.711 – 776). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vollmeyer, R. & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education*, 14, 541–554.
- Vollmeyer, R. & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, 10, 293–309.
- Vollmeyer, R. Burns, B.D. & Holyoak, K. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, 20, 75-100.
- Vollmeyer, R., Rollett, W., & Rheinberg, F. (1997). How motivation affects learning. In M. G. Shafto, & P. Langley, *Proceedings of the nineteenth annual conference of the Cognitive Science Society* (pp. 796–801). Hillsdale, NJ: Erlbaum.
- Vollmeyer, R., Rollett, W., & Rheinberg, F. (1998). Motivation and learning in a complex system. In P. Nenniger, R. S. Jäger, A. Frey, & M. Wosnitza, *Advances in motivation* (pp. 53–67). Landau: Verlag Empirische Pädagogik.
- Waerm, Y. & Cañas, J.J. (2003). Microworld task environments for conducting research on command and control. *Cognition, Technology & Work*, 5, 181-182.
- Wagener, D. (2001). Intelligence and complex problem solving performance: new findings using a system that was optimized for assessment – Poster presented at 2001 ISSID Conference, Edinburgh.
- Wagner, K. & Wechsung, G. (1986). Computational Complexity. In: *Mathematische Monographien 19*, VEB Deutscher Verlag Wissenschaften, Berlin.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). London: Penguin.
- Wason, P. C., & Shapiro, D. A. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63–71.
- Wenke, D. & Frensch, P.A. (2003). Is success or failure at solving complex problems related to intellectual ability? In J.E. Davidson & R.J. Sternberg, (Eds.), *The psychology of problem solving* (pp. 87-126). New York: Cambridge University

Press.

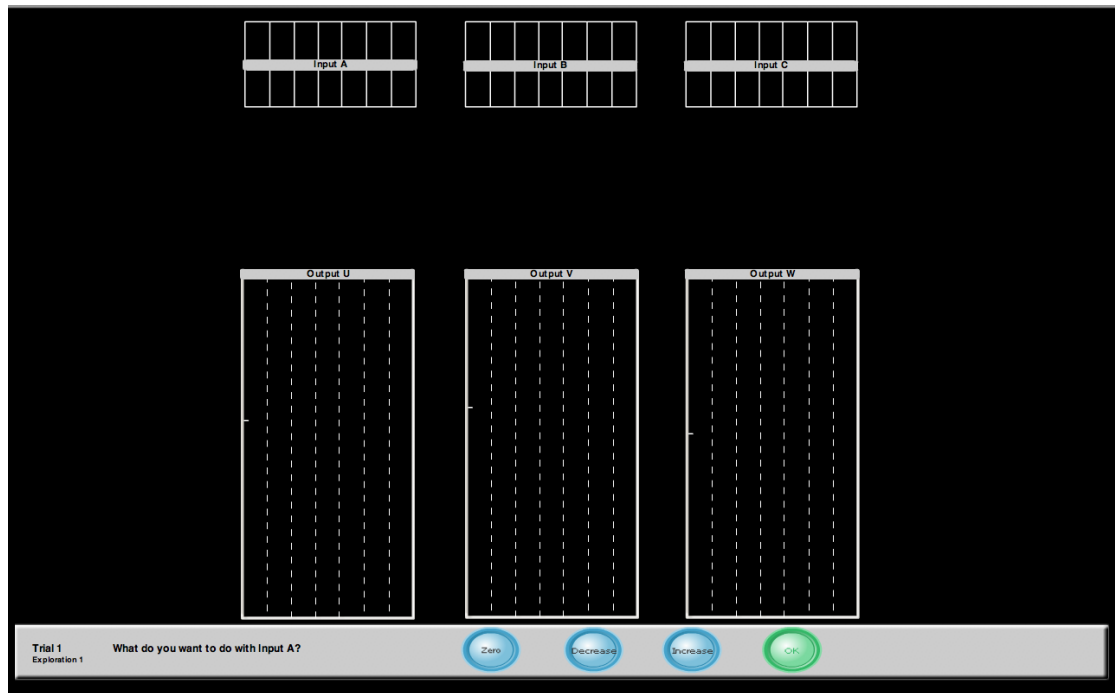
- Wenke, D., Frensch, P.A. & Funke, J. (2004). Complex problem solving and intelligence: Empirical relation and causal direction. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 160-187). New York: Cambridge University Press.
- Wirth, J. & Funke, J. (2005). Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme. In E. Klieme, D. Leutner & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern* (pp. 55-72). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wirth, J., Künsting, J. & Leutner, D. (2009) The impact of goal specificity and goal type on learning outcome and cognitive load. *Computers in Human Behaviour*, 25, 299-305
- Wittmann W.W. & Süß, H.M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry. In P.L. Ackerman and R.D. Roberts, (Eds.), *Learning and individual differences. Process, trait content determinants* (pp. 77–104), APA, Washington.
- Wittmann, W.W. & Hatrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioural Science*, 21, 393-409.
- Wohl, J.G. (1982). Maintainability prediction revisited: Diagnostic behaviour, system complexity, and repair time. *IEEE Transactions on Systems, Man and Cybernetics*, 12, 241-250.
- Wolfe, J. & Rogé, J.N. (1997). Computerized general management games as strategic management learning environments. *Simulation & Gaming*, 28, 423-441.
- Wood, R.E., Beckmann, J.F. & Birney, D.P. (2009). Simulations, learning and real world capabilities. *Education and Training*, 51, 491-510.
- Wood, R.E., Mento, A.J. & Locke, E.A. (1987). Task complexity as a moderator of goal effects: A meta-analysis. *Journal of Applied Psychology*, 72, 416-425.
- Wood, R.E. (1986). Task complexity: Definition of the construct. *Organizational Behaviour and Human Decision Processes*, 37, 60-82.

- Woods, D. (1988). Coping with complexity: the psychology of human behaviour in complex systems. In L.P. Goodstein, H.B. Andersen, S.E. Olsen (Eds.) *Tasks, errors, and mental models* (pp.128 - 148). Bristol: Taylor & Francis.
- Xiao, Y., Hunter, W.A., Mackenzie, C.F. & Jefferies, N.J. (1996). Task complexity in emergency medical care and its implications for team coordination. *Human Factors*, 38, 636-645.

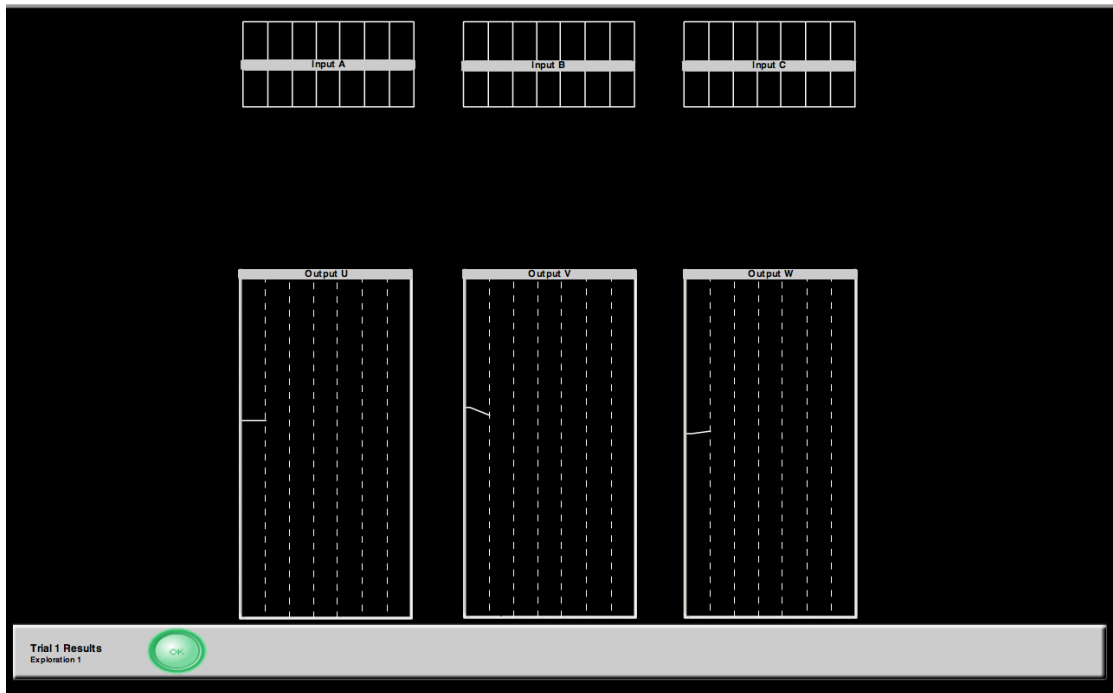
APPENDIX A

INSTRUCTIONS IN THE COMPLETE INFORMATION CONDITION

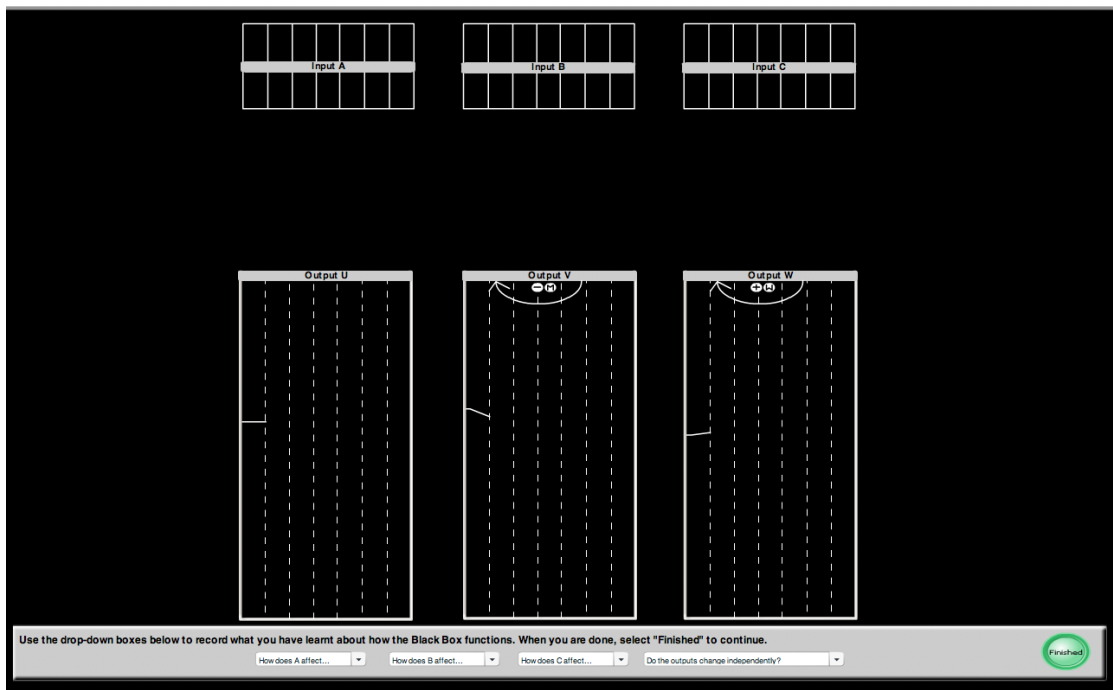
The following screenshots and text show the instructions given in the information condition for Study 1, and in the complete information condition in Study 2 and 3. The text was presented as a concurrent narration.



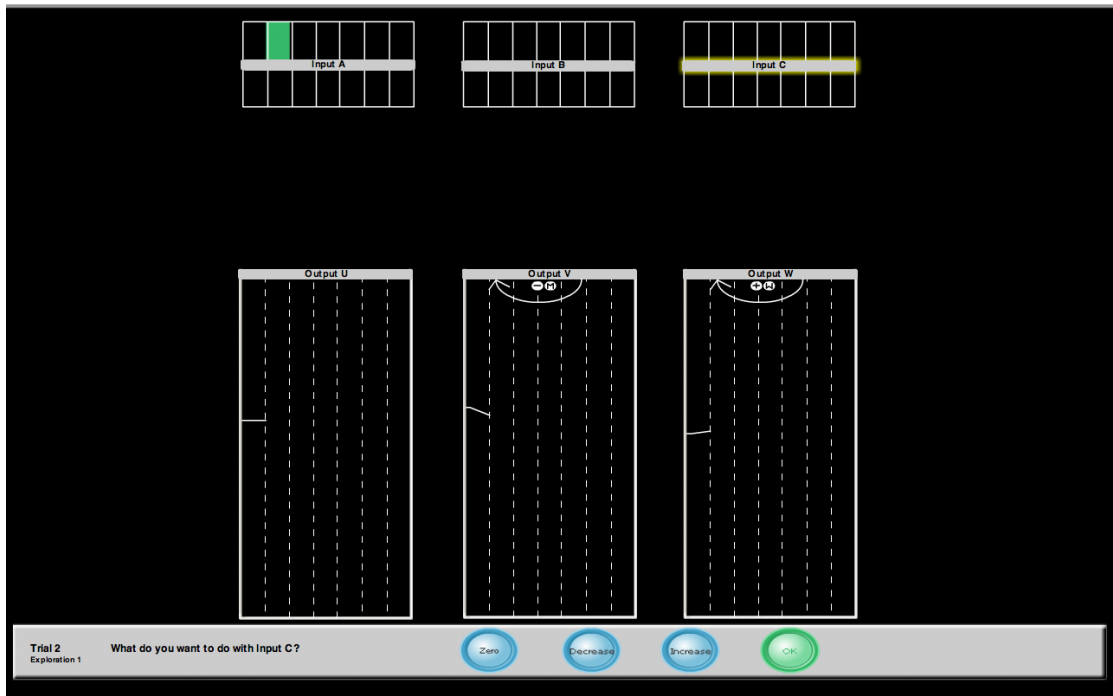
On the first trial, we should determine whether any of the outputs change independently of the inputs. So we should set each input to zero, and see if the outputs change.



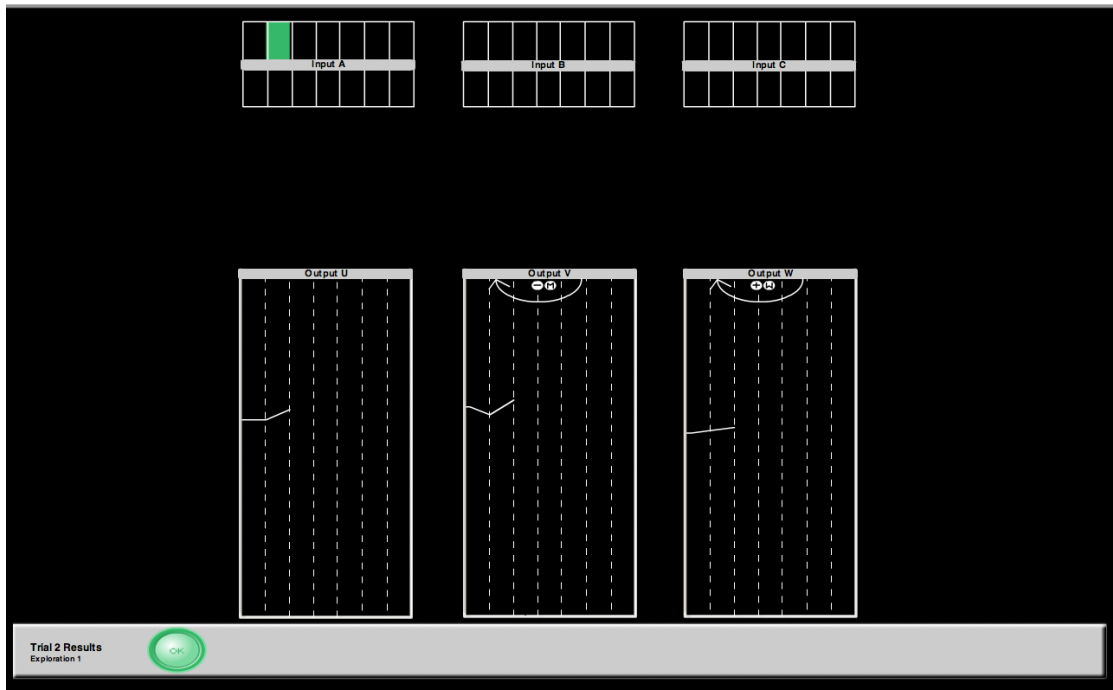
So we set each of the inputs to zero, and U didn't change, V decreased, and W increased. So it looks like V and W change on their own over time.



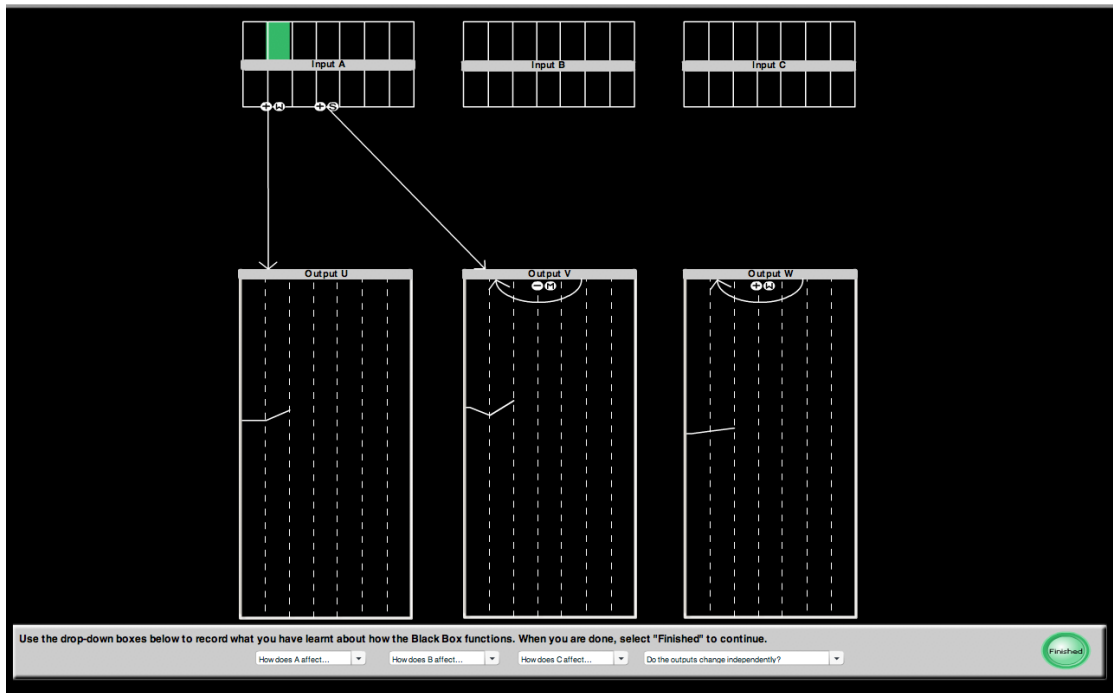
Let's record that result. So we found that V and W changed. So output V, yes it changes. It decreased, so that's a negative effect, and I'm going to say that it's a medium effect. Output W, yes it changed, it increased, so that's a positive effect, and it's quite small, so I'm going to say it's a weak effect.



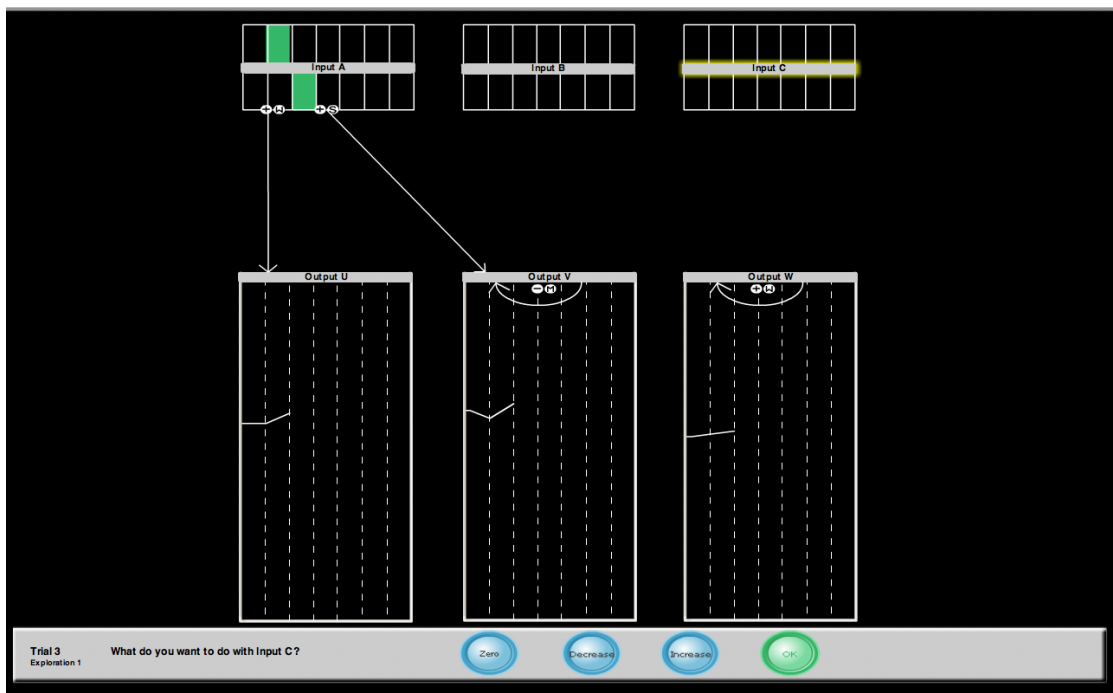
This time I'm going to look at the effect of input A on the outputs. So I'm going to increase A to maximum, and set each of the other inputs to zero.



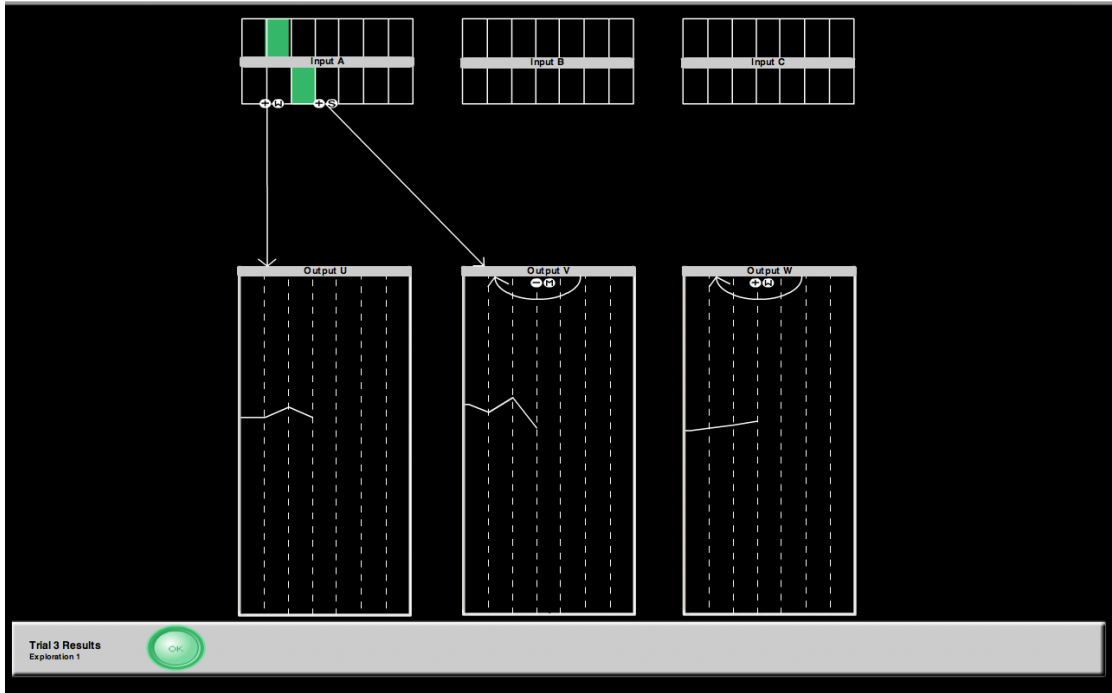
o I increased A to maximum, and U increased, V increased and W increased. Now we would expect W to increase on its own anyway, so it has not effect there. So we can say A effects U and V.



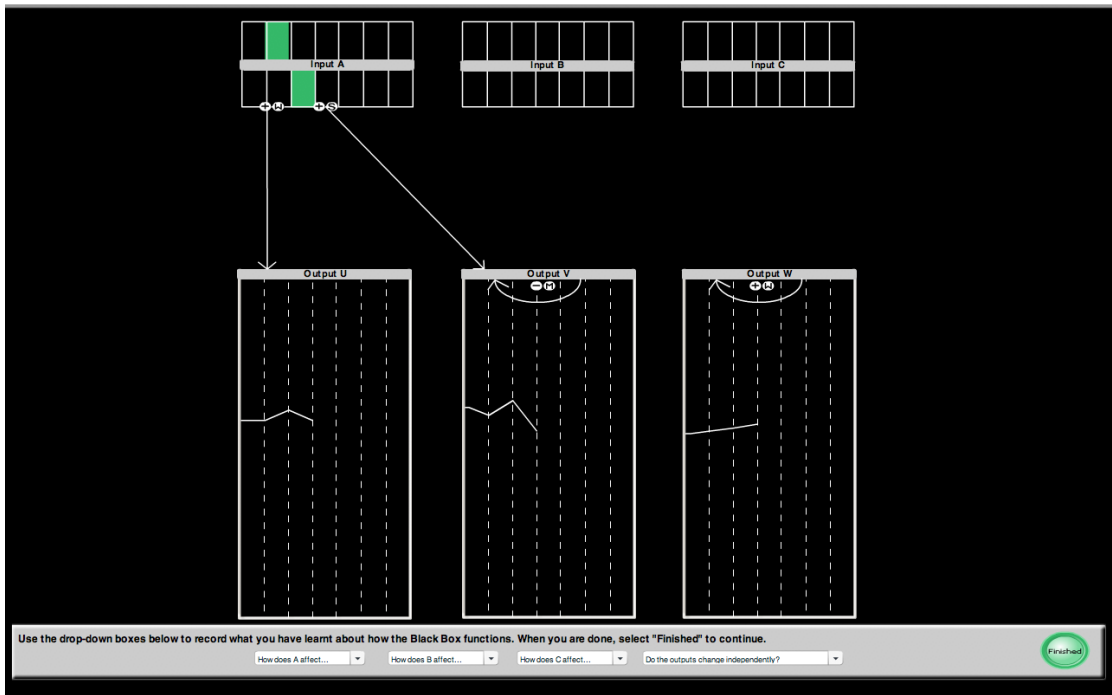
Let's record that result. So we found that input A effects output U. Yes there's an effect, increasing A increases U, so it's a positive effect. And I'm going to say it's a weak effect. We also found that A effects V. Increasing A increases V, yes there's an effect, it's a positive effect, and I'm going to say it's a strong effect because it has to work against the independent decline in V.



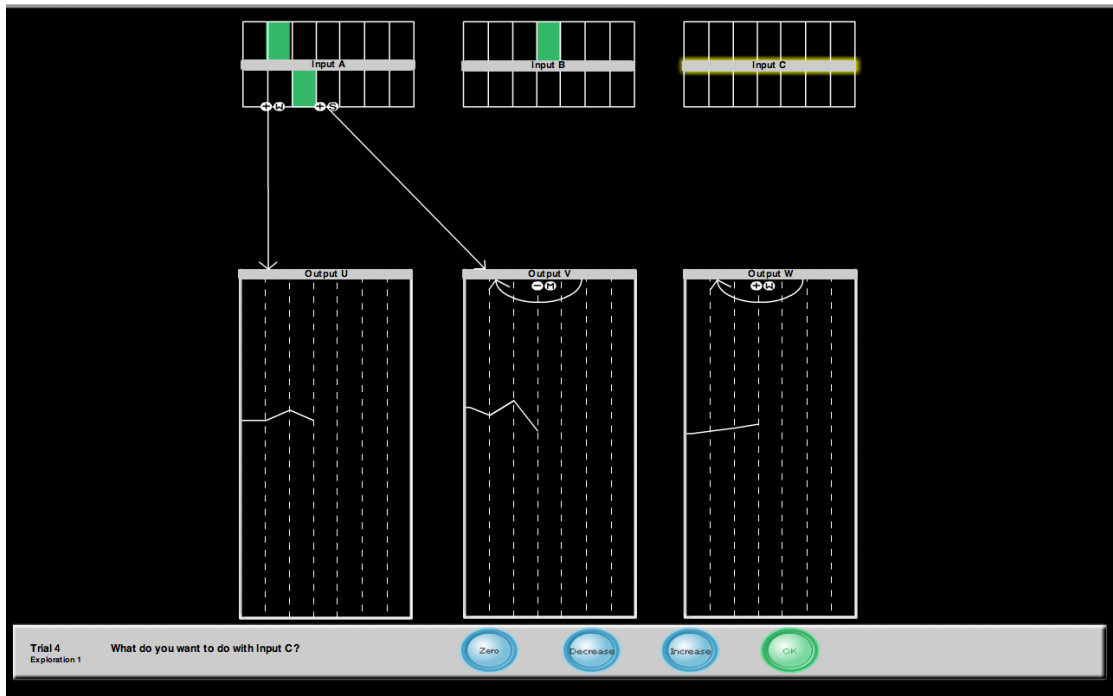
So this time I'm going to decrease A to minimum and set the inputs to zero.



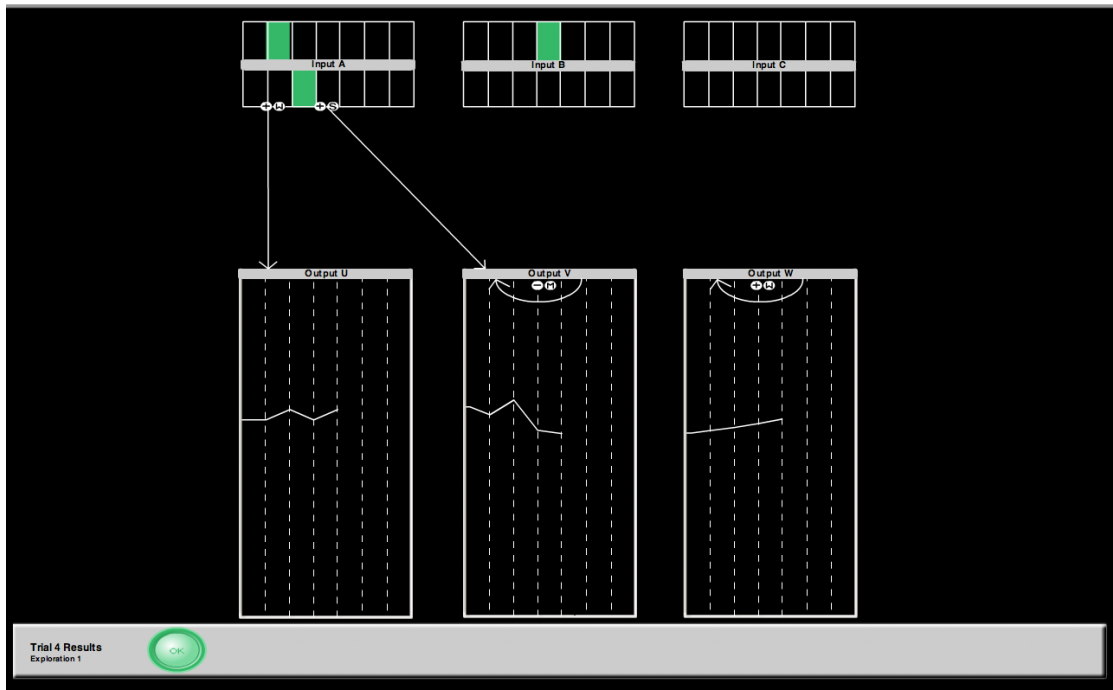
So we decreased A, and U decreased, V decreased and W increased slightly.



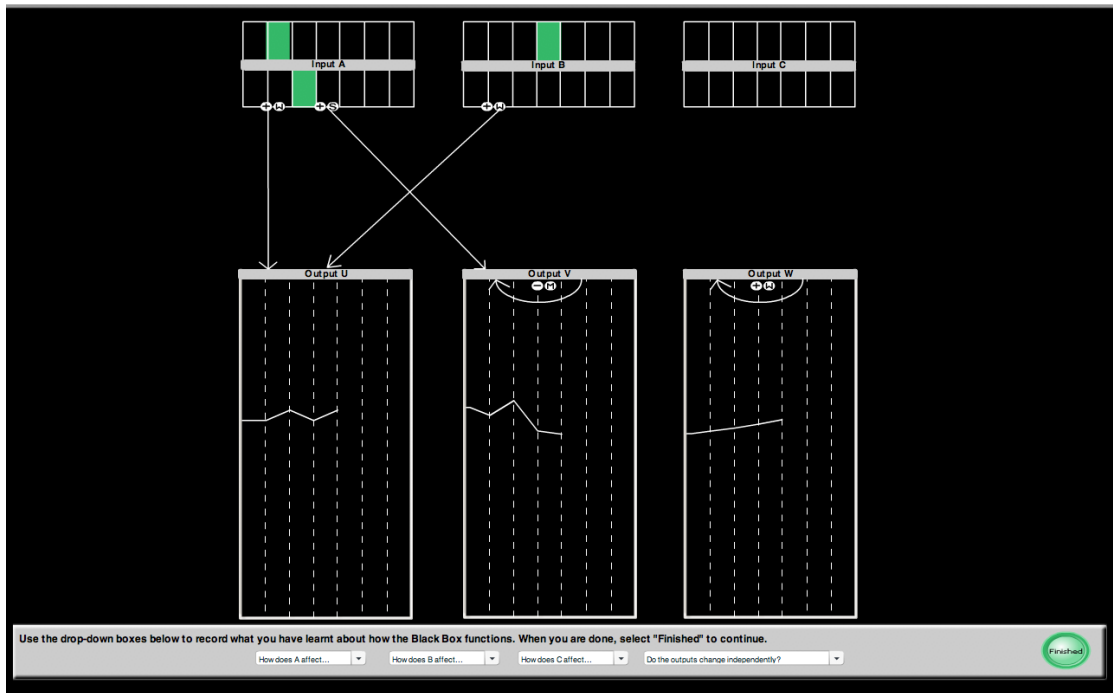
So I've already recorded that result that A affects U and V.



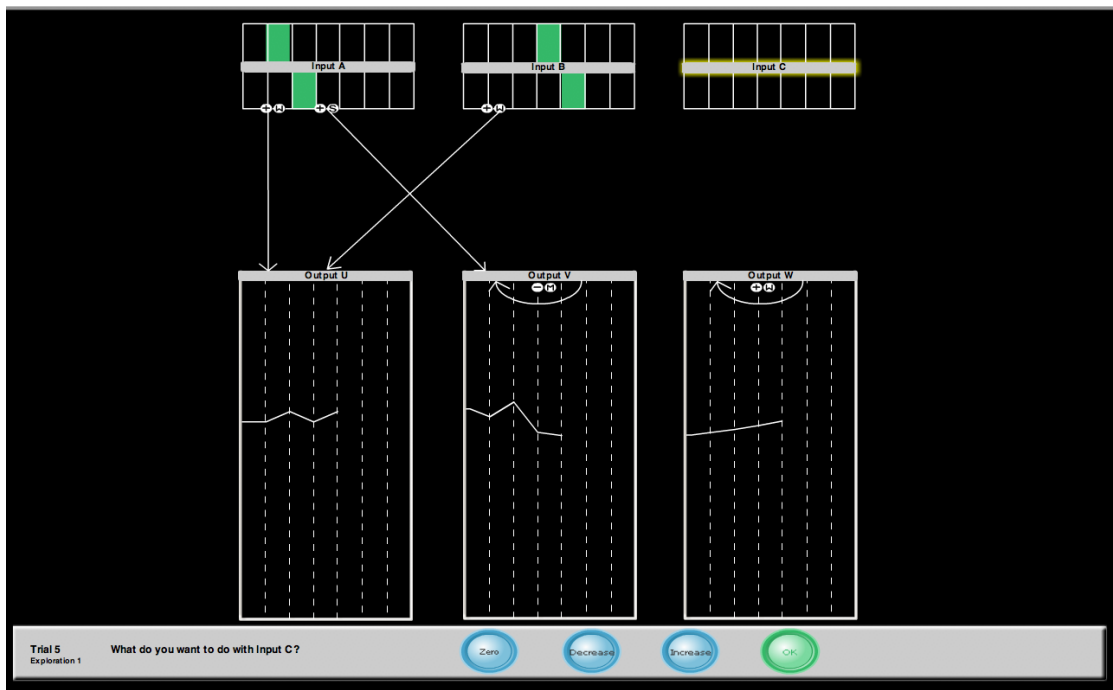
This time I'm going to look at the effect of B on the outputs. So I'm going to set A at zero, increase B to maximum, and set C at zero.



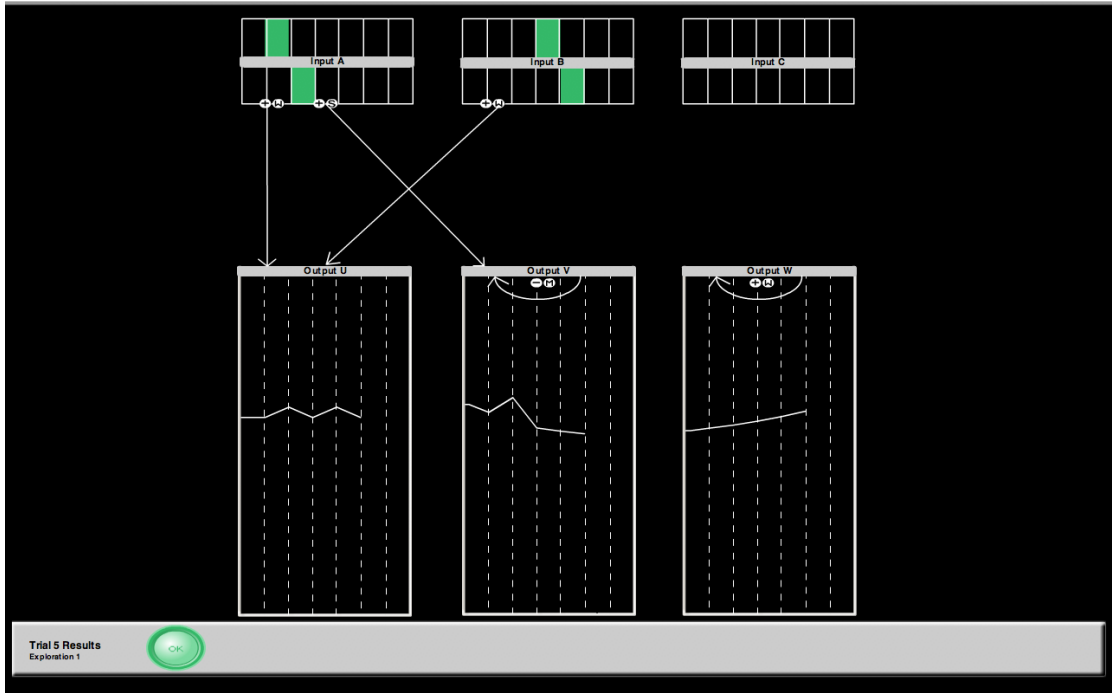
So increased B, and U increased, V decreased and W increased slightly. We know that V and W change on their own anyway, so it looks like B only effects U.



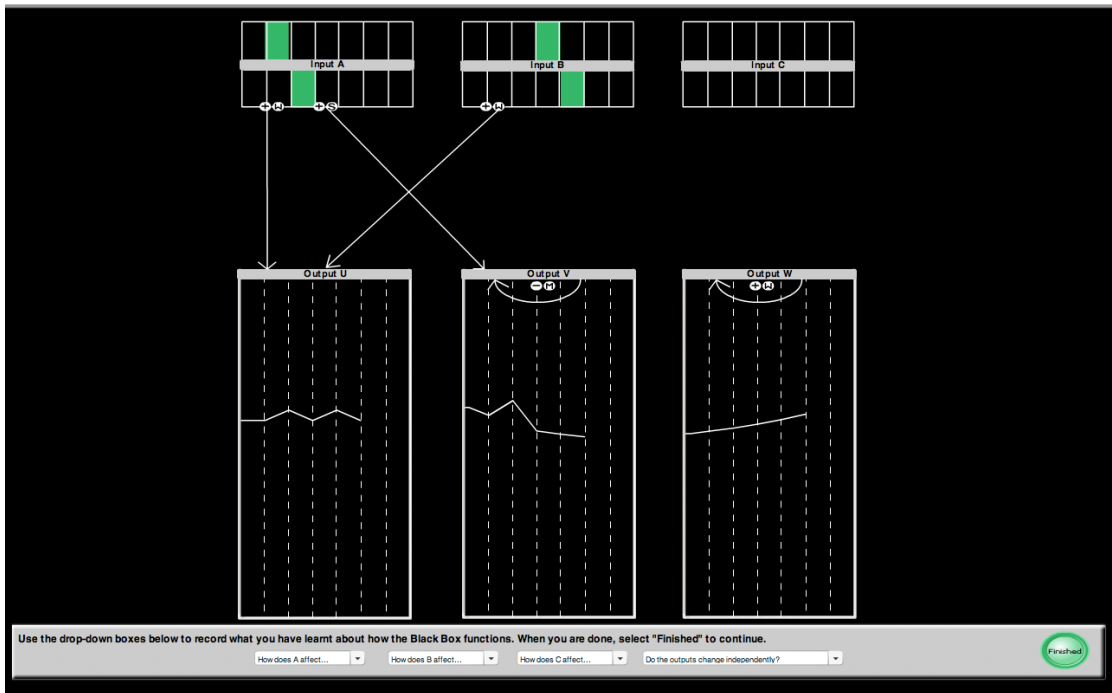
So let's record that result. B only affected U. Yes there was an effect, we increased B and U increased, so that's a positive effect, and I'm going to say it's a weak effect.



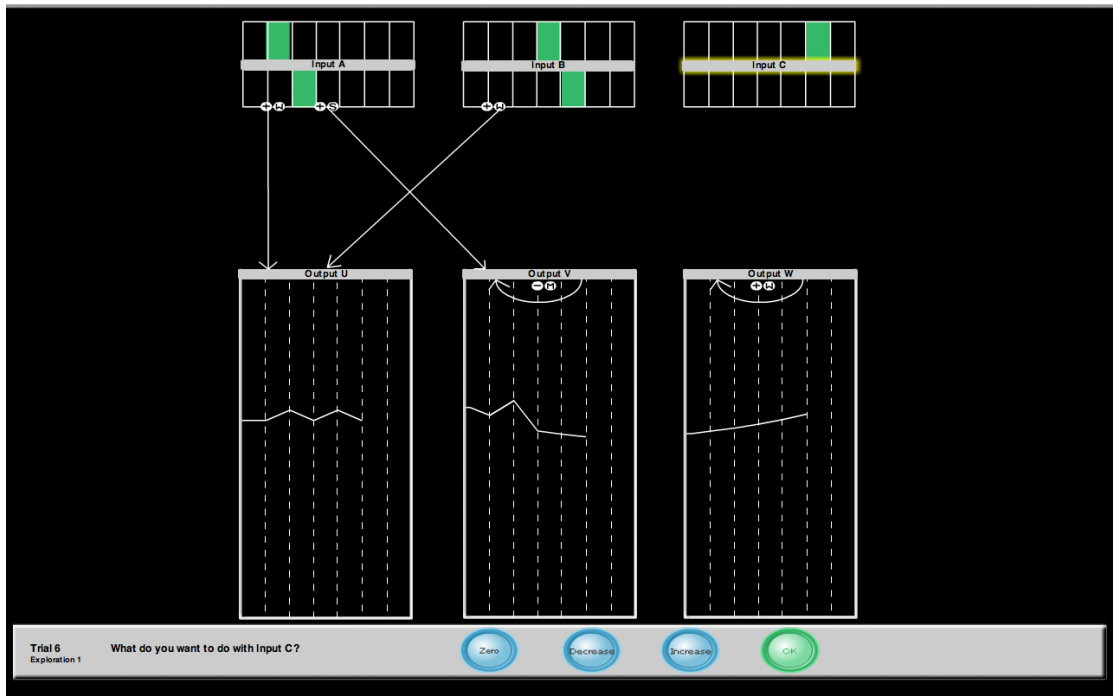
This time I'm going to look at the effect of B again, but I'm going to decrease it to minimum and set the other inputs at zero.



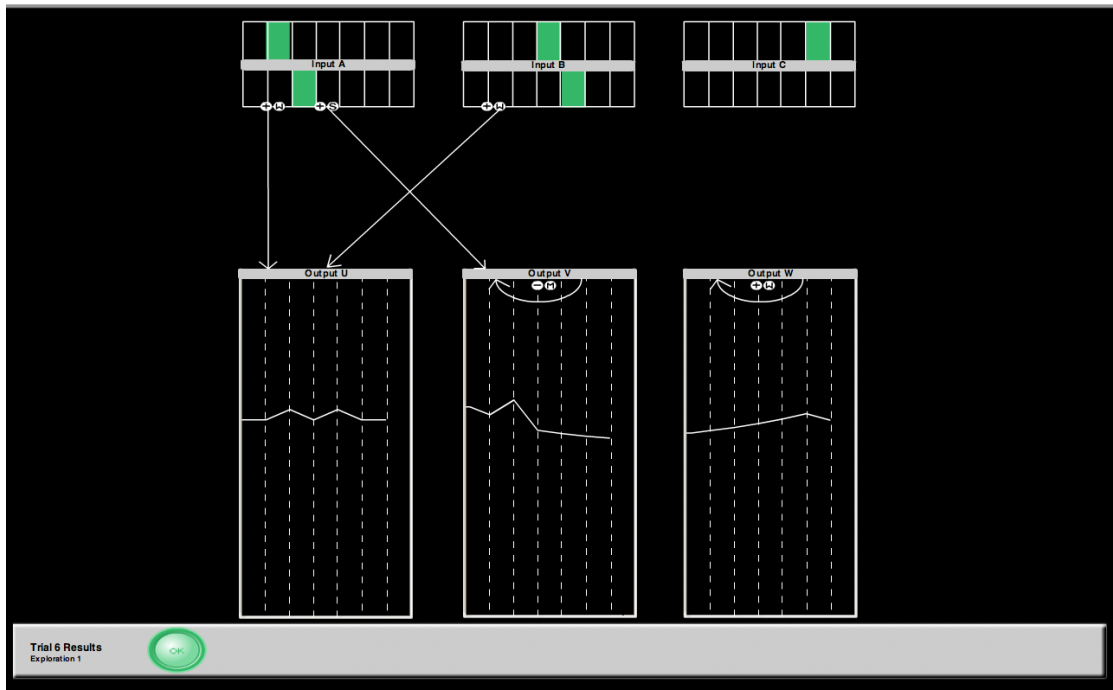
So I decreased B, and U decreased, V decreased and W increased slightly.



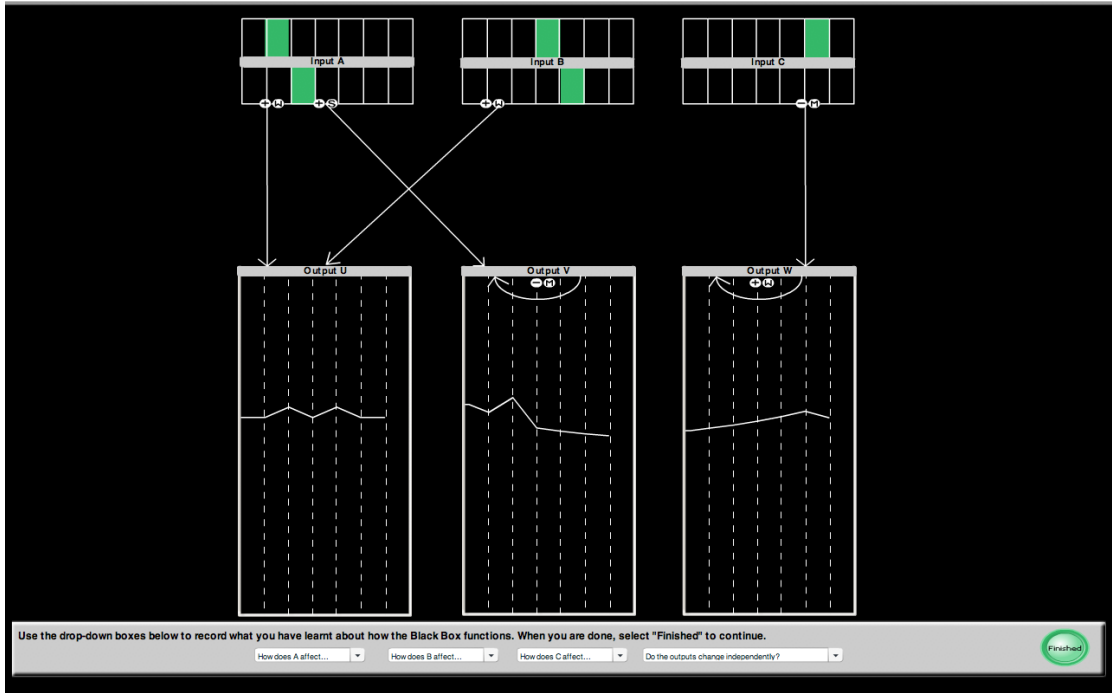
So I've already recorded that B only effects U.



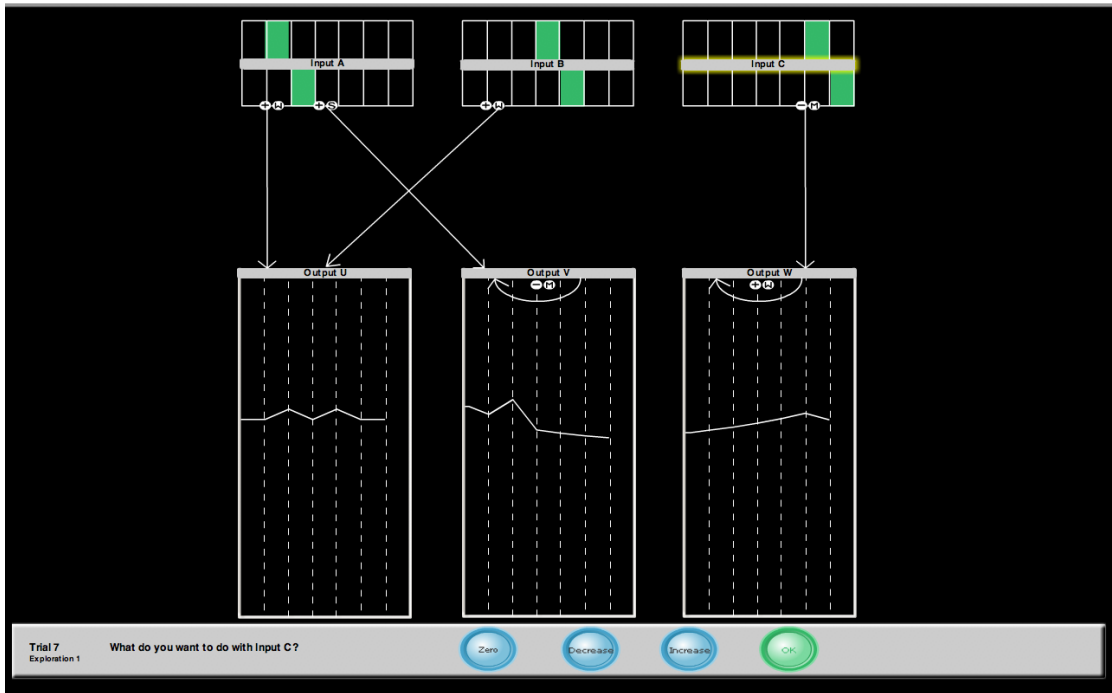
This time I'm going to look at the effect of C on the outputs. So I'm going to set A to zero, B to zero and increase C to maximum.



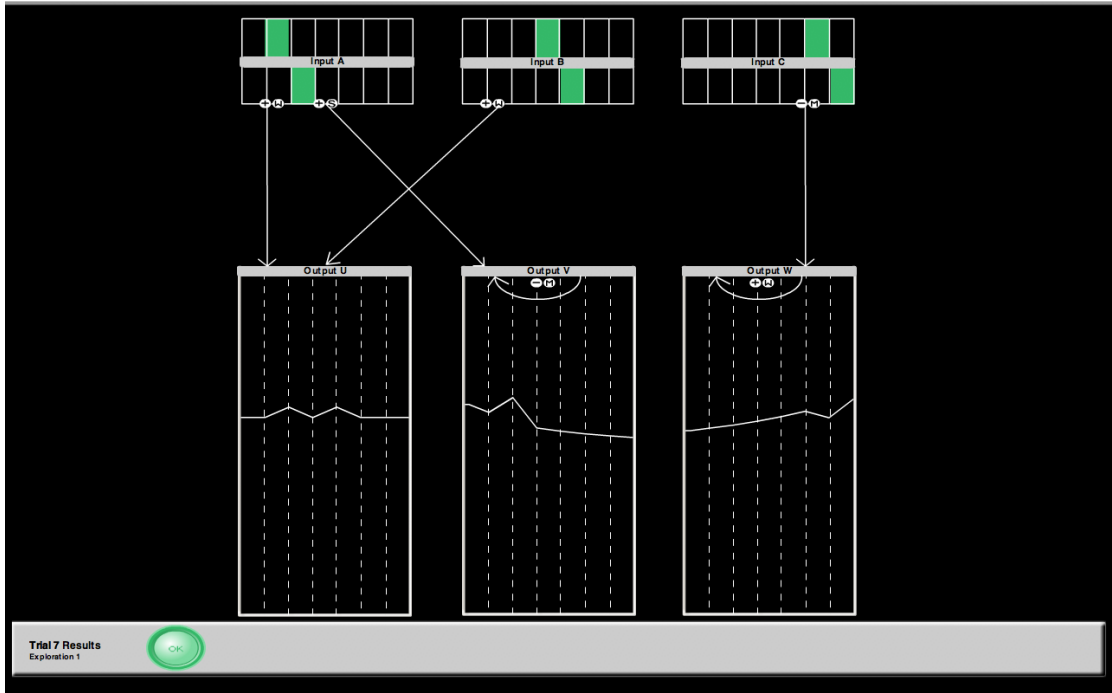
So I set C to maximum, U stayed the same, V decreased slightly, and W decreased. So I would expect V to decrease on its own anyway, so there's no effect there. And this time W decreased, so we found that C effects W.



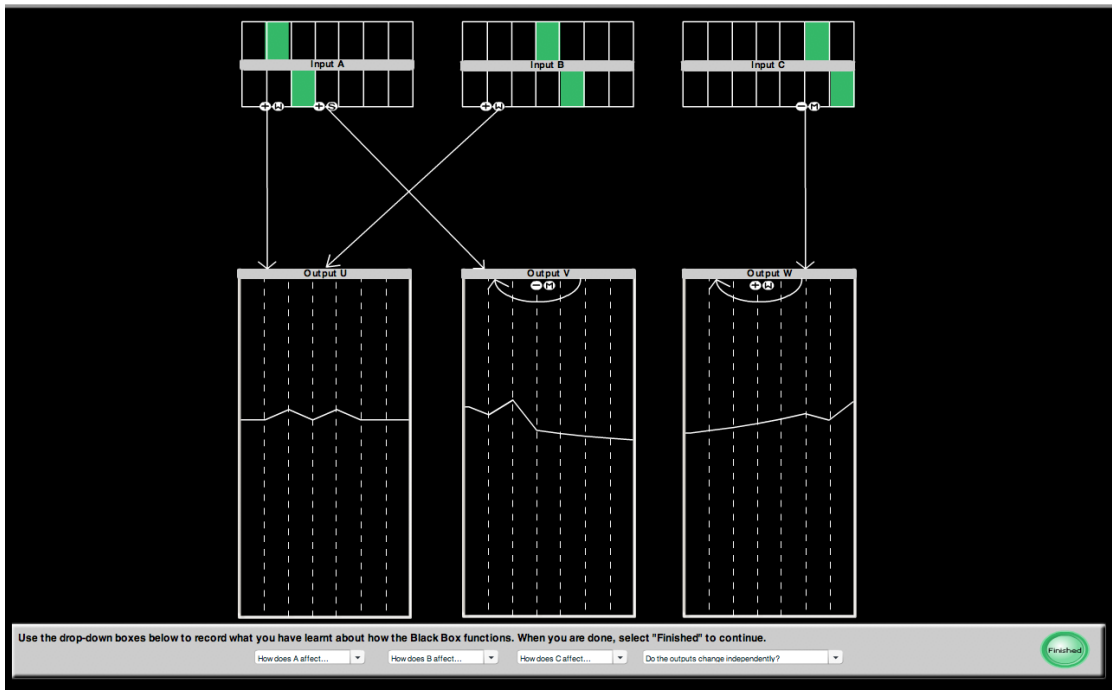
So let's record that result. We found that C affected W. Yes there was an effect. We increased C and W decreased, so that's a negative effect, and I'm going to say it's a medium effect.



OK, so looking at the effect of C again, setting A to zero, B to zero and decreasing C to minimum.



We decreased C to minimum. U stayed the same, V decreased slightly and W increased slightly. So C only effects W.

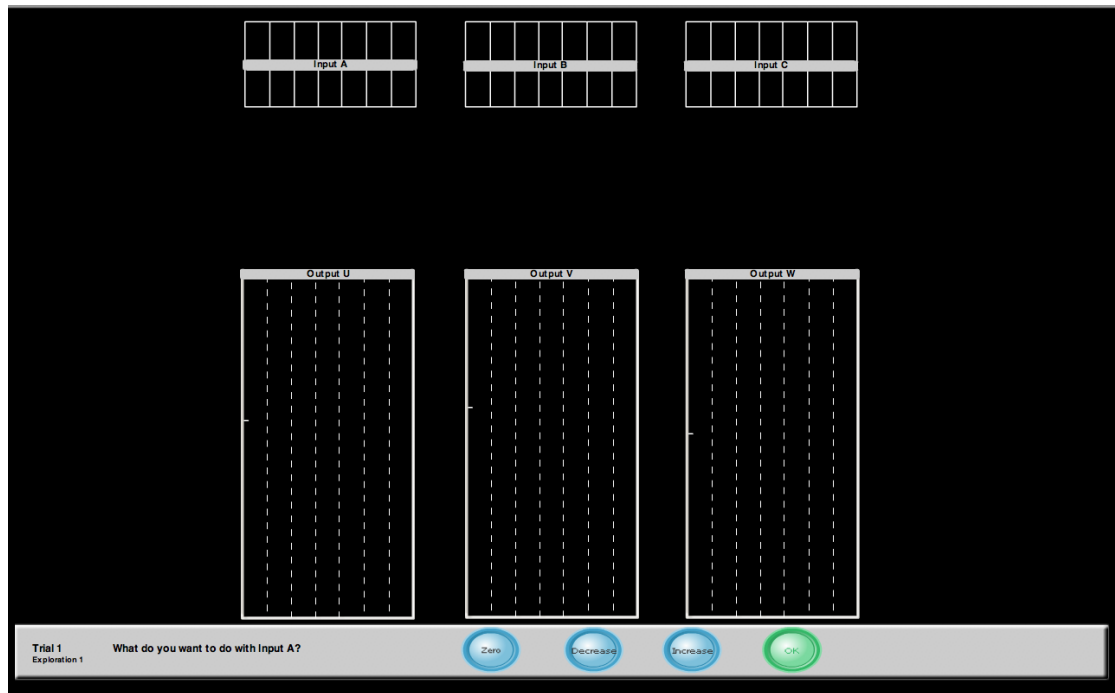


So I've already recorded that C has a positive, medium effect on W.

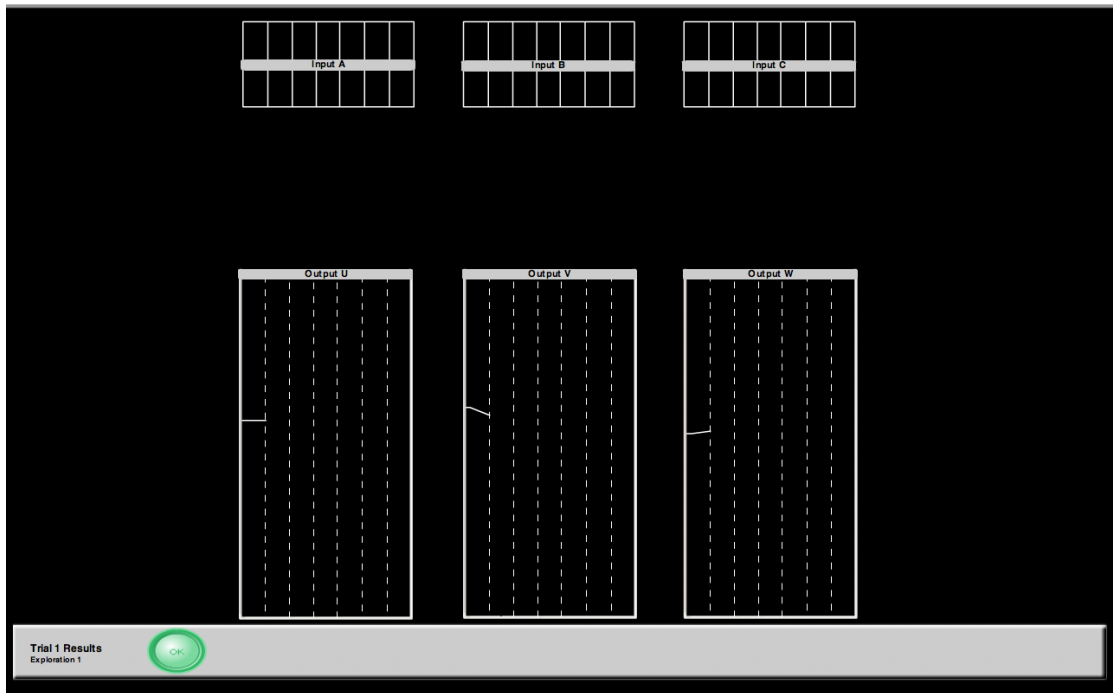
APPENDIX B

INSTRUCTIONS IN THE PARTIAL INFORMATION CONDITION

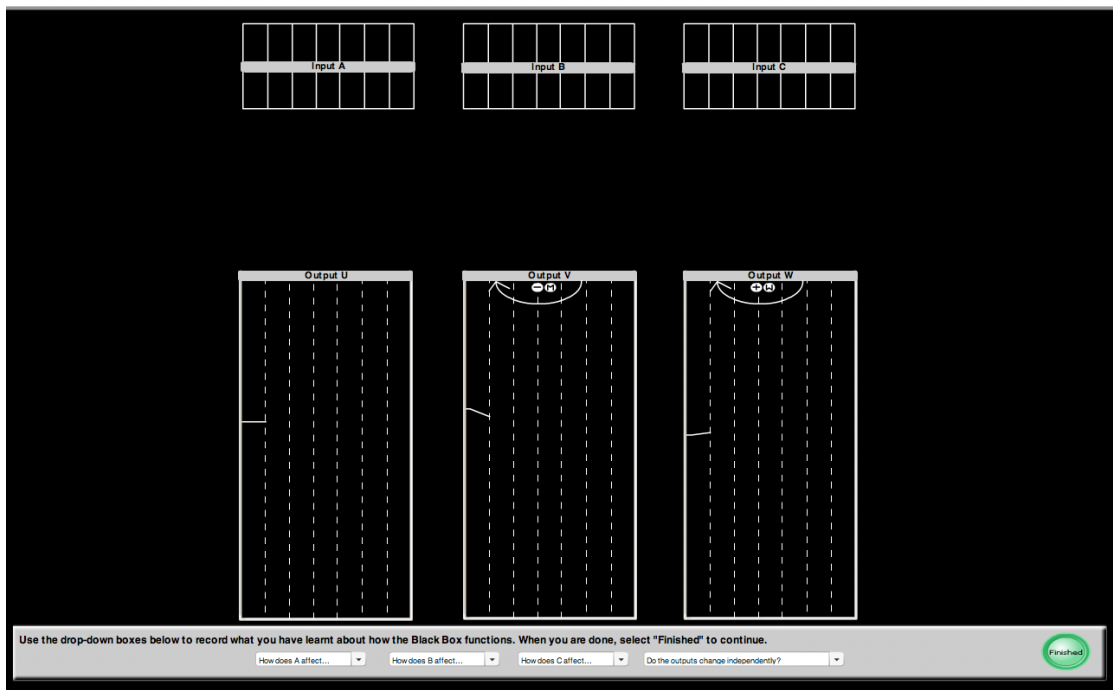
The following screenshots and text show the instructions given in the partial information condition (Study 2 and 3). The text was presented as a concurrent narration.



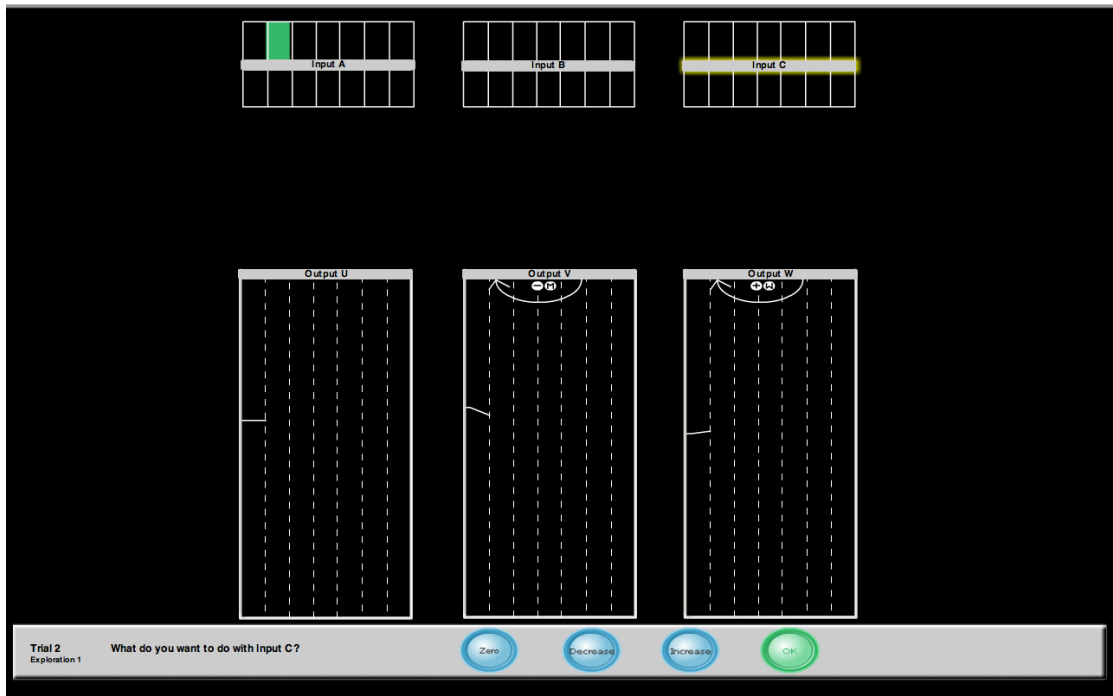
On the first trial, we should determine whether any of the outputs change independently of the inputs. So we should set each input to zero, and see if the outputs change.



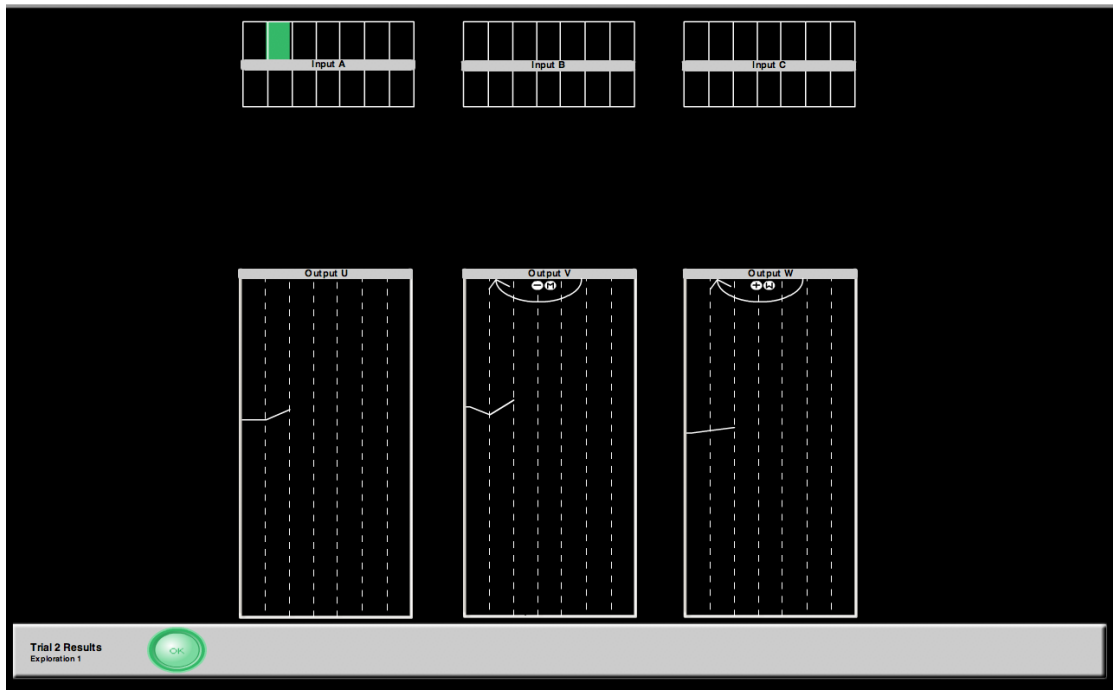
So we set each of the inputs to zero, and U didn't change, V decreased, and W increased. So it looks like V and W change on their own over time.



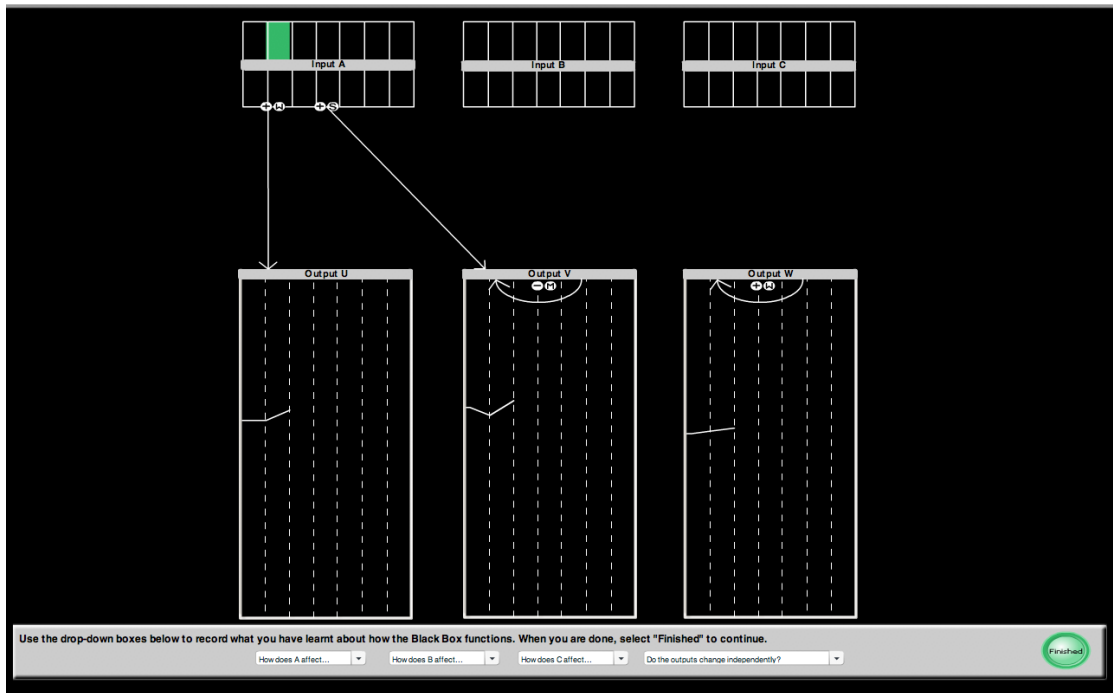
Let's record that result. So we found that V and W changed. So output V, yes it changes. It decreased, so that's a negative effect, and I'm going to say that it's a medium effect. Output W, yes it changed, it increased, so that's a positive effect, and it's quite small, so I'm going to say it's a weak effect.



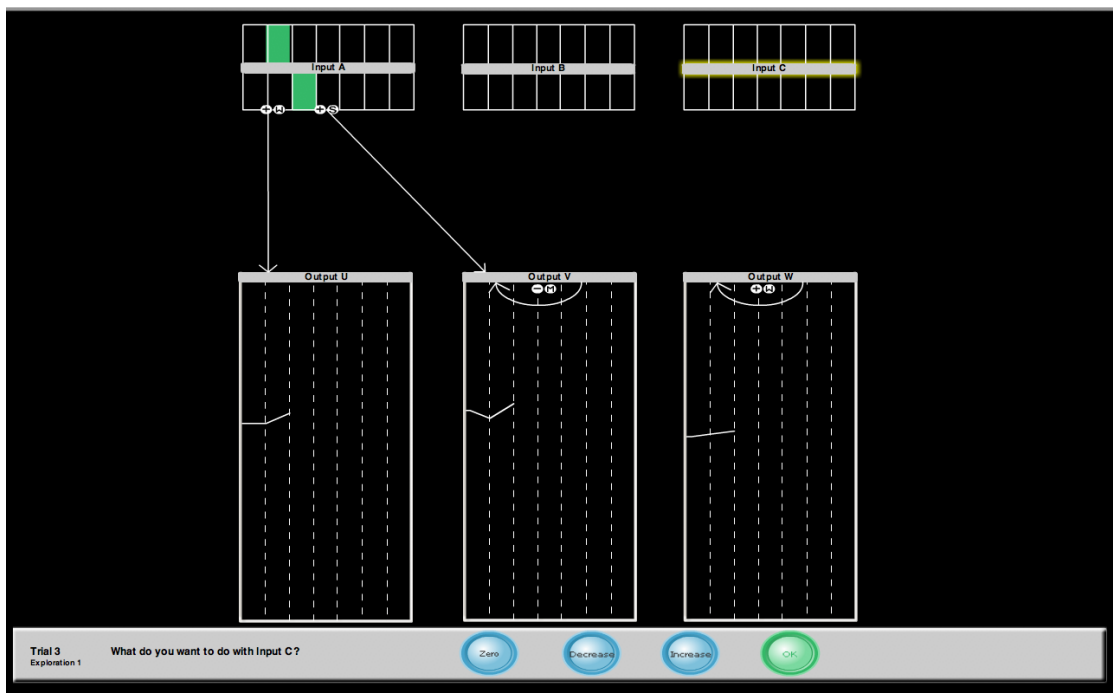
This time I'm going to look at the effect of input A on the outputs. So I'm going to increase A to maximum, and set each of the other inputs to zero.



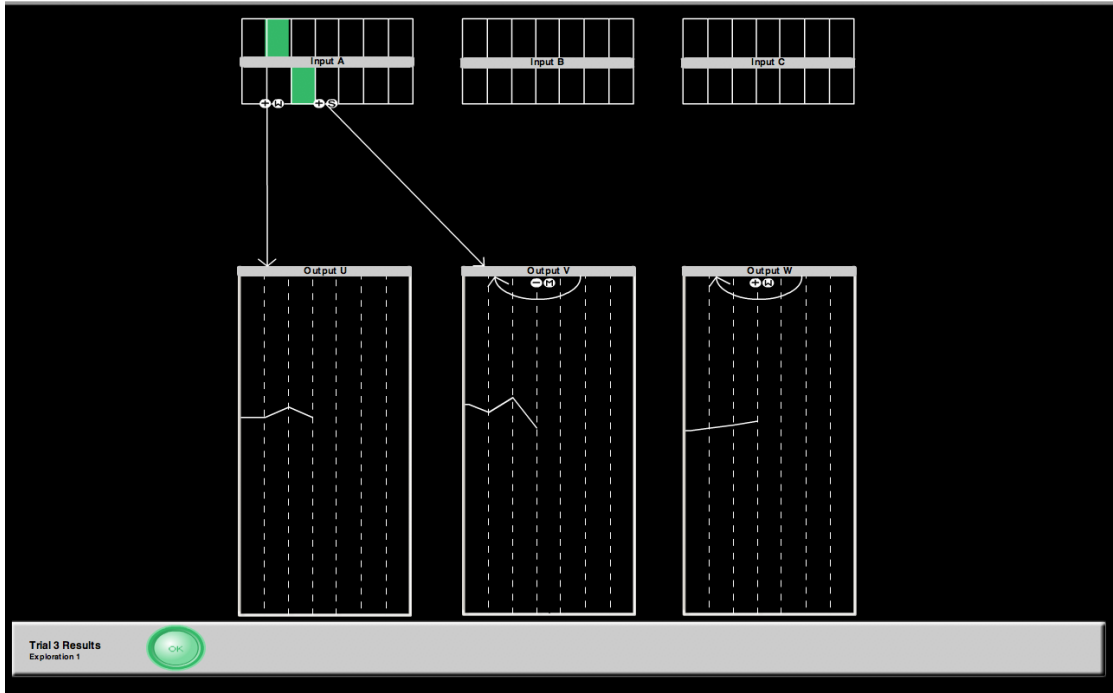
So I increased A to maximum, and U increased, V increased and W increased. Now we would expect W to increase on its own anyway, so it has not effect there. So we can say A effects U and V.



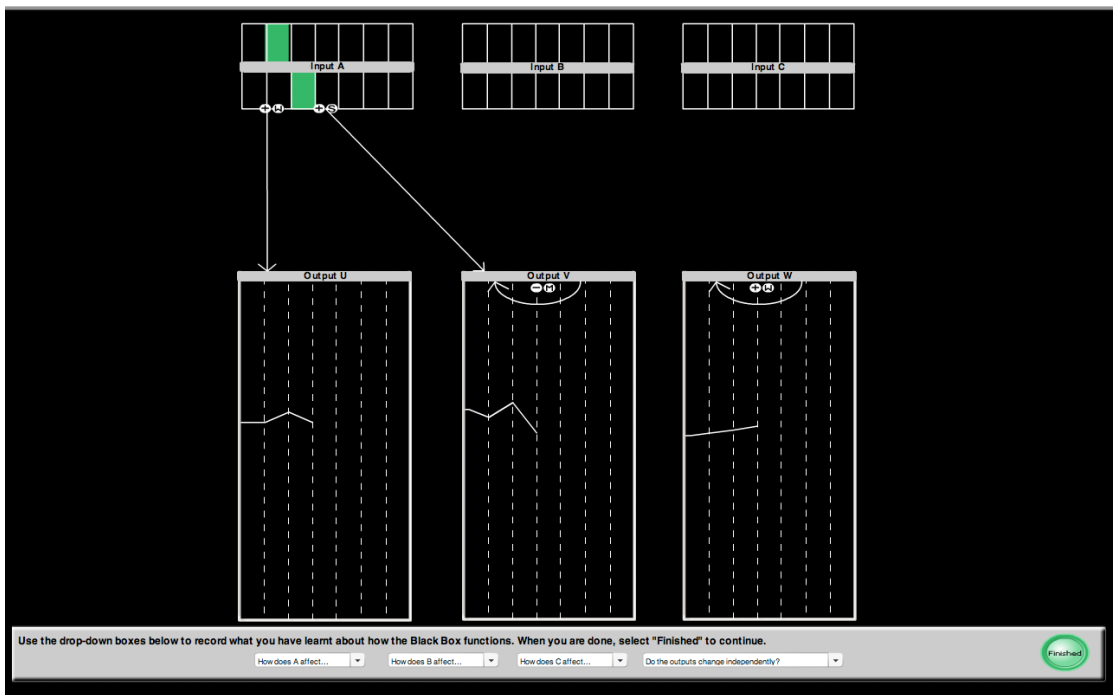
Let's record that result. So we found that input A effects output U. Yes there's an effect, increasing A increases U, so it's a positive effect. And I'm going to say it's a weak effect. We also found that A effects V. Increasing A increases V, yes there's an effect, it's a positive effect, and I'm going to say it's a strong effect because it has to work against the independent decline in V.



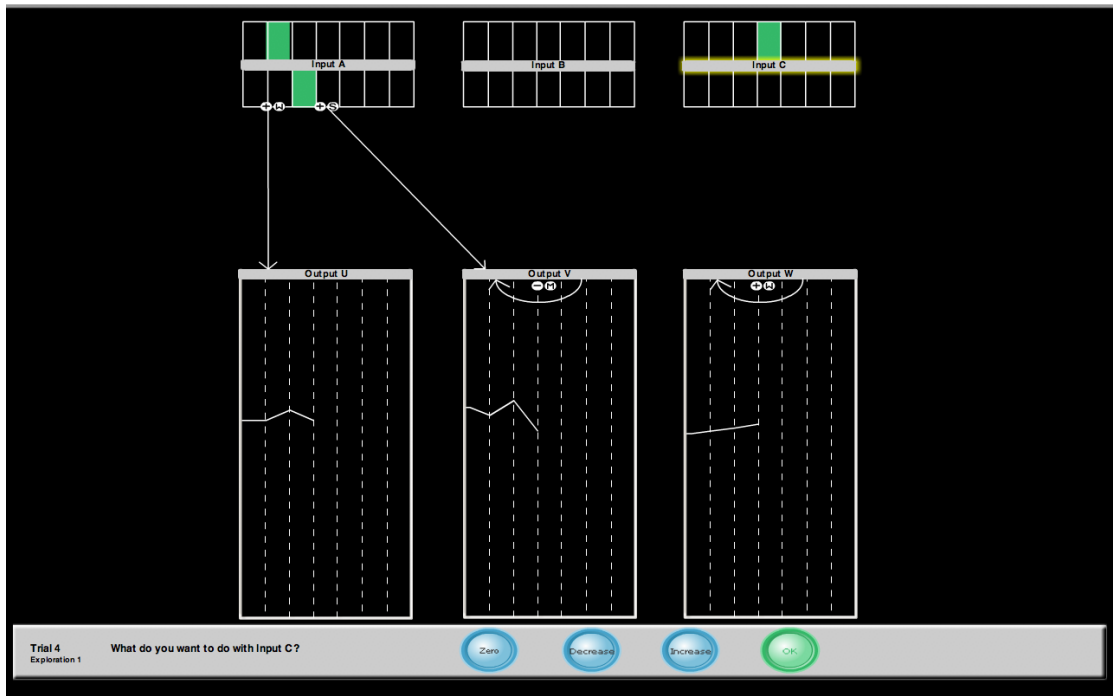
So this time I'm going to decrease A to minimum and set the inputs to zero.



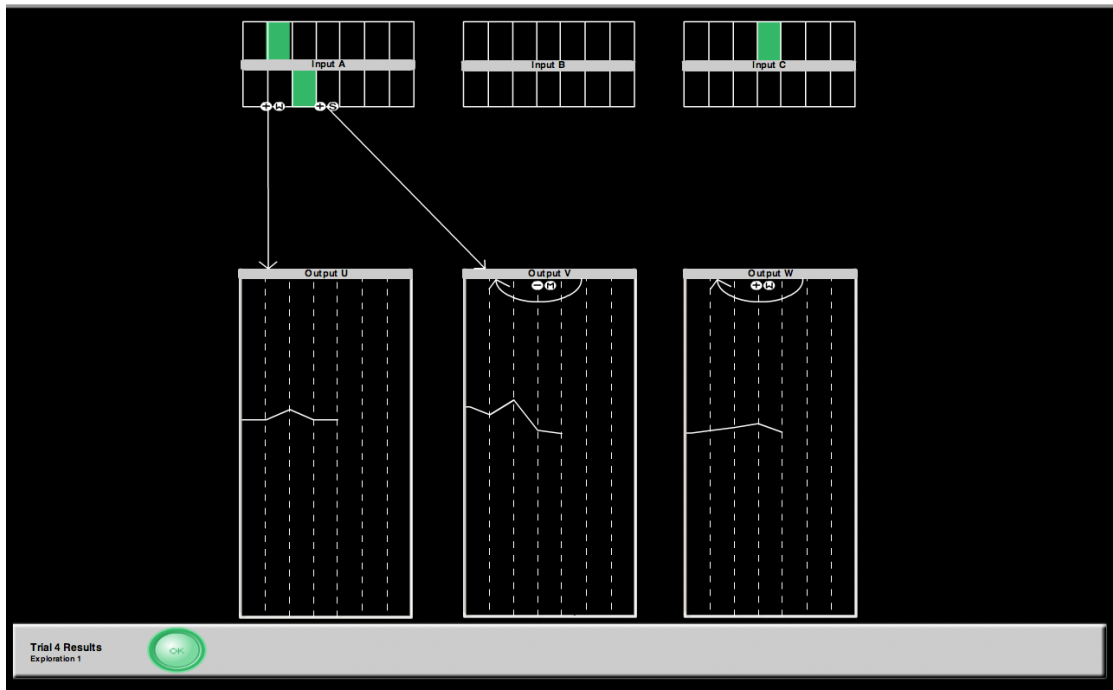
So we decreased A, and U decreased, V decreased and W increased slightly.



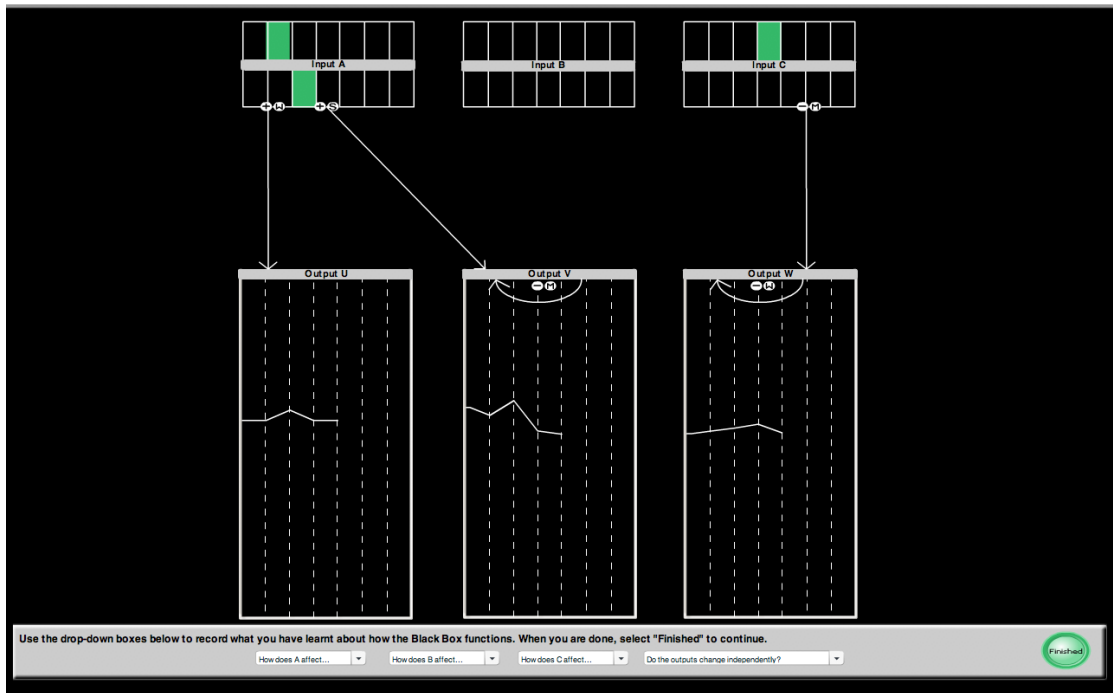
So I've already recorded that A affects U and V.



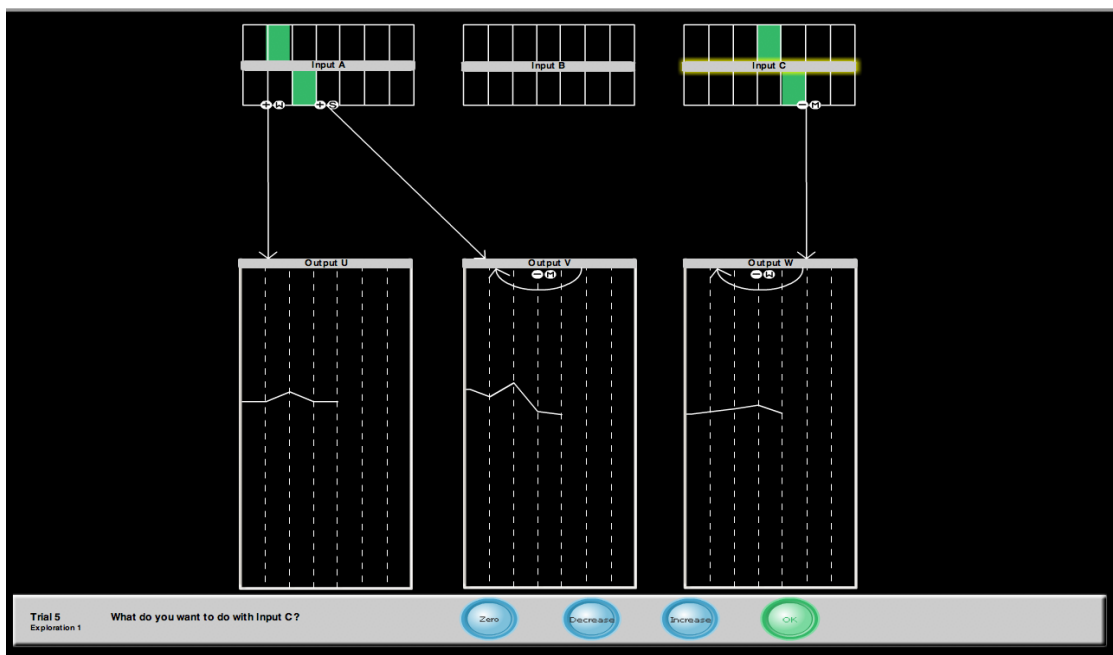
This time I'm going to look at the effect of C on the outputs. So I'm going to set A to zero, B to zero and increase C to maximum.



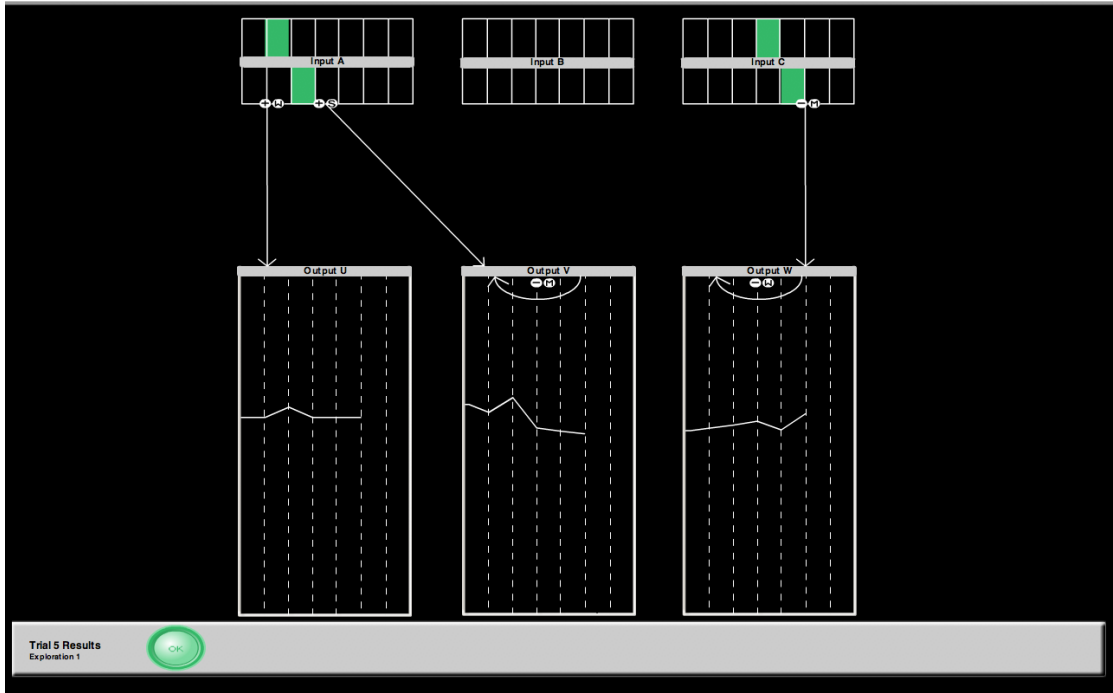
So I set C to maximum, U stayed the same, V decreased slightly, and W decreased. So I would expect V to decrease on its own anyway, so there's no effect there. And this time W decreased, so we found that C effects W.



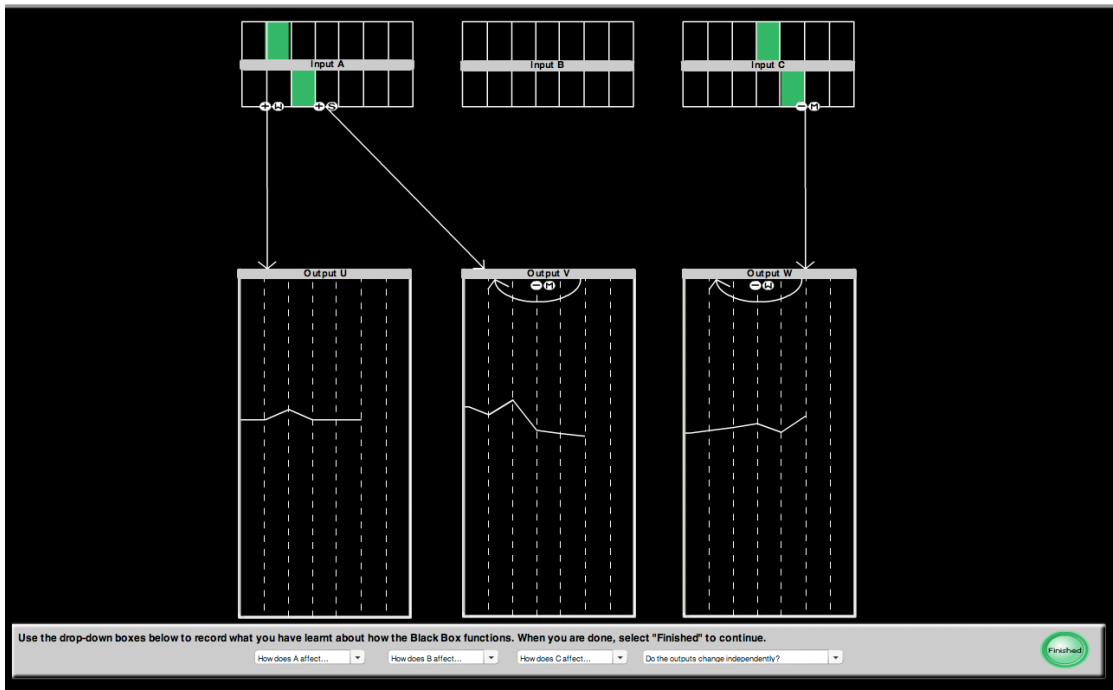
So let's record that result. We found that C effects W. Yes there was an effect. We increased C and W decreased, so that's a negative effect, and I'm going to say it's a medium effect.



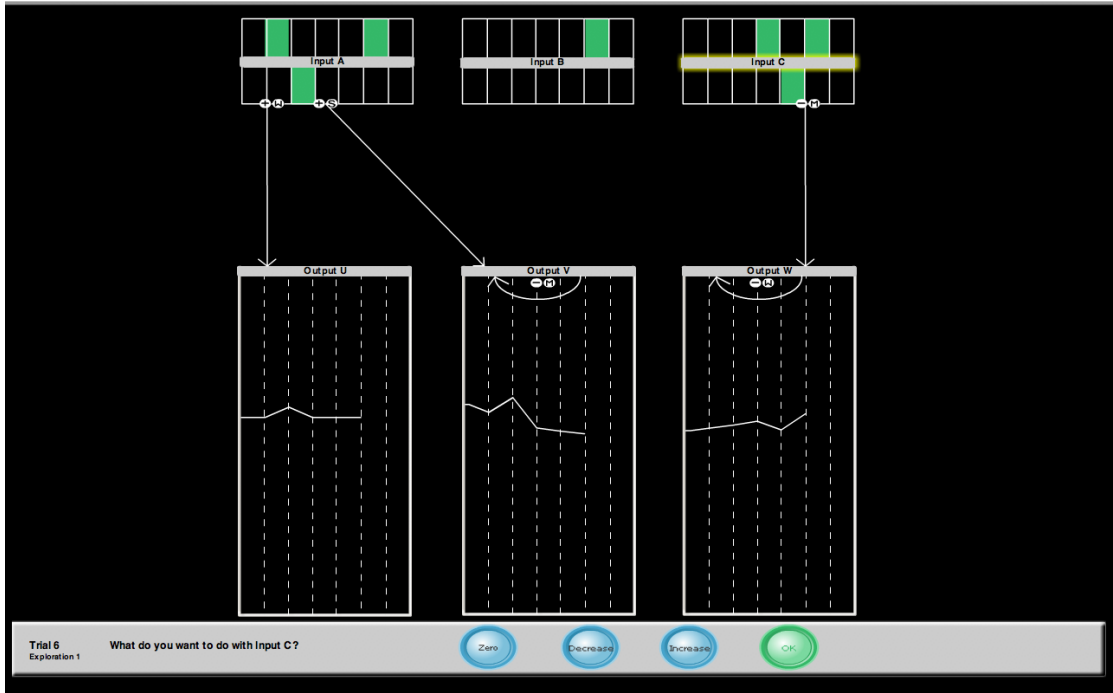
OK, so looking at the effect of C again, setting A to zero, B to zero and decreasing C to minimum.



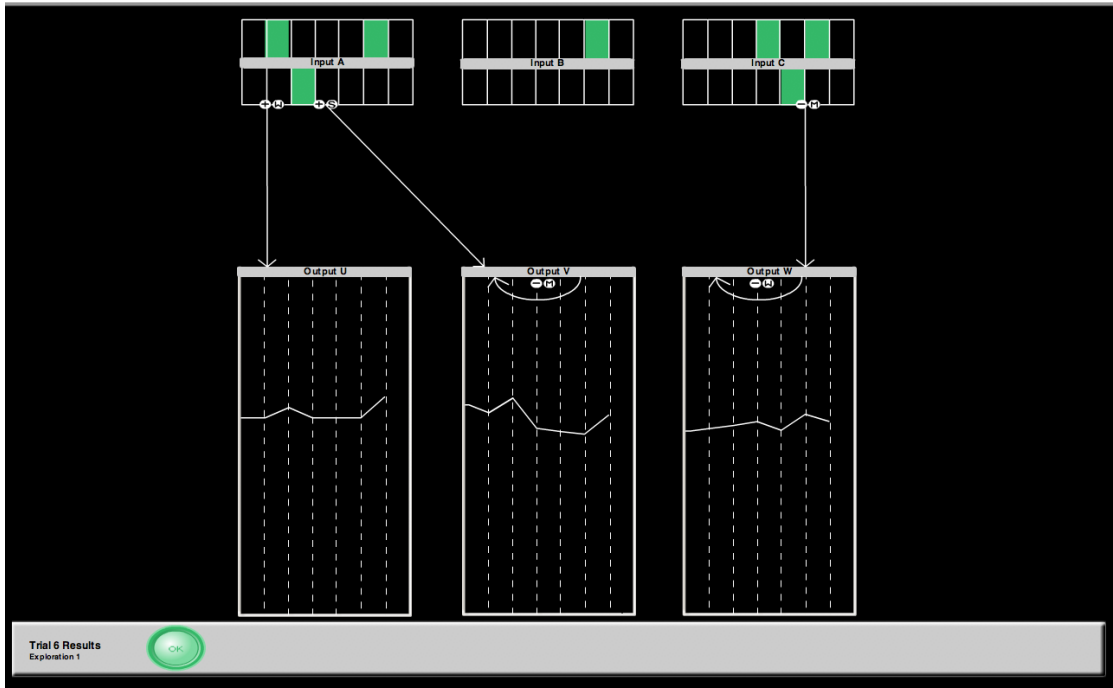
We decreased C to minimum. U stayed the same, V decreased slightly and W increased slightly. So C only affects W.



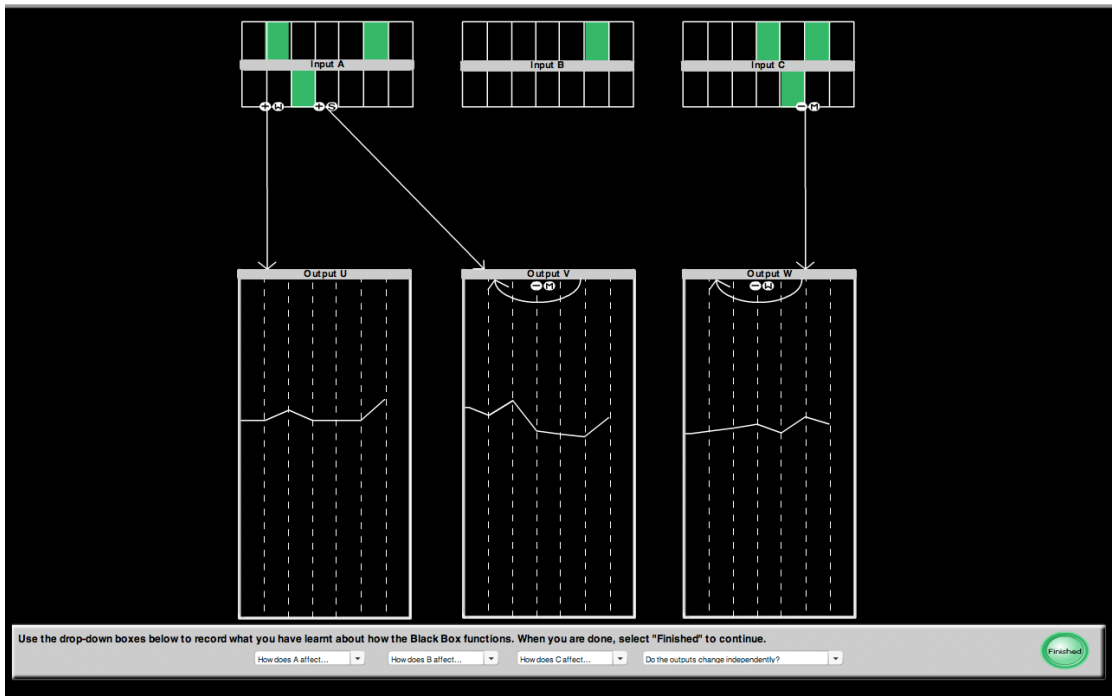
So I've already recorded that C has a medium, negative effect on W.



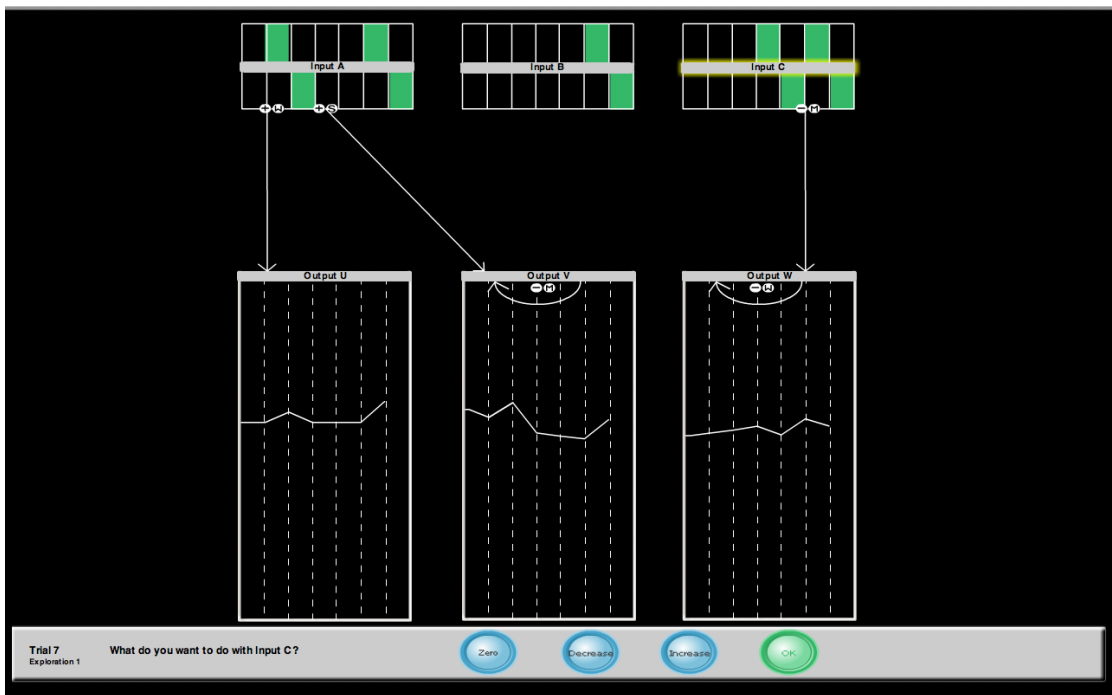
This time I'm going to increase all the inputs to maximum. So increasing A to maximum, B to maximum and C to maximum.



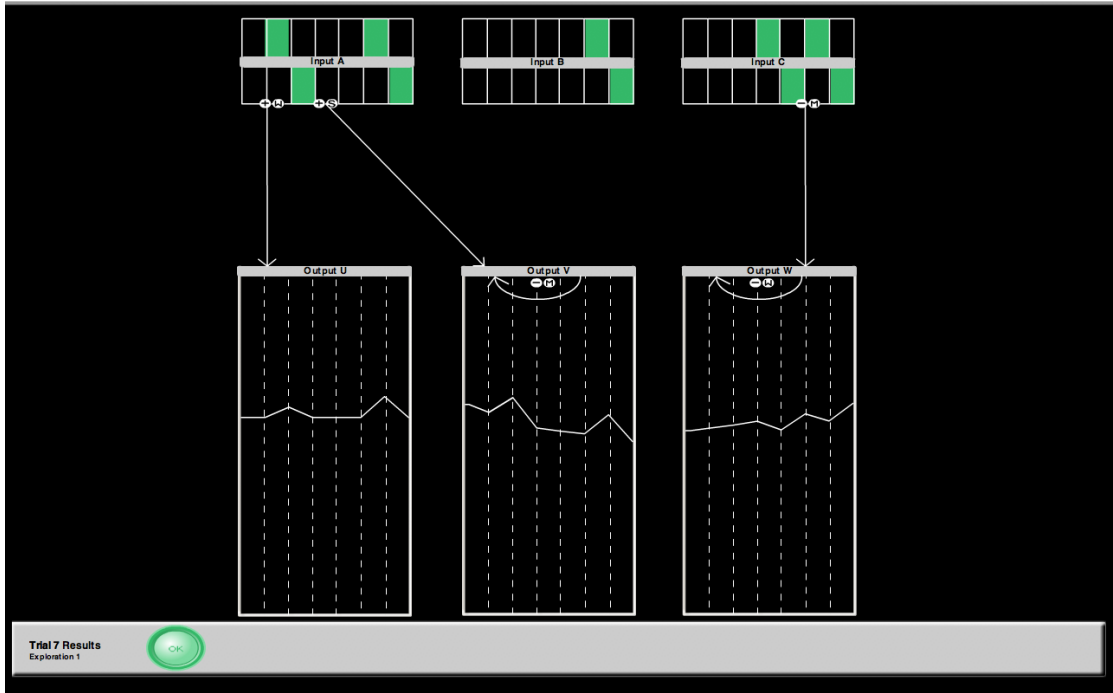
So all the inputs are set at maximum. U increases, V increases and W decreases.



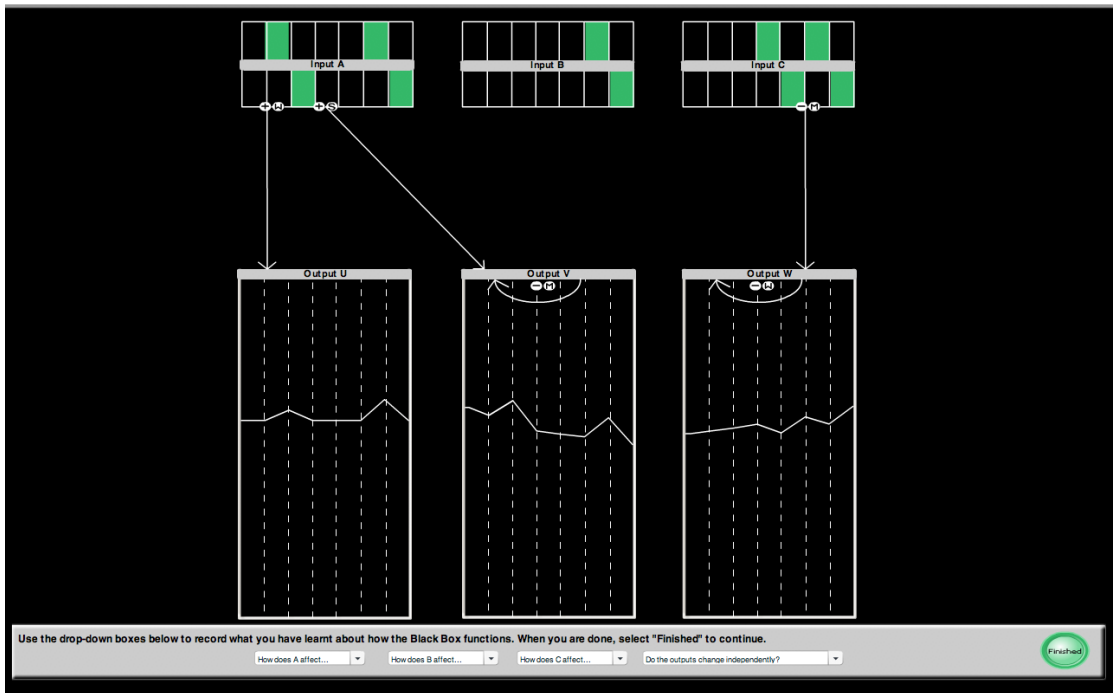
So nothing new to record on that trial.



This time I'm going to increase all the inputs to minimum. So increasing A to minimum, B to minimum and C to minimum.



So all the inputs are set at minimum. U decreased, V decreased and W decreased.

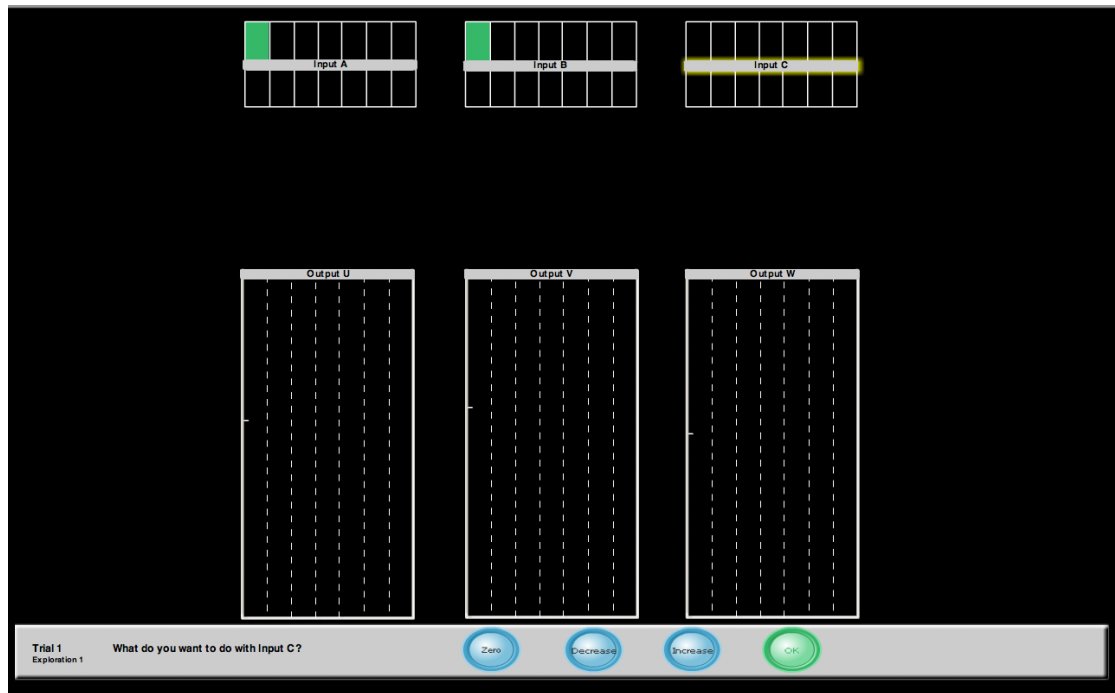


So nothing new to record on that trial.

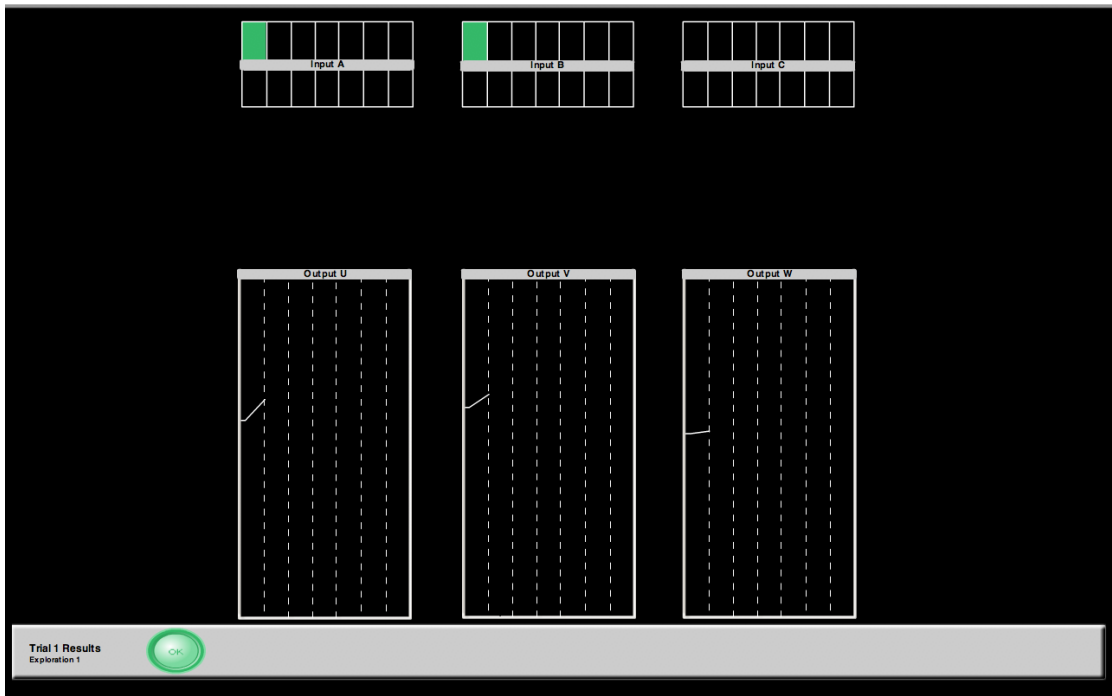
APPENDIX C

INSTRUCTIONS IN THE NO INFORMATION CONDITION

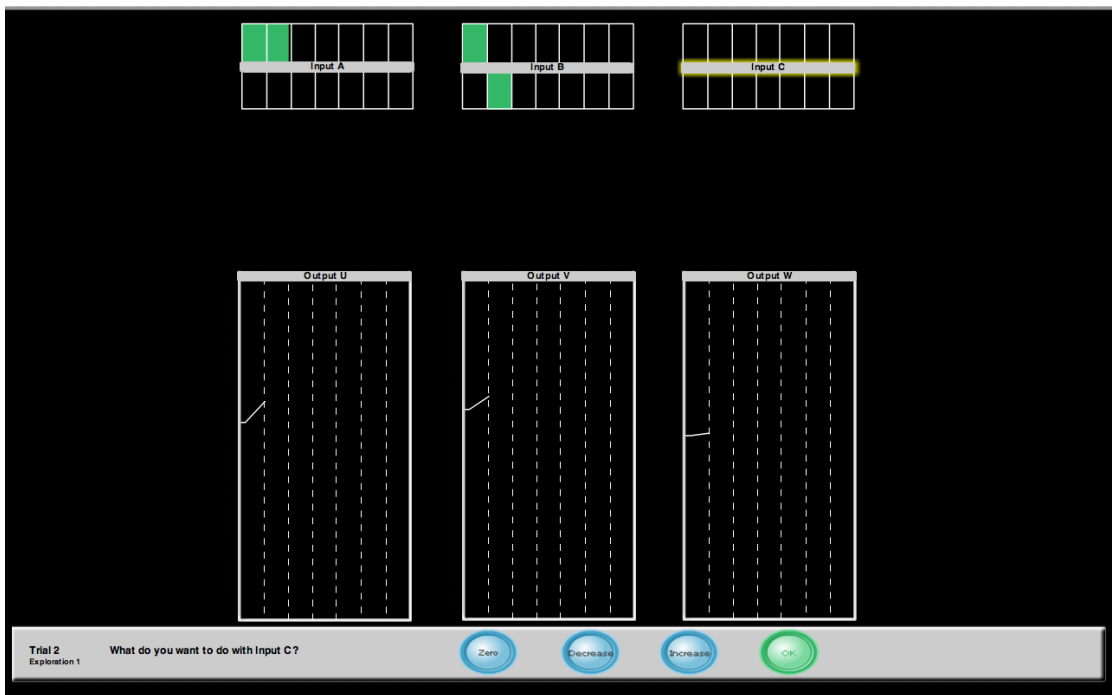
The following screenshots and text show the instructions given in the no information condition (Study 2 and 3). The text was presented as a concurrent narration.



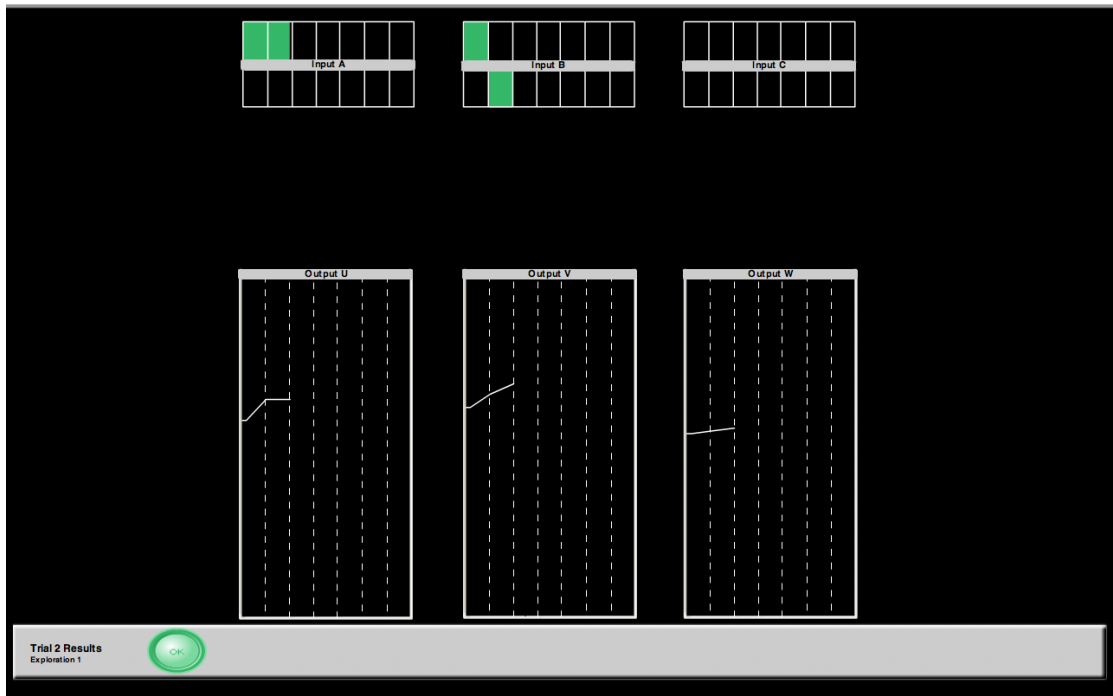
OK on the first trial I'm going to look at the effect of input A and B on the outputs. So I'm going to increase A to maximum, increase B to maximum and set input C to zero.



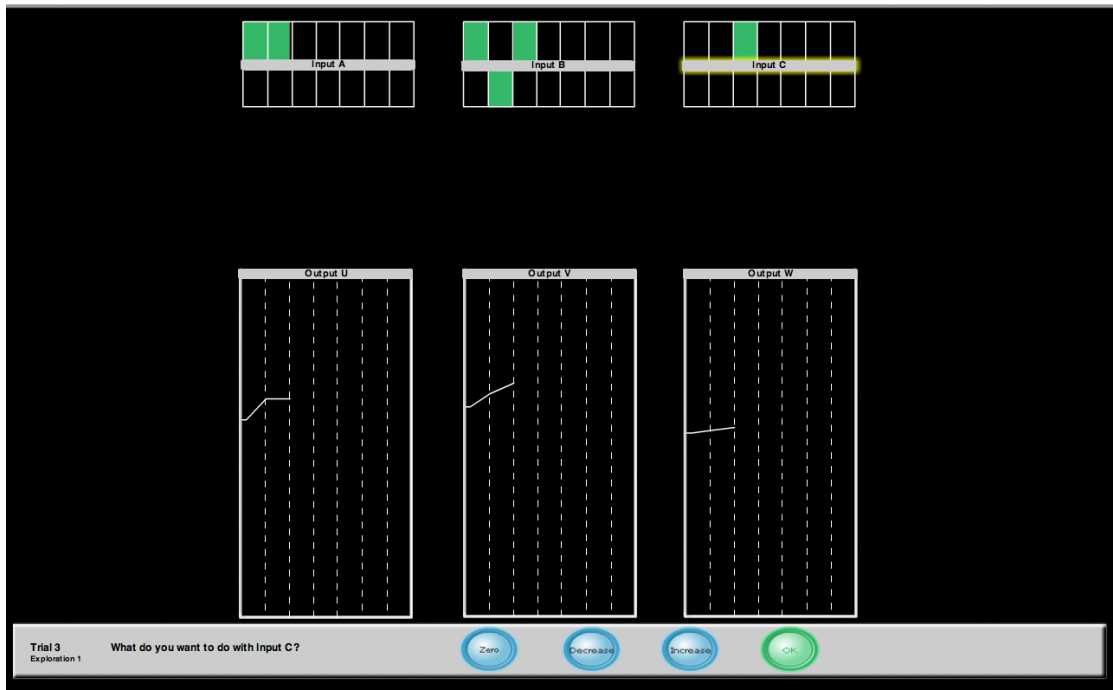
So we increased A and B, and U increased, V increased and W increased slightly.



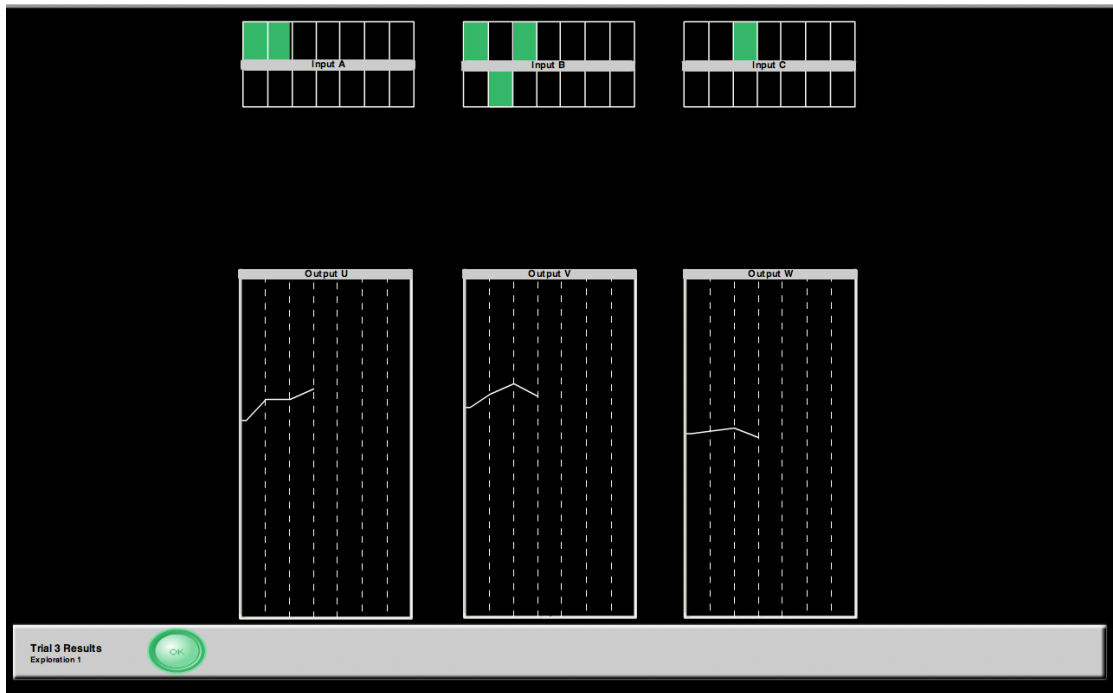
OK, I'm going to look at the effect of A and B on the outputs again. This time I'm going to increase A to maximum, decrease B to minimum, and set C to zero.



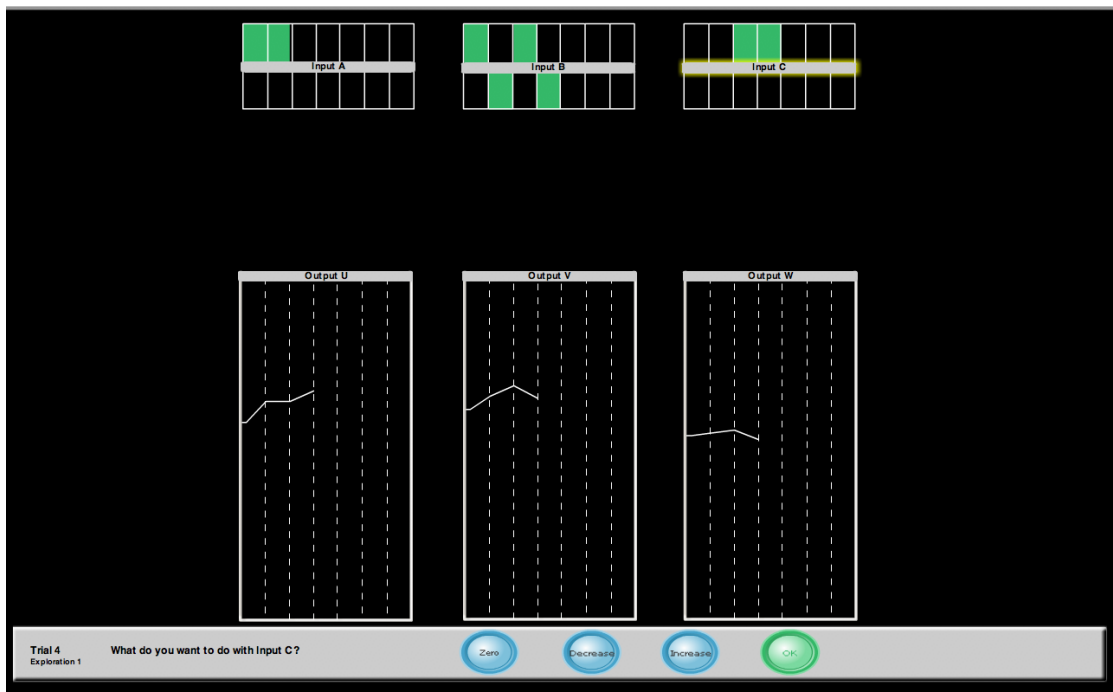
So we increased A and decreased B. U stayed the same. V increased and W increased slightly.



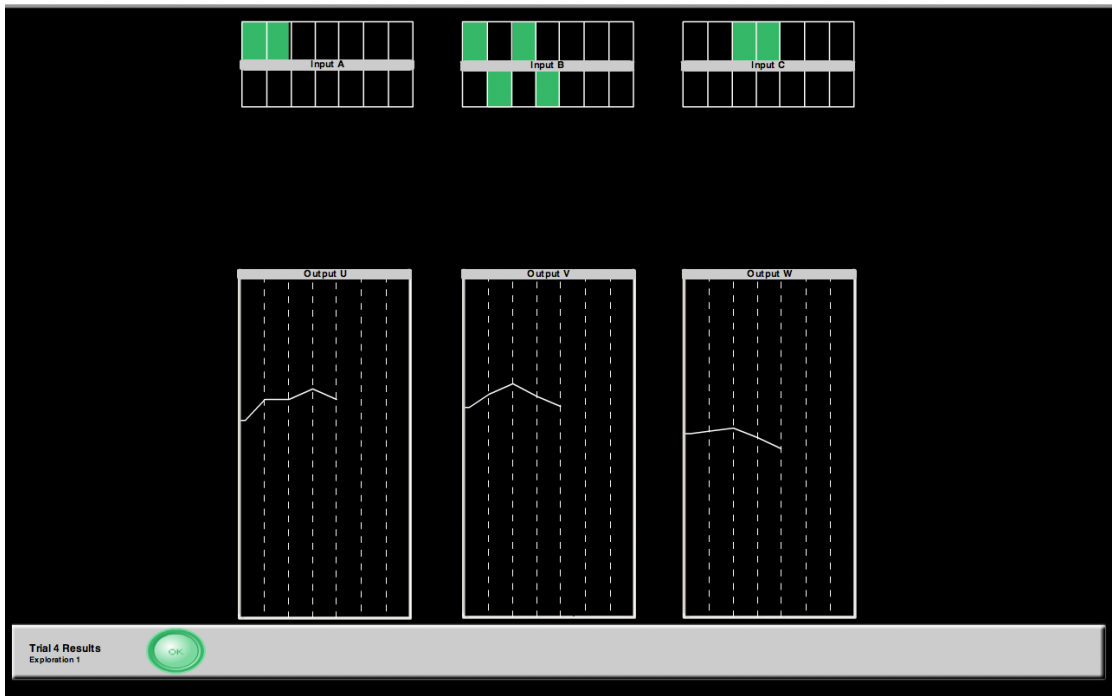
OK on the third trial I'm going to look at the effect of B and C on the outputs. So I'm going to set A at zero, increase B to maximum, and increase C to maximum.



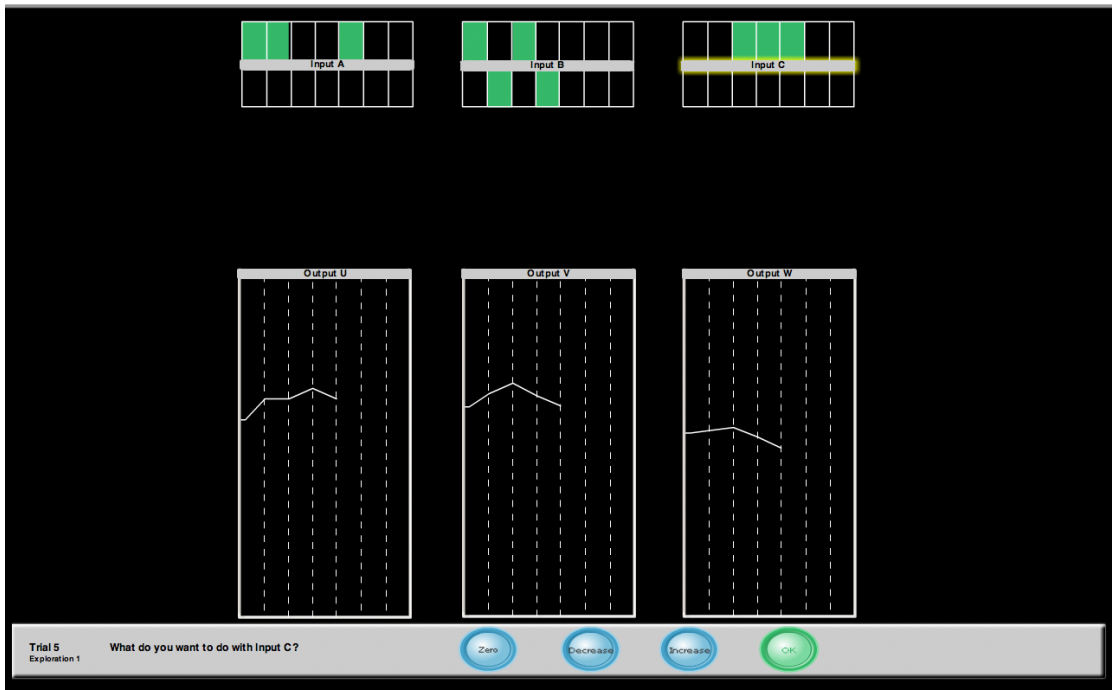
So we increased B and C and U increased, V decreased and W decreased.



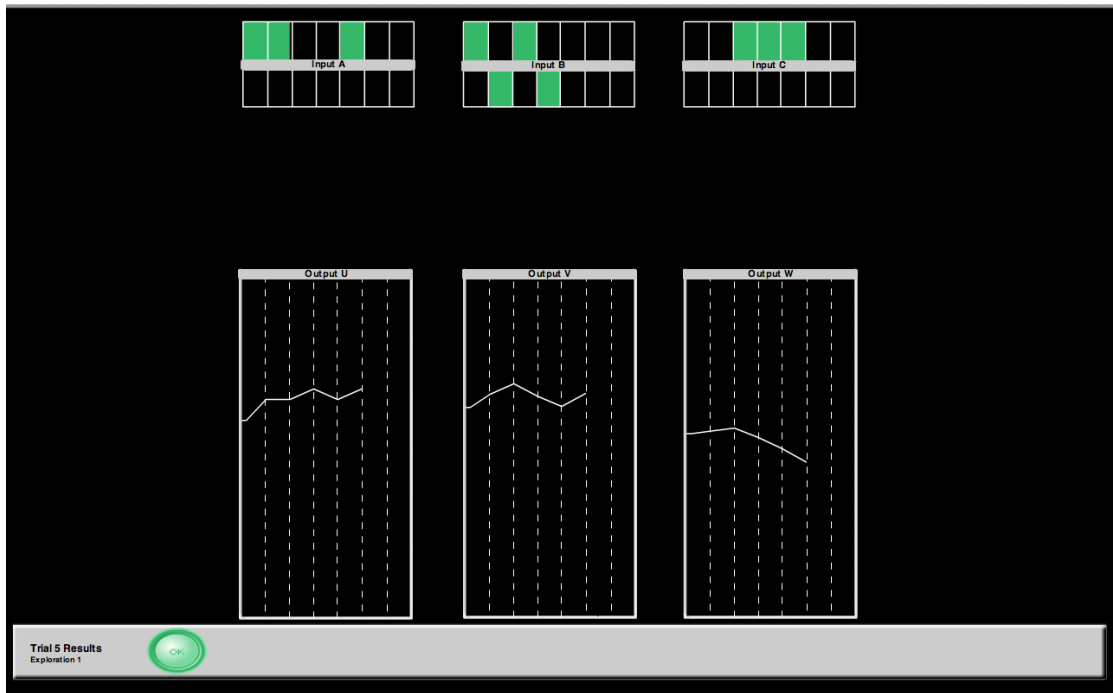
OK on the fourth trial I'm going to look at the effect of B and C on the outputs again. So I'm going to set A at zero, decrease B to minimum, and increase C to maximum.



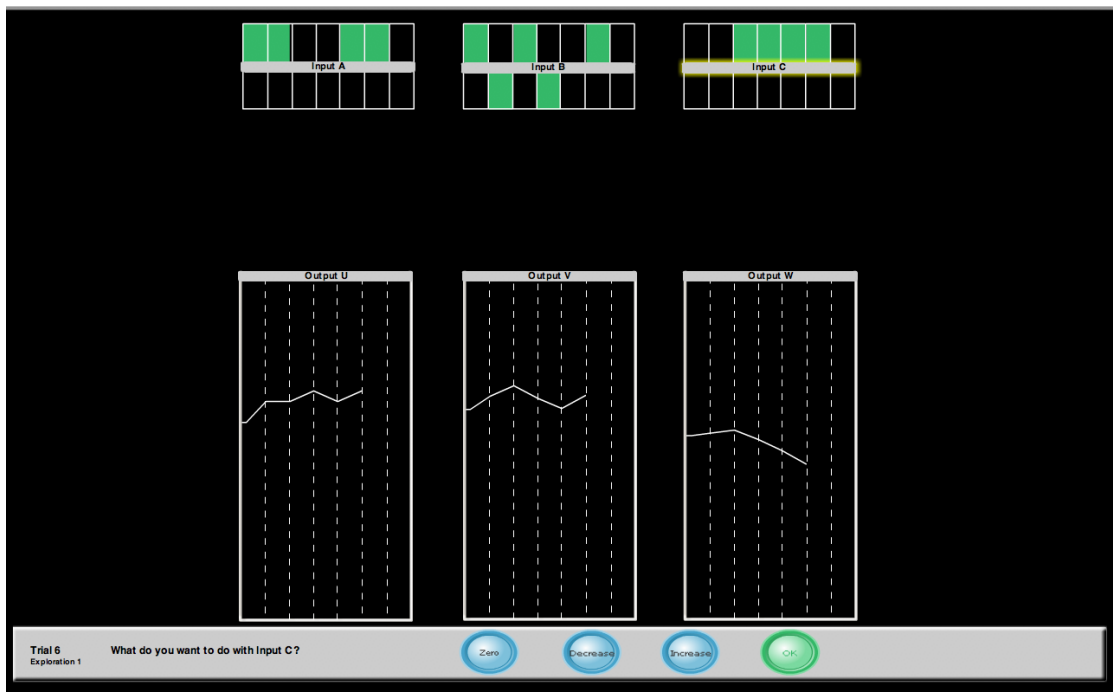
OK, so we decreased B and increased C and U decreased, V decreased and W decreased.



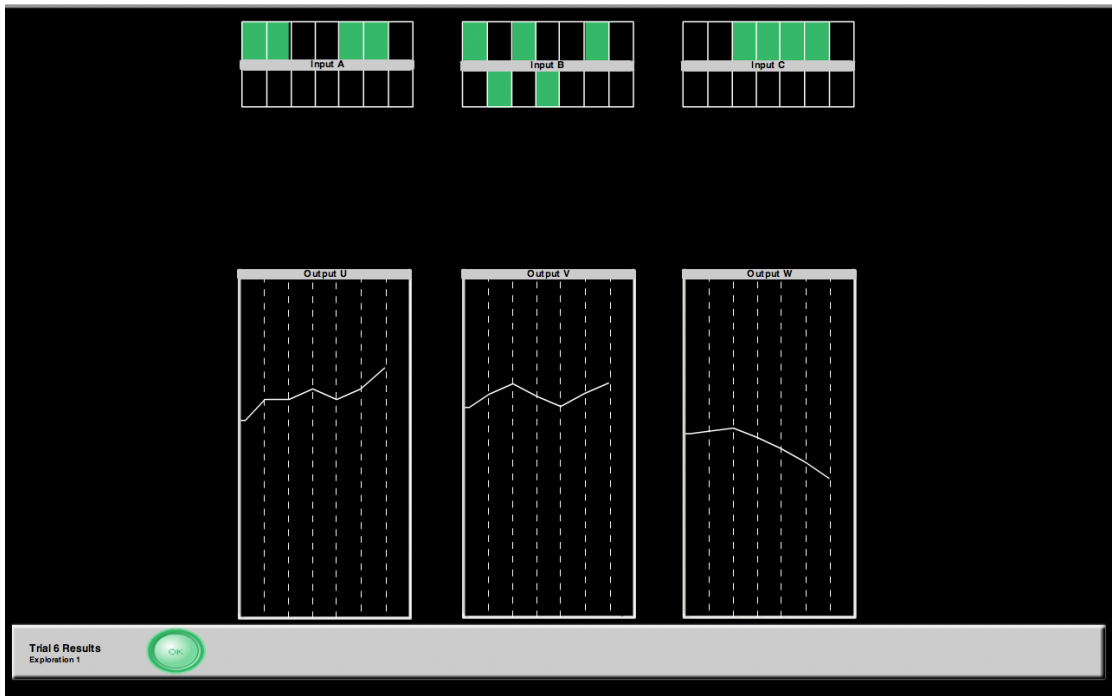
So on the fifth trial I'm going to look at the effect of A and C. So I'm going to increase A to maximum. Set B at zero, and increase C to maximum.



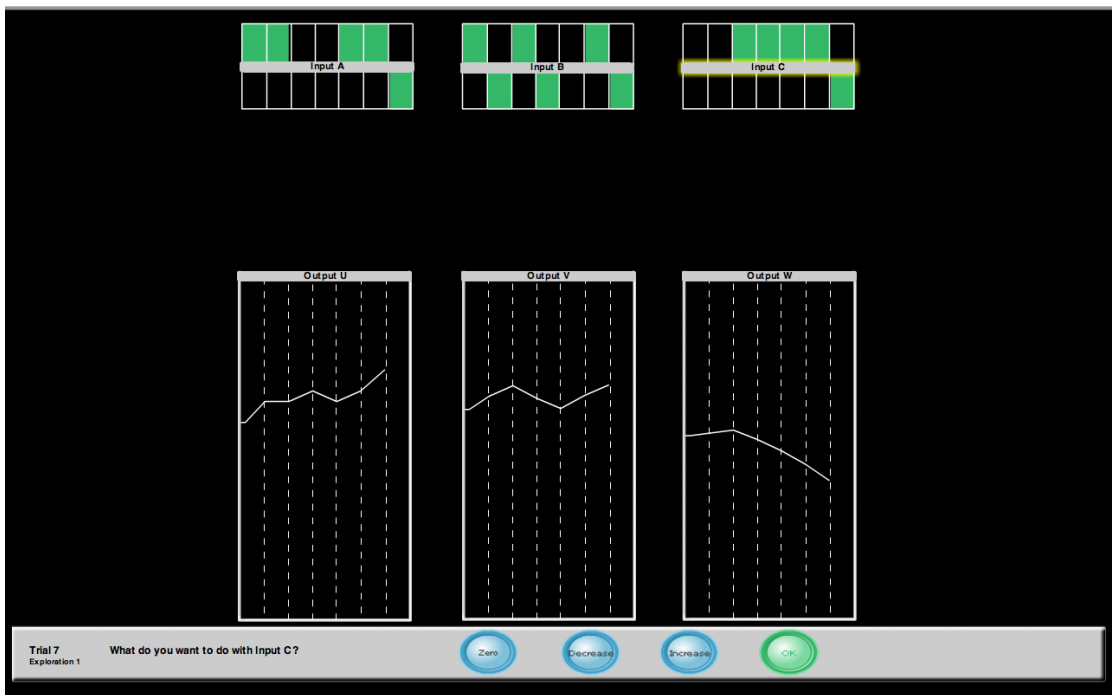
So we increased A and C, and U increased, V increased and W decreased.



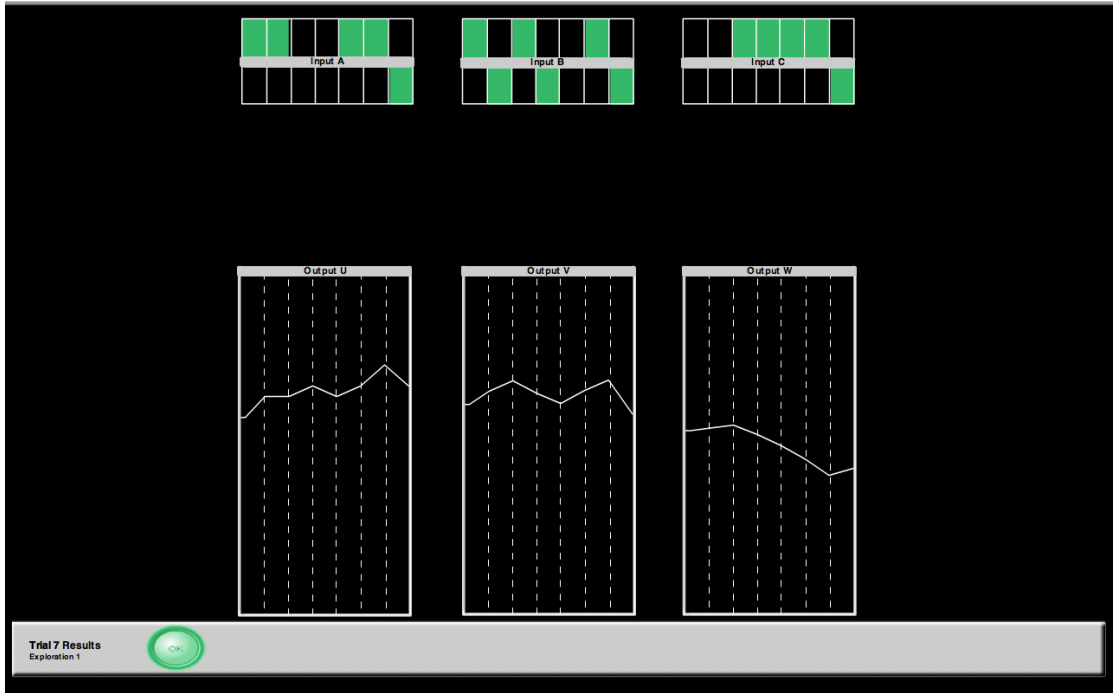
OK, so on the sixth trial I'm going to look at the effect of all the inputs on the outputs. So I'm going to increase A to maximum, increase B to maximum and increase C to maximum.



So we increased A, B and C, and U increased, V increased and W decreased.



OK, so this time I'm going to decrease all the inputs to minimum. So decreasing A to minimum, decreasing B to minimum, and decreasing C to minimum.



So I decreased A, B and C and U increased, V decreased, and W increased slightly.

APPENDIX D

THE CONCEPTUALISATION AND MEASUREMENT OF TASK COMPLEXITY

D.1 Introduction

Task complexity has been discussed as a potential moderator variable on performance in a wide range of applied and experimental settings, such as medical diagnosis (Xiao, Hunter, MacKenzie & Jefferies, 1996; Park & Jung, 2007), the maintenance of electronic systems (Wohl, 1982; Rouse & Rouse, 1979; Rasmussen & Lind, 1981), instructional design (Speier, 2006; Stahl, Pieschl & Bromme, 2006; Beckmann, 2010), goal-setting (Wood, Mento & Locke, 1987), problem solving (Kotovsky, Hayes & Simon, 1985; Sweller, 1976; Haerem & Rau, 2007; Vakkari, 1999; Spilsbury, Stankov & Roberts, 1990; Halford et al., 2005) and decision-making (Hu, Huhmann & Hyman, 2007; Olshavsky, 1979; Payne, 1976; Earley, 1985; Simon & Tversky, 1992; Mone & Shalley, 1995; Paquette & Kida, 1988). Currently, task complexity is conceptualised, measured, and operationalised in multiple ways, even within particular fields of research. Firstly, in this chapter we will review this diverse literature in order to clarify what is meant by task complexity. Secondly, the various frameworks that are available to estimate task complexity will be evaluated in terms of whether they are likely to provide accurate estimates of task performance. Overall, the purpose of this chapter is to derive a more comprehensive account of complexity than is currently given in the complex problem solving literature.

D.2 Dimensions of Task Complexity

Within the literature, task complexity is often treated as a multi-dimensional construct. The primary dimension, which is consistently referred to in all fields of research, captures notions of intricacy and inter-connectedness (Spilsbury, Stankov & Roberts, 1990). Under this view, typically, task complexity is defined as the number of components of the task or mental operations involved when solving the task. Different researchers tend to emphasise either the sheer number of procedures and/or task elements as the most important contributor to task complexity (e.g. Park & Jung, 2007), or the extent to which these procedures and/or task elements must be

integrated in a single step (e.g. Halford et al., 1998a) or a combination of the two (e.g. Wood, 1986; Campbell, 1988) in order to reach a particular goal.

Dynamics (Wood, 1986; Woods, 1988), risk (Woods, 1988), task structure (Bonner, 1994; Byström & Järvelin, 1995; Haerem & Rau, 2007) and uncertainty (Schroder, Driver & Streufert, 1967; Wood, 1986; Woods, 1988; Xiao, Hunter, Mackenzie & Jeffries, 1996; Osman, 2010) have been proposed as additional dimensions of task complexity by various researchers. Dynamics refers to whether the outcomes of the task change in response to the decisions made by the problem solver and independently over time. It may also refer to whether there are feedback delays in the task (Wood, 1986; Woods, 1988). It could be argued that the extent to which a task is dynamic essentially captures a temporal aspect of the primary dimension, relating to intricacy and inter-connectedness, and so, these two factors should not be considered to be independent.

Risk refers to the possibility that incorrectly enacting the procedures needed to reach a desired solution can have large costs in certain situations (Woods, 1988). For example, the negative impact would be considerable if workers fail to correctly perform the emergency procedures in a nuclear power plant. Woods (1988) argues that risk is typically low in most problem solving and decision-making tasks that are performed in a laboratory setting. Nevertheless, it could be argued that the evaluation of risk by the task-doer essentially increases the number of mental activities, or number of variables that need to be taken into account when performing the task, and thus is likely captured by the primary dimension of task complexity relating to intricacy and inter-connectedness.

Task structure usually refers to the level of specification of what is to be done in the task. This includes whether the goals are ill or well defined and whether the procedures needed to solve the task are known to the task-doer (Bonner, 1994; Byström & Järvelin, 1995). However, Haerem and Rau (2007) propose an alternative interpretation of task structure. In an extension of Chomsky's (1957) work on linguistics to an analysis of cognitive tasks they refer to task structure as the level of processing that is required in the task. In deep structure tasks, solution requires a search through the processes or procedures related to the task. In surface structure tasks, solution requires a search through the inputs and outputs of the task. Mixed

structure tasks require a combination of the two. As with risk, it seems reasonable to expect that task structure, under either definition, is also likely impact upon the intricacy and inter-connectedness of the task.

Uncertainty can be defined as either as an objective property of the task or a subjective experience of the problem solver. From an objective perspective, Wood (1986) argues that uncertainty exists when the relationships between the variables in the task change, or the relationships between actions and outcomes change. Similarly, but from a subjective perspective, Osman (2010) defines uncertainty as the extent to which the problem solver is confident in predicting the outcome of events and that an action will lead to an expected outcome. Alternatively, Guastello (2004) argues that uncertainty exists when “the decision maker knows what the possible outcomes would be, knows the probabilities associated with each outcomes, but is compelled to guess which of the outcomes will actually take place... [or when] the decision maker knows what the possible outcomes would be, but does not know the probabilities that each will take place.” The common theme across these conceptualizations of uncertainty appears to be reliability of the task, where subjective uncertainty occurs as a result of objective uncertainty. Of course, a situation can also be imagined in which the objective uncertainty of the task is low, while the subjective uncertainty of the task is high due to lack of information about the task (Osman, 2010).

Certainly, it seems likely that intricacy and interconnectedness, risk, task structure and uncertainty might all contribute to overall task difficulty. However, it could be argued that the inclusion of each of these factors within the construct of task complexity might reflect a somewhat misplaced desire to unite every aspect of task difficulty under the banner of task complexity. Such a strategy, however, is likely to be counter-productive because a) it is unclear whether each of these facets are independent, and thus whether they should be measured separately, and b) the construct may become too broad to posit as a clear explanation for task performance. Nevertheless, this problem is largely ignored in the literature as most of the frameworks that have been developed to estimate task complexity focus almost exclusively on the primary facet that relates to the intricacy and inter-connectedness of the task.

In the following sections, the frameworks that have been developed to estimate the intricacy and inter-connectedness of task (i.e. complexity) will be reviewed, with the general organising principle as to whether task complexity is treated as an objective property of the task or an interaction between psychological and task characteristics. Objective treatments of task complexity claim to consider the task independently of any task doer, although it will be argued that this assumption is not met in most objective formulations of task complexity. Interactive perspectives consider the task in relation to the psychological processes that must be performed in order to reach a particular goal.

D.3 A starting point: The relationship between difficulty and task complexity

Firstly, however, it is necessary to differentiate between task complexity and difficulty. It should be noted that difficulty is not synonymous with task complexity, although sometimes the two terms are sometimes used interchangeably (e.g. Kluge, 2008a; Woods, 1988; Byström & Järvelin, 1995; Earley, 1985; Sweller, 1976). Tasks can become more difficult, without necessarily becoming more complex (Rouse & Rouse, 1979; Wood, 1986; Campbell, 1988; Bonner, 1994; Spilsbury, Stankov & Roberts, 1990; Quesada, Kintsch & Gomez, 2005; Halford et al., 1998a). For example, Rouse and Rouse (1979) note that “While lifting weights, running ten miles and monitoring a radar screen for long periods of time may be difficult tasks, they are not necessarily complex tasks. A task can be difficult because it is a physically hard thing to do, or hard to put up with (i.e. boring) or perhaps because it is complex. Thus, complexity should be viewed as a possible component of difficulty” (p.724). Essentially, complexity is not the only factor that influences difficulty; therefore, it is not appropriate to use the terms synonymously.

However, given our conceptualization of complexity as the intricacy and inter-connectedness of the task, a more complex task should also be one that is more difficult. If hypothesised changes in complexity do not result in empirical changes in difficulty, then we might have reason to suspect that our operationalisation of complexity may be lacking. In the following sections, alternative conceptualizations of task complexity will be evaluated with respect to this assumption.

D.4 Complexity as an objective property of the task

D.4.1 Computational Complexity

Computational complexity theory examines the complexity of effective computations for solving problems that consist of a set of instances (i.e. goals for performance) and a set of solutions to these problem instances. Computational complexity can be expressed as a function of the amount of time needed to solve a particular instance of a problem using the most efficient computation or algorithm. Similar calculations can be made in terms of the computer resources needed to solve an instance with a particular algorithm. Problems that cannot be solved within a certain time frame are referred to as intractable problems. Such problems are assigned the highest level of computational complexity, and hence difficulty (Gary & Johnson, 1979; Wagner & Wechsung, 1986).

In the early stages of theoretical development, it was argued that computational complexity might provide an index of task complexity for human problem solvers, and hence task difficulty (Gary & Johnson, 1979; Rouse & Rouse, 1979; Wagner & Wechsung, 1986). However, from a theoretical perspective, the concept of computational complexity seems inadequate for this purpose, for at least three reasons. Firstly, computational complexity essentially provides an estimate of how difficult it is to apply an algorithm, and ignores the difficulty of algorithm selection. This is problematic because many tasks require human problem solvers to identify *and* apply solutions, and it is well known that solution identification accounts for a large proportion of task difficulty. To cite an extreme example, the effort to create an algorithm that can solve the game of checkers (8 x 8 draughts), in the sense that perfect play on both sides leads to a draw, started in 1989 and was only recently completed in 2007 (Schaeffer, Burch, Björnsson, Kishimoto, Müller, Lake, Lu & Sutphen, 2007). Secondly, it has been argued that computational complexity is limited in its application to human problem solving because it is only relevant for algorithms that guarantee a solution to every instance (i.e. every possible goal state) of a problem. Human problem solvers, on the other hand, tend to search for solutions that are relevant to a particular goal state, unless instructed otherwise (Brattico, 2008). Thirdly, in mathematical proofs "...not only is it necessary to make explicit the changes an action brings about, it is also necessary to make explicit the

things that do not change, and for most actions this will be a great deal” (Shanahan & Baars, 2005, p.158). It seems doubtful that at a day to day level human problem solvers explicitly consider the trivial non-effects of their actions (the frame problem, McCarthy & Hayes, 1969; Shanahan & Baars, 2005). Thus, the problems that are relevant to the formulation of mathematical proofs or their application are quite unlike those that are relevant to human problem solvers in many cognitive tasks. As such, it seems unacceptable to appropriate the concept of computational complexity to estimate task complexity.

Secondarily, the empirical evidence shows that humans can solve many problems that are considered to be intractable with little difficulty, and conversely, many computationally “simple” problems are quite difficult for human problem solvers (Halford et al., 1998a; Quesada, Kintsch & Gomez, 2005). For example, one of the most frequently cited examples of an intractable problem is the traveling salesman problem. In this problem, subjects are given a list of European cities and their distances, and the task is to find the shortest possible route between them, in which each city is visited only once. As the number of possible routes increases exponentially with the number of cities to be visited, determining the shortest route is usually impossible for a computer. Human subjects, on the other hand, solve this problem, and its variants, relatively easily (Pizlo, Stefanov, Saalweachter, Li, Haxhimusa & Kropatsch, 2006). In contrast, some logic problems such as Wason’s (1966; Wason & Shapiro, 1971) selection task, that have a low computational complexity, are extremely difficult for many human problem solvers (Quesada, Kintsch & Gomez, 2005). These results suggest that computational complexity does not provide a useful metric for operationalising task complexity in the context of human problem solvers, or for explaining why one task is more complex than another.

D.4.2 Size of the problem space

Newell and Simon (1972) proposed that problems, and by extension tasks, could be represented by a “problem space”, which consists of an initial state, a goal state, all the possible states of the problem, and a number of operators that contain constraints for moving between states. This idea has been utilised in a study conducted by Buchner and Funke (1993) to estimate the complexity of a complex

problem in which they argued that “...complexity is determined by the number of states that a device can be in, and by the number of interventions with different consequences possible for a given system state. The number of different interventions corresponds to the number of potential user decisions given a particular state” (p.89). Similarly, in decision-making tasks, task complexity is often operationalised as the number of alternatives and attributes that must be considered in order to make a choice (Olshavsky, 1979; Payne, 1976; Simonson & Tversky, 1992; Kerstholt, 1992). This essentially reflects the size of the problem space that must be considered.

From the problem space approach has emerged the idea that the size of the problem space determines problem difficulty (Kotovsky, Hayes & Simon, 1985), and by extension, it has been proposed that the size of the problem space may provide a useful proxy for task complexity (Buchner & Funke, 1993; Schoppek, 2002; Quesada et al., 2005; Jonassen & Hung, 2008). However, a range of empirical evidence suggests that the size of the problem space is not a good indicator task difficulty. Firstly, Kotovsky, Hayes & Simon (1985) argue that problems with relatively small problem spaces can be extremely difficult to solve. They give the example of the Missionaries and Cannibals problem that has a problem space of 16 nodes and the Tower-of-Hanoi problem that has 27 nodes, and argue that both of these problems are quite difficult for problem solvers (see also Quesada, Kintsch & Gomez, 2005). Similarly, Buchner and Funke (1993) found that subjects were unable to successfully control a dynamic system that had only 9 possible states of the system variables. Most strikingly, subjects who controlled this system performed no better than subjects who controlled a system that had 46 possible states. These findings show that there is not a direct relationship between the size of the problem space and task difficulty, and hence that the size of the problem space is not a good indicator of task complexity.

Secondly, the cover story of the problem appears to have a strong effect on task difficulty, even though this has no impact on the objective size of the problem space. Findings show that problem isomorphs of the Tower-of-Hanoi with different cover stories but identical problem spaces have substantially different difficulties. These problem isomorphs can be categorised into two broad classes, according to whether the operators that constrain the transitions between states involve *moving* a variable,

or *changing* a variable. Problems that involve changing a variable take approximately twice as long to solve as those that involve moving a variable, although they have identical problem spaces (Hayes & Simon, 1974; Kotovsky et al., 1985). Similarly, as discussed previously, the difficulty of acquiring structural knowledge in dynamic systems tasks with different cover stories but identical structures, and thus problem spaces, can vary considerably (Hesse, 1982; Beckmann, 1994; Beckmann & Guthke, 1995; Lazonder, Wilhelm, & Hagemans, 2008; Lazonder, Wilhelm & Van Lieburg, 2009). Strong content effects have also been observed in numerous studies on decision-making in static tasks (see Goldstein & Weber, 1995 for a comprehensive review). This suggests that the content of the cover story has an influence on task difficulty, which is not accounted for by the size of the problem space.⁷

Thirdly, findings from a number of different areas of research have shown that how the task is visually represented has a strong effect on the difficulty of the task. For example, Stock and Watson (1984) found that accountants were more accurate in predicting the bond rating of a stock if it was represented schematically, rather than numerically. Similarly, in dynamic systems research, graphical presentations of the system variables have been shown to result in better control performance than numerical presentations (for reviews see Vessey, 1994; Speier, 2006). Again, these manipulations should have no effect on the size of the problem space, yet they result in significant effects on problem difficulty.

Finally, increasing the number of possible decisions open to the problem solver, and thus the size of the problem space, does not necessarily hinder performance. In decision-making tasks, there is an inverted U-shaped relationship between the number of decision alternatives and task difficulty. Numerous researchers have found that decision making times increase with the number of possible decision alternatives that must be evaluated up until a certain point, after which people actually take less time to make a decision (e.g. Naylor, 1968; Payne, 1976; Hogarth, 1975; Bruce & Johnson, 1996). This suggests that if there is a relationship between the size of the problem space and task difficulty, it may not be

⁷ Obviously, these objections are also relevant to the computational complexity approach, which would also not predict strong content effects.

linear. Overall, these results indicate that the size of the problem space does not indicate task difficulty, as thus cannot be used to estimate task complexity.

D.4.3 Campbell's (1988) Information Processing Framework

Campbell (1988) has proposed an Information Processing Framework for determining the complexity of cognitive tasks. Task complexity is conceptualised as a function of the information processing demands that the task places on the individual, which can be quantified in relation to the *information load*, *information diversity* and the *rate of information change*. In the framework, any task characteristic that results in an increase in any of these factors implies an increase in task complexity.

Campbell (1988) identifies four task characteristics that result in either a high level of information load, diversity or rate of information change, and hence imply different levels of task complexity: Multiple paths to a desired goal state, multiple desired outcomes, conflicting paths to different desired outcomes and a high level of uncertainty between the relationships between inputs and outputs. A typology of tasks is proposed that categorises tasks in terms of whether they contain each of these characteristics, or combinations of these characteristics. Campbell (1988) argues that as tasks contain more of these characteristics, they become more complex.

This framework is insufficient for the purposes of estimating task complexity for a number of reasons, many of which have already been discussed in relation to the computational complexity and the problem space approaches. Firstly, the strong content effects that are observed in problem solving research cannot be explained within this framework. Secondly, the specific limitation of this framework is that it is unclear how tasks that are classified as the same within Campbell's (1988) typology of tasks might be differentiated in terms of their task complexity. For example, all items in the Raven's Progressive Matrices have only one path to the desired goal state, contain one desired goal state, should not have conflicting paths, and contain no uncertainty. Clearly, however, the least difficult items are less complex than the most difficult items. Thirdly, it could be argued that multiple paths to a desired goal state might actually result in a decrease in task difficulty, unless the task-doer is required to determine which of the paths is most efficient. Otherwise, it seems unlikely that the task-doer will necessarily consider every possible path to the desired

goal state. Finally, the emphasis on task characteristics results in a framework that is descriptive, rather than explanatory. No explanation is given for why these particular task characteristics might result in increases in information processing load.

D.4.4 Summary of objective measures of task complexity

Although this is by no means an exhaustive review of objective measures of task complexity, it serves to highlight the limitations of this perspective. Firstly, objective measures of task complexity may not reflect how the problem solver actually approaches the problem. In particular, there is little evidence that problem solvers develop mathematically optimal algorithms for solving problems or exhaustively search through the problem space for the best solution. Secondly, the empirical evidence suggests that task difficulty is dependent on the problem solvers' internal representation of the task, as strong content effects have been shown on problem difficulty. Hence, any measure of task complexity that does not take into account how human problem solvers internally represent a task is likely to be inadequate. Finally, these approaches do not make any predictions about the limits of human performance, as they do not specify when we are likely to encounter a ceiling effect on the relationship between complexity and difficulty. Thus, the conceptualisation of task complexity as an objective property of the task is likely to prove inadequate for the purposes of cognitive research.

D.5 Complexity as an interaction between psychological and task characteristics

With these limitations in mind, a number of researchers have attempted to conceptualise complexity as an interaction between psychological and task characteristics. Although a number of different approaches have been developed, the common theme is the description of the task in "terms of the behaviours and cognitive processes that subjects need to perform in order to reach a certain criteria of success" (Hackman, 1969, p. 103).

D.5.1 Relational Complexity

Relational complexity theory (RCT) has been developed by Halford and associates (Halford et al., 1998a) in order to account for the limited nature of human information processing. In RCT, the task is broken down into a series of sub-tasks, which must be completed in order to arrive at the solution. Each sub-task consists of

a number of relations and sources of variation that must be integrated in order to formulate a solution to the sub-task. A unary relation consists of a single argument and one source of variation. A binary relation consists of two arguments and two sources of variation and so on. For example, a cat has the property of being a certain size (unary relation), as does a dog. If we specify that a cat is smaller than a dog, then this constitutes a binary relation. If we further specify that a cat is smaller than a dog, under the condition that the breed of the cat and dog is known, then this constitutes a tertiary relation. The key claim is that complexity is essentially a function of the demand imposed by integrating different sources of information, and the relational complexity of a task is defined by the sub-task with the largest number of arguments that the problem solver represents in parallel in order to arrive at a solution. The relational complexity metric has been applied to a wide range of cognitive tasks, and the empirical evidence supports the assumption that increases in relational complexity lead to increases in task difficulty (e.g. Halford et al., 1998a; Andrews & Halford, 1998; Birney & Halford, 2002; Halford et al., 2005).

RCT offers a psychological explanation as to why increases in complexity should lead to decrements in performance, as the amount of information that can be processed in parallel has long been recognised as a critical constraint on human performance. Miller (1956) first suggested that human capacity is constrained to processing a small number of chunks of information, where a chunk is a unit of information of arbitrary size. For instance, the word “cat” can be represented as a single chunk or as three chunks if the letters are encoded separately (Halford et al., 1998a). Initial estimates placed the limit of human information processing at seven chunks, plus or minus two (Miller, 1956). However, this estimate has since been revised downwards in the light of further empirical evidence (e.g. Broadbent, 1975; Fisher, 1984; Halford, Maybery & Bain, 1986; Halford et al., 2005), which suggests that human information processing capacity is likely to be constrained to a soft limit of processing four relations in parallel (Halford et al., 1998a; Halford et al., 2005). Thus, the theory also makes empirically based predictions for the limitations of performance in relation to the complexity of tasks.

In order to overcome these limits on processing, RCT proposes that *conceptual chunking* and *segmentation* can reduce the number of relations that must be processed in parallel. Conceptual chunking occurs when concepts are recoded into

fewer dimensions through learning over time (Halford et al., 1998a). For example, experienced readers chunk the individual letters of a word together in long-term memory, so that they are processed as a single dimension when encountered in text. Halford et al., (1998a) argue that the cost of conceptual chunking is that the individual dimensions “...become temporarily inaccessible” (p.810). Segmentation occurs when relations are broken into steps so that they can be processed serially. For example, in multi-digit addition, experienced problem solvers break the process down into sub-tasks by grouping like digits and blocking units into groups of 10 so they can be handled more easily. As with conceptual chunking, such strategies are acquired through experience with the particular problem type. As conceptual chunking and segmentation can be used to reduce the relational complexity of a task, “...the number of arguments in a relation does not immediately translate into effective complexity” (Halford et al., 1998a, p.811). Rather, it is dependent on the problem solvers’ internal representations and strategies for solution. These mechanisms could potentially explain the strong content effects that have been observed in relation to problem solving and decision-making.

As comprehensive as RCT may be, it does have some general practical limitations. Firstly, not all tasks involve processing relations so the theory is limited in its application (Halford et al., 1998b). Secondly, the researcher must have a comprehensive model of how the task is performed in terms of the sub-tasks that must be executed for successful performance. Thirdly, as experienced relational complexity is influenced by the problem solvers’ internal representation and acquired strategies, this makes it hard to predict the difficulty of tasks where problem solvers have a large amount of domain knowledge and experience (Halford et al., 1998a). Fourthly, unless the internal representation of a particular individual is known, then it is impossible to determine the complexity of the task as it is experienced by that individual (Sweller, 1998).

D.5.1.1 Complexity frameworks similar to RCT

D.5.1.1.1 Element interactivity

In cognitive load theory (CLT), *element interactivity* has been proposed to account for the load on working memory imposed by the demands of instructional material (Sweller, 1994; Sweller & Chandler, 1994; Sweller, 2010). In the early

stages of its theoretical development (e.g. Sweller, 1994; Sweller & Chandler, 1994), element interactivity was used to explain only the intrinsic cognitive load imposed by the task, which is the “...natural complexity of information that must be understood and material that must be learned...” regardless of the instructional format (Sweller, 2010, p.124). However, more recently it has been argued that it can be used to estimate the cognitive load of the task as a whole (Sweller, 2010).

Element interactivity is the extent to which information elements in learning material interact, and hence must be processed in parallel in order to be learned, where “an element is anything that needs to be or has been learned, such as a concept or a procedure” (Sweller, 2010, p.124). Thus, in CLT the emphasis is on the narrower concept of “learning”, in comparison to the broader concept of “processing” in RCT, which includes learning, but also encompasses other mental activities such as reasoning.

In addition, the concept of an “element” is less clearly defined than a “relation”. However, it could be argued that if an element can consist of a procedure to be learned then it is likely to be composed of a number of (possibly interacting) relations. Hence, the examination of tasks at the relational level provides a finer grained level of analysis than an examination of tasks at the level of elements.

As in RCT, the level of element interactivity depends, to some extent, on the internal schemas of the learner. To return to an example given earlier, Sweller (2010) similarly argues that while novice readers may process the letters of a word as individual elements, an expert reader will process it as a single element. As in RCT, CLT proposes that element interactivity, and thus cognitive load, can be reduced through the design of efficient instructional materials, in order to encourage chunking or segmentation so that interacting elements can be processed serially, or at a lower level of interactivity (Sweller, 2010). A range of empirical evidence supports the claim that reductions in element interactivity lead to better learning in instructional settings, and hence reductions in task complexity (for a review see Sweller, 2010). Considering that the tenants of CLT in relation to element interactivity are essentially identical to those proposed by RCT, these empirical results add weight to the utility of relational complexity as an effective means of estimating task complexity in different domains.

D.5.1.1.2 Step complexity

Park and colleagues (Park, Jung & Ha, 2001; Park, Jung, Kim, Ha & Shin, 2001; Park, Jung, Kim & Ha, 2003; Park, Jung, Ha & Park, 2002; Park, Jung, Kim & Ha, 2003) have proposed a similar metric to RCT for estimating the complexity of emergency operating procedures, which they refer to as *step complexity*. The task to be performed is broken down into a series of sub-tasks, each with separate goals, which are further decomposed into a series of procedural steps. Procedural steps are at an equivalent level of analysis to relations in RCT. The complexity of each sub-task is then estimated as a function of the amount of information that is needed to complete the sub-task, the logical structure of the procedural steps to be performed and the number of procedural steps that need to be executed. As in RCT, the complexity of the task is equated with the step complexity for the most complex sub-task. The utility of the measure was demonstrated in a series of studies conducted by Park and associates (Park et al., 2001; Park et al., 2002; Park et al., 2003), who found that it was a significant predictor of response times in accomplishing the procedural steps associated with each sub-task in the emergency operating procedures of a nuclear power plant.

One problem with this measure is that it is difficult to quantify the complexity of a procedural step (Park & Jung, 2007). For example, is a step that entails the comparison of two values more complex than a step that entails the prediction of a value, given a certain state in the system? Park and Jung (2007) suggest a rather complex answer to this problem that involves modelling individual procedural steps through graph entropy, a method that has traditionally been used to evaluate the complexity of software. However, as mentioned previously, as procedural steps are at the same level of analysis as a relation in RCT, it would be possible to quantify the complexity of procedural steps as a function of the number of relations that must be integrated in order to instantiate them. In psychological terms, this has a clear advantage over the graph entropy method, as it provides an explanation for the complexity of procedural steps with regards to cognitive processes. In comparison, the application of the graph entropy procedure is likely to suffer from the same problems that were discussed in relation to computational complexity earlier, as the assumption that we (human problem solvers) process information in the same way as computers is highly questionable.

D.5.2 Wood's (1986) conceptualisation of task complexity

Wood (1986) proposes that tasks can be described in terms of their required acts, products and information cues. Acts are the mental activities and behaviours that are undertaken to produce a given product, where a product is the measurable result of an act. Information cues are the knowledge relevant for performance and feedback from the task. Overall, task complexity is conceptualised as a function of the information processing demands of the task to be performed.

It is argued that three types of complexity contribute to the overall information processing demands of the task: *Component complexity*, *coordinative complexity* and *dynamic complexity*. Component complexity is "...a direct function of the number of distinct acts that need to be executed in the performance of the task and the number of distinct information cues that must be processed in the performance of these acts" (Wood, 1986, p.66). It can be reduced if there is a degree of overlap among the acts or information cues. For example, if a particular act must be repeated at multiple stages of the task. Coordinative complexity is the extent to which acts require precise scheduling and timing. For example, coordinative complexity increases if acts must be performed in a specific order. Dynamic complexity is the extent to which there is uncertainty between the relationships between acts and products. A mathematical metric is provided to assess each component of task complexity, although Wood (1986) notes that at the current stage of research it is not clear how each type of complexity interacts with the other, and it is unlikely that the factors are orthogonal. In comparison to RCT, the emphasis here is on the demand imposed on the task-doer by the entire task, rather than the demand imposed by the most complex sub-task.

One essential problem with this approach involves determining what constitutes an elementary level of analysis in terms of acts and information cues. With regard to the analysis of acts, Wood (1986) notes that an act can be defined at varying levels of abstraction, ranging from the very to specific to the more complex. For example, "inference" is considered to be one act because it produces a single product, although it is likely composed of multiple mental processes. As with the step complexity metric discussed earlier, it then becomes difficult to compare the complexity of different acts. As the quantification of the complexity of acts is not specified, this provokes the question of whether a task that requires many acts that

are composed of few processes is more complex than a task that requires a single act that is composed of many interacting processes. This issue remains unresolved in Wood's (1986) task complexity metric.

With regard to the representation of information cues, Wood (1986) does not explicitly address how expertise or domain-relevant knowledge might affect the formation of information cues, and thus overall task complexity. However, he does acknowledge that domain-relevant knowledge is likely to have a significant impact on the number of information cues that are perceived by the task doer. The framework also specifies that a comprehensive model of how the task is internally represented by the problem solver is required before adequate estimates of task complexity can be made using this metric.

Bonner (1994) argues that Wood's (1986) conceptualisation of task complexity is insufficient because it does not adequately capture the uncertainty of real world tasks. He proposes an alternative approach in which tasks are described in terms of their *inputs*, *processes* and *outputs*, which are essentially equivalent to Wood's (1986) information cues, acts and products, respectively. As in Wood's (1986) approach, task complexity is described as a function of the number of each of these elements. The main extension is that task complexity is also related to the *clarity* of each of these elements. Although Bonner (1994) names a number of distinct factors that may influence clarity, it essentially seems to reflect to the extent to which the situation is uncertain. For example, the extent to which there are consistent relationships between inputs and outputs, and the extent to which cues and goals are specified. Thus, although Wood (1986) metric for dynamic complexity captures some of these factors, it does not reflect all of them.

The empirical validity of Wood's (1986) metrics is not addressed in the original paper. In addition, although a Google Scholar search undertaken on the 25th June 2010 indicated that Wood's (1986) paper is cited in 473 separate papers, few researchers have actively applied the framework as it is set out in its original formulation. This perhaps reflects the difficulty of meaningfully decomposing a task into required acts, products and information cues. It has, however, been used to classify tasks on a post hoc basis in a meta-analysis conducted by Wood, Mento & Locke (1987) on the effect of goal-setting in cognitive tasks. Across the studies

included in the meta-analysis, setting challenging goals led to higher levels of performance in simple tasks, but the effect was much smaller in more complex tasks. This indicates that task complexity, as defined by Wood (1986), does appear to have a psychologically meaningful impact on task performance across a wide range of cognitive tasks.

In a more recent study, Beckmann (2010) has applied Wood's (1986) approach to quantify the complexity of figural series completion tasks. In these tasks, subjects are shown a series of geometric shapes and must describe the next shape in the series. The tasks were presented differently in two conditions, "Standard" and "Memory Eased". In the former condition, subjects had to describe all the features of the next shape in a single answer, while in the latter condition, verbal descriptions of the next shape could be made gradually, and as the subject named a feature, such as the colour green, it appeared on screen. An information-processing approach was used to describe the tasks in both conditions in terms of their component acts, products and information cues, and subsequently, Wood's (1986) component complexity metric was used to determine their level of task complexity. This approach was contrasted with a task analysis based on the notion mental load as indicated by element interactivity in CLT, which was discussed earlier. Different complexity estimates were derived for the tasks presented under different conditions, and these estimates lead to the prediction that difficulty would be lower in the "Memory Eased" than in the "Standard" condition. In contrast, mental load estimates did not predict such differences. The results showed that tasks presented in the memory eased format did have a lower level of difficulty than in the standard format. This suggests that Wood's (1986) approach provides an empirically valid means by which to provide an a priori estimate of the effect of the task and the situation on performance.

D.5.3 Summary of complexity as a function of psychological and task characteristics

RCT and Wood's (1986) framework converge on the idea that task complexity is a function of the information processing demands imposed by the task, which is indicated by the number of processes that need to be executed in the performance of the task, and the dependency among those processes. This is determined by the internal representation of the task-doer and objective requirements of the task.

Hence, decrements in performance related to task complexity are explained by constraints on the human information processing system.

In general, the limitations of each these frameworks are similar. Firstly, the researcher must have a model of ideal task performance. Secondly, as complexity is influenced by the problem solvers' internal representation of the task and acquired strategies, this makes it hard to predict the difficulty of tasks where problem solvers have a large amount of domain knowledge and experience. Thirdly, unless the internal representation of a particular task-doer is known, then it is impossible to determine the complexity of the task as it is experienced by that task-doer (Sweller, 1998). In cognitive research, these problems could be overcome through the development of comprehensive models of task performance, and utilising tasks that are novel for all problem solvers.

The main difference between the two main frameworks discussed above is in terms of their applicability. RCT is limited in application to tasks that involve the processing of relations. In comparison, Wood's (1986) explicit aim was to develop a general theoretical model for the description of tasks that are used to study human behaviour. As "required acts" and "information cues" can be specified at varying level of abstraction, the framework is flexible enough that it has the potential for application to a wide range of cognitive and motor tasks. For example, an "act" may involve the processing of a relation, but it could equally be used to describe the generation of a strategy, or the retrieval of declarative knowledge. Therefore, Wood's approach may be applied to a wider range of cognitive tasks than RCT.

D.5 Conclusions

At a conceptual level, complexity has been aligned with the intricacy and interconnectedness of tasks. In turn, this can be conceptualised as an objective property of the task, or as an interaction between psychological and task characteristics. The adoption of the first perspective is problematic, as it essentially ignores how the task is performed by the task-doer. The second perspective emphasises a process-orientated account of task complexity. Under this view, complexity is seen as a function of the information processing demands of the task to be performed. This approach is likely to prove more successful in the prediction of task performance, as it takes into account the internal representation of the task-doer and the objective

requirements of the task, which are both likely to impact upon the overall difficulty of the task to be performed.

APPENDIX E

EQUATIONS FOR THE SYSTEMS USED IN STUDY 3

V3x3-R-3-C-1: 3 inputs, 3 outputs, 3 relations, connectivity 1

$$U = U_{t-1} + 0.8A$$

$$V = V_{t-1} + 1.6B$$

$$W = W_{t-1} - C$$

V3x3-R-6-C-2: 3 inputs, 3 outputs, 6 relations, connectivity 2

$$U = U_{t-1} + 0.8A + 0.8B$$

$$V = 0.8V_{t-1} + 1.6A$$

$$W = 1.2W_{t-1} - C$$

V6x6-R-12-C-2: 6 inputs, 6 outputs, 12 relations, connectivity 2

$$U = U_{t-1} + 0.8A + 0.8B$$

$$V = 0.8V_{t-1} + 1.6A$$

$$W = 1.2W_{t-1} - C$$

$$X = X_{t-1} + 0.8D + 0.8E$$

$$Y = 0.8Y_{t-1} + 1.6D$$

$$Z = 1.2Z_{t-1} - F$$

V3x3-R-7-C-3: 3 inputs, 3 outputs, 7 relations, connectivity 3

$$U = 0.8U_{t-1} + 0.8A + 0.8B$$

$$V = 0.8V_{t-1} + 1.6A$$

$$W = 1.2W_{t-1} - C$$