The effects of task complexity on performance in constraint satisfaction design

Victoria Smy
BSc (Hons) Psychology, MSc Occupational Psychology

A Thesis submitted for the
Degree of Doctor of Philosophy

School of Psychology, Cardiff University
Tower Building, Park Place
Cardiff, CF10 3AT

April 2013

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ……………………VASmy… (Candidate) Date …26/06/2013……………

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD

Signed ……………………VASmy… (Candidate) Date …26/06/2013……………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. A bibliography is appended.

Signed ……………………VASmy… (Candidate) Date …26/06/2013……………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……………………VASmy… (Candidate) Date …26/06/2013……………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed ……………………VASmy… (Candidate) Date …26/06/2013……………

ACKNOWLEDGEMENTS

I would like to thank Professor John Patrick for his patience, guidance and encouragement. I would also like to thank John for all of the opportunities for personal and academic development he has opened up for me.

I would like to thank all the people who have provided assistance in some way during my research. In particular, I would like to thank L. Tiley, L. Cowley, H. Seeby, Dr. P. Morgan, H. Rolph, and Dr. J. Burton, for their help in participant recruitment, video transcription, data checking and for providing well-timed advice and keeping me sane. I would also like to thank I. Kaloskampis and A. Anthopoulos for their programming skills and technical support.

Further thanks go to my parents, sisters and friends for all their support throughout the years.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

SUMMARY

The aim of this thesis is to examine the importance of constraints in design activity, and more specifically, in constraint satisfaction tasks. Constraints are involved in all design tasks and denote criteria on what constitutes a good design outcome. Within the present research, a constraint stipulates a restriction on how a design element may be assimilated into a design. For instance, given a spatial office layout design, the positioning of an employee may be restricted by a constraint stipulating that they should be in close proximity to a particular area. This thesis attempts to address a gap in the design literature by examining the effect of such constraints on design performance using experimental methodology. Two research threads are addressed; the effects of constraints on performance, and how constraint satisfaction performance can be improved by training. In the first research thread, the theoretical framework adopted concerns Newell and Simon’s (1972) problem space theory and more recent suggestions by Halford, Wilson and Phillips (1998) concerning relational complexity. These are used to predict the effects of increasing the number of design constraints (Experiments 1-2) and the number of types of constraint (Experiments 3-4). Both these factors together with a reduction in the degree of cognitive fit (Vessey & Gellata, 1991) between the constraints and the external representation (Experiment 5) were found to reduce design performance. The second research thread examines whether training can improve design performance. Practice only was found to improve performance on a near transfer task relative to a control group but not a dissimilar, far transfer task (Experiment 6). The subsequent Experiments examined the effect of what has been labelled ‘metacognitive’ training on performance. Findings indicated that a training intervention aimed at encouraging either reflective self-explanation (Experiment 7), or aimed at improving planning strategy (Experiment 8) improved performance in comparison to both control and practice only groups. The implications of these results are discussed together with future research directions.

OVERVIEW OF THESIS

The present experimental thesis is concerned with design problem solving. More specifically, it is concerned with the cognitive processes of humans when satisfying external design constraints. The structuring of relevant theoretical threads and empirical research are organised as follows:

- Chapter 1 outlines what is meant by the term design problem solving and the situations in which it occurs. The main features of design problem solving are described and the notion of variation in the level of design task structure is detailed. Following from this, preliminary evidence into human cognition in design problem solving is briefly reviewed and research aims are outlined.
- Chapter 2 introduces some psychological theories that may shed light on how variation in external constraints impacts upon design complexity. In addition, two experimental paradigms are described in terms of their relevance to applied situations and their psychometric properties.
- Chapters 3-5 describe experimental work investigating whether variation in design constraints affects measures of design performance. This captures the first experimental theme of this thesis.
- A foreword to the remaining experimental work outlines a change in focus, from constraint factors affecting design performance, to training initiatives that may aid design performance. Relevant theories are introduced briefly.
- Chapters 6 and 7 present empirical work investigating how practice and training interventions affects design performance. This constitutes the second research theme. Both the intrinsic feedback available through practice (without further training

instruction), and practice plus metacognitive training interventions, specifically self-explanation techniques, were examined.

- Finally, Chapter 8 summarises the empirical research undertaken presently and general conclusions and implications are drawn. Afterwards, research limitations are addressed and avenues for future research are identified.

CHAPTER ONE

Introduction to design problem solving

1.1 What is design problem solving?

Design is concerned with investing effort in order to create a new product, or new specifications, given no set procedure for doing so (Ball, Evans, Dennis & Ormerod, 1997; Purcell & Gero, 1996). As such, design problem solving is concerned with the cognitive process(es) engaged in when a structure or object is being created. Likewise, research into human design processes concerns itself with identifying sources of cognitive complexity in design. Such research may enable the development of training or support methods that may ease the level of cognitive complexity encountered.

Design is both a formal discipline and an everyday activity (Visser, 2004). It permeates many technical and professional disciplines such as engineering, architecture and graphical design. In such disciplines, large-scale projects may affect many users. Here consequences of bad design may be wide-reaching. Such design failures may be costly, dangerous or may require the initiation of time-consuming redesign. As an everyday activity, design is abundant in multiple situations, albeit on a smaller scale. The functional arrangement of any living accommodation will incorporate some design principles. Indeed, any form of spatial arrangement may be considered a design process. The scheduling and organising of multiple activities may also be considered a design activity. Route-planning is a further form of common design activity. The consequences of everyday design failures here may not necessarily be as extreme as professional design failures. However, consequences of bad design may, at best, still be inconvenient or unproductive.

Design problem solving is a notoriously difficult activity (Chandrasekaran, 1990; Guindon, 1990; Römer, Leinart & Sachse, 2000). Typically involving ill-specified, incomplete or ambiguous goals, and no obvious solution (Ball *et al.*, 1997; Purcell & Gero, 1996; Restrepo & Christiaans, 2004), the design of any new item will require a designer to strategically utilise various cognitive abilities including analysis, prioritisation and decision making (Ball, Lambell, Reed & Reid, 2001; Hacker, 1997, Römer, Pache, Weiβhahn, Lindermann & Hacker, 2001). Designers must be able to tolerate ambiguity and uncertainty whilst still showing good judgement (Dym, Agogino, Eris *et al.*, 2005). Smith and Browne (1993) highlight five crucial elements of design that are inherently intertwined with the cognitive processes of designers when undertaking design problem solving. These are goals, constraints, alternatives, representations and solutions. Each element is now briefly described.

Goals: Design problem solving necessarily involves the presence of unmet goals. These goals are helpful in that they allow a designer to ascertain criteria for evaluating design solutions (Smith & Browne, 1993). Additionally, goals may be decomposed into sub-goals to allow the designer to create structure to direct their design efforts more effectively (Liikkanen & Perttula, 2008). Chan (1990) suggests that the prioritisation of, and focus on, a design sub-goal can be considered a strategic way of starting the design process. Nevertheless, goal identification in design problem solving situations is complex. Goals may be multiple, incomplete, and of varying and sometimes unknown importance. Further difficulty may arise when there is no apparent match between a goal or sub-goal and possible solutions or design alternatives.

Constraints: Constraints may derive from different sources and are usually concerned with what is required or feasible in a design situation. Smith and Browne (1993) refer to the existence of external and internal constraints. External constraints may refer to social and environmental policies, the wishes of the client/stakeholders and the context in which designing is taking place. External constraints may also be necessary restrictions imposed upon a design task. Internal constraints are defined as the essential functions, or requirements of the artefact being developed (Smith & Browne, 1993). In the field of cognitive psychology, internal constraints are most often conceptualised in terms of the cognitive limitations of the problem solver. One such constraint is the presence of limited working memory (i.e., Miller, 1956; Sweller, van Merrienboer & Paas, 1998), or the conditions surrounding the willingness to deploy memory-based problem solving strategies (i.e., Waldron, Patrick & Duggan, 2010).

Whilst the presence of external design constraints may aid a designer in narrowing down their search for sensible solutions (Visser, 2004), the presence of overly restrictive constraints may lead to less innovative and less effective design (Smith & Browne, 1993). Extensive numbers of constraints may be expected to diminish design possibilities. The issue arising from this is how to determine which constraints may not be relaxed, as ignoring such constraints will inevitably result in flawed designs. It may be that some proposed constraints cannot be practically or effectively assimilated into a design. Also, some design constraints may not be apparent at the onset of design activity but may arise as a product of exploring differing design alternatives. Indeed, in conceptual design where novel products or services are created, the process of problem decomposition and the exploration of design alternatives often lead to the identification of additional design constraints (Chan, 1990). The discovery of more constraints at these points may necessitate a revision of design problem solving strategy.

Alternatives: Alternatives are possible courses of action that act as intermediaries between the current state of the design problem and the goal (or sub-goal) that the designer is trying to achieve. Smith and Browne (1993) point out that alternatives do not ‘pre-exist’, rather they emerge via the process of goal-directed search. In design problem solving, the often ill-specified nature of the goals may lead to an undisclosed amount of possible alternatives. These alternatives may themselves be vague, incomplete or dependent on the resolution of related design issues (Ullman & D'Ambrosio, 1995). While prioritisation of goals and constraints may serve to filter out less desirable alternatives, there is always the possibility that some overlooked constraint may cause a desirable alternative to be overlooked. Often design entails an iterative process of alternative generation and alternative evaluation, with subsequent design decisions made on the basis of trade-offs between what is desirable (goals) and what is possible (constraints).

Representations: Representations refer to a depiction of the problem space in which design problem solving takes place (Smith & Browne, 1993). Representations may be internal or external. Internal representations correspond to a designer's mental model of the relevant design knowledge (goals and constraints) and the possible alternatives which can be generated from this knowledge. External representations may contain similar information but can take various forms including sketches, diagrams, and graphs. In addition, representations may be complete, entailing all relevant knowledge and alternatives for the problem in question, or incomplete containing unknown parameters and possibilities (Goldschmidt, 1997). Visser (2006) describes representations as 'operative' in that they are constantly under revision as new information, constraints, or design developments are integrated into it. As such, representations should be flexible and restructuring, a form of representation translation, can occur (Akin, 2002; Jones & Schkade, 1995; Visser, 2006).

Solutions: Solutions are a designers attempt to fulfil design goals. They are necessarily complex, usually entailing many attributes and interrelations (Smith & Browne, 1993). Due to the uncertainty inherent in the nature of design problem solving, design solutions are likely to be merely satisfactory rather than optimal (Cross, 2001). Often, formulated designs are simply effective until a situation arises whereby the end design runs into unforeseen circumstances that prevent it from functioning as intended. At this point a design may be recycled, modified, or the design process may begin again.

It is the interlinked nature of all of Smith and Browne's (1993) crucial elements which make design inherently complex. In an ideal design situation, a designer would endeavour to gain a full understanding of design goals and constraints leading to the construction of a comprehensive representation of the design problem space. Having achieved this, the designer should be able to weigh the costs and benefits of the alternatives available to them. However, in practice, design is never as straight forward as this. The ill-defined and ill-specified nature of most design tasks (Ball *et al.*, 1997; Visser, 2006) may make exhaustive problem scoping impossible. Some aspects of the problem representation or some design alternatives may never be identified or explored. Designers may be sensitive to cost-benefit trade-offs. As design is computationally costly, cognitive short-cuts are likely to be applied. However, there is some suggestion that a degree of structure can be imposed onto a design problem solving task, as knowledge about that design task is accumulated (i.e., Restrepo & Christiaans, 2004). Design researchers have attempted to detail the design process in terms of a prescribed set of systematic and sometimes iterative stages, whereby a design task may vary in its degree of structure at any stage. A couple of examples of prescriptive accounts of design problem solving are now discussed.

Dym's (1992) prescriptive account describes the beginning of the design process as (1) the 'conceptual stage' involving the identification and prioritisation of goals and constraints, the exploration of design alternatives, the gathering of further information and the resolution of trade-offs between conflicting goals and constraints. This is followed by (2) a 'preliminary design' period in which the components and design sub-parts are identified. Then follows (3) 'detailed design', whereby a designer outlines the specific parts needed to construct the end product. Finally (4) 'analysis and optimisation' occurs. This involves the testing and evaluation of the design. Coming from an engineering background, Dym (1992) essentially captures the design process, from its start, to the point at which a design can be mass-produced, mass-distributed, refined and recycled. Hacker (1997) offers a similar, although somewhat simplified, description of design that focuses on its' initial stages rather than later manufacturing processes. Hacker's (1997) stages are, (1) the identification and clarification of the design problem, (2) the development of a frame of conceptual solutions, (3) the design of a favoured solution, and (4) working out details so as to make the end design both functional and feasible.

It appears that it is the early stages of a design process that are the most ill-specified and ill-structured periods of design. This is especially true of the early stages of conceptual design (Dym, 1992), where a novel product must be created from scratch. It is here that uncertainty surrounds many of the crucial design features outlined by Smith and Browne (1993). In later stages, some of the uncertainty has been alleviated due to sourcing of new design information and the setting of new design parameters. This should result in the emergence of more definite design goals and requirements, and the introduction or acknowledgment of relevant constraints. As such, a more accurate representation of the design task should emerge, and subsequent design problem solving efforts may be directed more effectively. Towards the latter stages of design, an increased level of structure has been imposed onto the task (Fernandes & Simon, 1999; Restrepo & Christiaans, 2004). Therefore, structure might not be an invariant feature in design problem solving. The section that follows discusses properties of well-structured and ill-structured tasks.

1.2 Properties of task structure

There are several published papers that attempt to highlight the crucial differences and distinctions between well-structured and ill-structured problem solving tasks (Dorst, 2006; Fernandes & Simon, 1999; Goel & Pirolli, 1992; Shin, Jonassen & McGee, 2003). Borrowing heavily from the researchers referenced, typical characteristics of well-structured and ill-structural problem solving tasks are now reviewed. Table 1.1 depicts the differences between

Table 1.1. Important distinctions between well-structured and ill-structured tasks or problems

Feature	**Well-structured**	**Ill-structured**
Knowledge of problem elements	Complete knowledge of problem elements. All problem features are presented. Goals and constraints are known and there is no need to source further information.	Incomplete knowledge of problem elements. Fails to present all or some problem elements. Goals and constraints may be vague or may conflict.
Presence of solution and evaluative criteria	There is a definite, correct solution. This clear solution provides evaluative criteria that can be used to gauge progress towards the solution.	No optimal solution. Several satisfactory solutions may exist. The lack of a clear solution criteria leads to uncertainty concerning how to monitor progress towards task solution. As such, there may be no clear criteria for terminating problem solving.
Solution pathways	There is a set, direct pathway to the solution that may be navigated from application of task rules and restrictions and logic. Any divergence from this pathway is indicative of cognitive errors.	The lack of a definite solution criteria means that there are no obvious pathways to a solution. As such, there is no simple way of reaching a solution. The potential viability of multiple pathways may make it harder to identify any errors made.
Task rules or principles	Productive problem solving moves can be readily identified from logical application of explicit task rules and constraints. These rules and constraints are known throughout problem solving.	The productivity of problem solving moves may be unknown. Further rules and constraints may continue to emerge as problem solving progresses. Any constraints or rules indentified may require a re-evaluation of task goals and current progress. Task rules or constraints may conflict. Therefore, prioritisation of emerging rules and constraints may occur.

well-, and ill-structured tasks on four crucial task features; knowledge of problem elements, solution criteria, solution pathways and task rules or principles. The distinction between the levels of structure on all of these criteria is the degree of specification and clarity. Well-defined problem will have clear goals, clear constraints and also clear rules for attaining the desired solution. In contrast, ill-defined problems are underspecified in respect to task goals, constraints and rules. Here there is no 'best' solution. Instead there may be a number of satisfactory solutions reachable by multiple and sometimes convoluted pathways. As such, ill-structured problem solving tasks are heavily under-determined in that they lack clarity about many problem aspects at the outset (Dorst, 2006).

Simon (1973) suggested that the distinction between well-structured tasks and ill-structured tasks is not set in stone. Instead, ill-structured and well-structured may represent the extremities of a continuum (Fernandes & Simon, 1999), with intermediate levels of task structure in between. Indeed, it may be necessary to impose structure upon a task before a solution can be reached. As previously alluded to, it is possible to gain (or to impose) structure within design problem solving contexts (Restrepo & Christiaans, 2004). Problem structuring, sometimes referred to as problem setting or problem framing, may occur during the early stages of a design cycle, and may reoccur periodically as more relevant information is assimilated (Fernandes & Simon, 1999; Goel & Pirolli, 1992; Schön, 1987). Problem structuring is the process of drawing upon knowledge, experience, or new information to compensate for unknown parameters (Restrepo & Christiaans, 2004; Schön, 1987). Table 1.2 captures and formalises some structuring effects within design contexts.

Table 1.2. Proposed structure of design tasks

Feature	**Ill-structured design** (early stages of the design process such as conceptual or creative/novel design)	**Semi-structured design** (later stages of the design cycle such as design detailing, configuration, constraint satisfaction design)
Knowledge of problem elements	Ill-specified problem elements. There may be open-ended goals and few explicit constraints.	Knowledge of all problem elements. Goals and external constraints known.
Presence of solution and evaluative criteria	No obvious solution criteria.	No obvious solution criteria.
Solution pathways	Multiple pathways to solution.	Multiple pathways to solution.
Task rules or principles	Vague knowledge of task rules, principles and constraints. Rules may conflict.	Knowledge of task rules, principles and constraints. Rules may conflict.

Early conceptual design, sometimes referred to as creative design (i.e., Goel & Pirolli, 1992), exhibits all of the characteristics of ill-structured tasks outlined in Table 1.1. However, later stages of design, such as Dym’s (1992) detailed design stage, or Hacker’s (1997) working out stage, have a greater degree of structure. Here, parameters concerning goal specifications and relevant constraints have been defined more clearly. Some researchers refer to such forms of design as configural design (Bayazit, 2004) or constraint satisfaction tasks (Sabin & Freuder, 1996; Visser, 2004). Whilst these design activities are semi-structured, they are not fully specified. Some of the complexities inherent in ill-structured designs are still evident. There are still multiple pathways to solution and no clear stopping criteria. In addition, uncertainty may still surround the legality of design actions due to conflicting or complex task rules.

So far, discussion has centred around situating design as a problem solving activity comprised of a number of key features (Smith & Browne, 1993) and a less than optimal level of task structure (i.e., Restrepo & Christiaans, 2004). Attention now turns towards empirical evidence and documentation of the design process. In particular, the following section focuses on empirical research concerning human cognition in design contexts.

1.3 Empirical evidence concerning design problem solving

Many design activities do not easily lend themselves to empirical investigation. Some design cycles are lengthy. They may take place over a number of weeks or, for huge design projects, perhaps even years (Pahl, Beitz, Feldhusen & Grote, 2007). Design may also involve multiple designers, or stakeholders, each performing a different role. Also, in applied situations, the design process is not typically replicated and may not be stringently documented. As such, comparisons of similar design processes may be difficult. Nevertheless, there is a slowly expanding body of research focusing on human cognition in design. A brief literature review detailing some key research themes are now summarised in the following sub-sections.

Design strategies

Following on from prescriptive models of design, the notion that design is carried out as a strategic activity has been examined. Strategic designers are expected to perform thorough decompositions of design problems and then work towards accomplishing the goals and sub-goals identified. A distinction has been made between top-down, depth-first approaches and top-down, breadth-first approaches to design (Ball & Ormerod, 1995; Ball *et al.*, 1997). In depth-first approaches, a designer may work towards satisfying one sub-goal before considering how to achieve the next. In breadth-first approaches a designer will work on many sub-goals simultaneously, at increasing levels of detail. The breadth-first approach is generally favoured (Ball & Ormerod, 1995), as this approach preserves the interrelations between sub-goals whilst also monitoring and resolving any conflicting design aspects. In the case of depth-first approaches, the piecemeal nature of the design may result in part solutions that need to be reworked and reconfigured to achieve a satisfactory and functional end design (Lee, Eastman & Zimring, 2003). However, it is possible that some designers switch between various top-down approaches, or fail to engage in a strategic top-down strategy altogether.

Ball *et al*. (1997) and Lee *et al.* (2003) found evidence to support a top-down, breadth-first approach to design. Ball *et al*. analysed the verbal protocols of six electronic engineers designing a novel integrated electrical circuit, whilst Lee *et al.* examined the protocols of two groups of architects redesigning an architectural studio. Ball *et al.* report that a depth-first approach was demonstrated by a progression from problem understanding activities to developing high-level solutions. These solutions were then refined into progressively greater levels of detail. Lee *et al.* interpreted a frequent level of switching between design activities producing student workspaces, and those producing group workspaces, as indicative of a breadth-first approach. They further attributed better design to breadth-first activities. Guindon (1990) also examined the notion that strategic top-down activity occurs in design problem solving by analysing the verbal protocols of two practicing software designers programming a lift to move between floors based on a list of rules. Guindon (1990) reported that although designers did evidence a tendency towards top-down breadth-first design strategies, there were considerable deviations from this pattern. Here, designers intertwined solution development with reconsideration of the design requirements. Guindon suggested that opportunistic deviation from top-down approaches took place for several reasons, including insufficient or conflicting information, the recognition of interrelated parts of a design, or the economical use of newly discovered solution options.

Further research has attempted to isolate and identify the prevalence of design activities such as problem decomposition more closely. Liikkanen and Perttula (2008) propose that the recognition of design principles should both encourage, and follow from, explicit decomposition. As such, explicit decomposition should be an effective design strategy. Liikkanen and Perttula observed mechanical engineering students performing a design problem. Designers were instructed to sketch and annotate design concepts. Sketches were then analysed, and verbal protocols coded, to determine whether designers proceeded through identifiable stages of problem decomposition, the justification being that an explicit form of decomposition would be evident from evidence attesting to a progression of design activity from analysis, through goal-setting, to design solution development. Only three of the 16 designers explicitly decomposed the design problem, but their design productivity levels benefitted from this process. The authors imply that more experienced designers may be more proficient in performing explicit problem decomposition. On a related line of enquiry, Kruger and Cross (2006) observed the cognitive activities of nine industrial designers whilst designing a rubbish disposal system. Inspection of verbal protocols led the author to propose four strategic design approaches:

(1) Problem driven design, characterised by high levels of data gathering, constraint identification and solution generation

(2) Information driven design, characterised by high levels of data gathering and constraint identification, but little solution generation

(3) Solution driven design, characterised by high levels of solution generation and relatively little data gathering

(4) Knowledge driven design, characterised by a high level of modelling activity

Kruger and Cross (2006) found that designers adopting problem and information driven strategies, thereby showing a greater degree of explicit problem decomposition, produced designs of greater overall quality. Whilst this does not allow the conclusion that experienced designers necessarily engage in more explicit problem decomposition, it appears to support the notion that engaging in more thorough problem decomposition benefits design efficiency.

Atman, Chimka, Bursic and Nachtman (1999) examined whether experience differentially affects designer's cognitive activities by comparing 1st and 4th year engineering students designing an urban playground. Videotapes were analysed for evidence of several cognitive activities relating to problem analysis, solution generation and solution development. Atman *et al.* noted some key differences between novice and experienced designers. Novices spent more time defining the problem whereas experienced designers spent more time gathering relevant information. Both groups spent equivalent proportions of time on solution development although experts produced more design alternatives. As could be expected, experienced designers were credited with better quality designs.

Studies surrounding design strategies are necessarily exploratory. A lack of statistical analysis within this area means inferences are often made on the basis of qualitative analysis utilising small samples. Prescriptive models of design imply design should be approached strategically. Indeed, the initial studies (Ball *et al.*, 1997; Guindon, 1990; Lee *et al.*, 2003) concluded that breadth-first approaches were adopted, albeit with some opportunistic deviation (Guindon, 1990). Further studies suggest that problem decomposition, an activity believed to occur at the beginning of the design process may not be undertaken thoroughly or explicitly, regardless of the level of design experience (Atman *et al.*, 1999; Liikkanen & Perttula, 2008). Nevertheless, some tentative differences between novices and experienced designers have been proposed (Atman *et al.*, 1999). Overall, mixed results, and the use of differing qualitative criteria, mean that firm conclusions regarding design strategies are hard to pinpoint. It appears that individuals vary greatly in their approach to design with no one strategy favoured consistently.

Sketching as an external cognitive aid

Many researchers suggest that sketching, as a form of external memory, may relieve cognitive load and prove helpful when revising and refining design ideas (Bilda, Gero & Purcell, 2006; Cross, 2001; Römer *et al.*, 2000). The benefits of sketching may be many and varied. For instance, Bilda *et al.* (2006) proposed that sketching plays a vital role in the acquisition and representation of design concepts. Goldschmidt and Smolkov (2006) studied the effects of being able to sketch on design outcomes using a sample of industrial designers. Half sketched freely, others just sketched their end design. Sketching did not appear to effect design originality. However, significantly higher practicality scores were found when sketching was permitted. In another study, Sachse, Leinert and Hacker (2001) asked students from technical disciplines to complete a regular and a complex computer aided design (CAD). Half were permitted to sketch. No benefit of sketching was found in the regular CAD task, but an effect was found within complex CAD tasks, as those permitted to sketch had shorter solution times and used fewer design moves. No difference in the solution quality was found. Of designers who sketched, most felt sketching to be unnecessary for the regular CAD problem, but all reported benefits when performing the complex design task. Benefits include easier problem analysis and support during the planning and structuring of design solutions. Both Goldschmidt and Smolkov (2006), and Sachse *et al.* (2001) propose that sketching is beneficial when tackling complex design problems. In a study of somewhat narrower focus, Römer *et al*. (2000) investigated whether sketching can lead to improved design problem analysis. Undergraduate designers analysed low and high complexity mechanical systems. Half were instructed to sketch system features. Results indicated that those sketching recalled more facts and interrelations. The effect was greater when recalling information about complex systems. Designers who sketched also perceived the tasks as less difficult. They further found that designers who sketched recalled more information only when allowed to retain their sketches. Römer *et al*. (2000) suggest that sketching is beneficial as an external representation of design problems when the sketches produced remain available for later inspection. When memory alone must be relied upon, previous sketching activities do little to aid detailed problem representation. Collectively, the experiments reported tentatively indicate that sketching is of some benefit during the solving of complex design problems but may not be linked to overall design quality.

External problem representation

There is some suggestion that the presentation in which a design problem is relayed has an effect on the quality of design solutions. To test this notion Carroll, Thomas and Malhotra (1980) had students perform a spatial design task and an isomorphic manufacturing scheduling task. Results indicated that designers tackling the spatial task achieved higher scores and completed their designs quicker than those completing the scheduling task. Carroll *et al.* suggest that spatial design tasks may be easier, as they involve designing at a physical, rather than an abstract, functional level. The authors also noted that all of the designers completing a spatial design task produced a graphic representation of the problem, whereas only two of the designers in the scheduling task produced such a representation. Carroll *et al.* suggest that spatial tasks are more easily translated into graphical representation, and that this process may aid solution generation. As a follow-up, Carroll *et al.* examined whether providing designers with an external diagram representing the design problem aids the generation of appropriate solutions. Forty-five students, completing the same spatial and scheduling tasks, were provided with a graphical representation of the design problem in the form of a matrix that solutions could be mapped onto. Results indicated higher design scores were still evident when performing spatial design tasks. However, the difference was reduced in comparison to previous results. Carroll *et al.* concluded that providing an appropriate external representation may mitigate some of the complexities of a problematic design task and make the generation of design solutions easier.

Römer *et al.* (2001) attempted to more closely identify the benefits afforded by external problem representations by analysing the self-reported use of such representations by engineering designers working in various industrial settings. Results indicated that 95% of those sampled used sketches during the early, conceptual stages of design. CAD representations were used by 67% of respondents, and over half of the sample constructed models during the initial stages of a design process. Of those utilising representations, sketching and CAD were used more frequently than models. The self-reported functional uses of external representations included; developing and testing solutions, checking requirements, supporting memory processes, and also documenting and communicating design ideas. Overall, sketching was credited with the biggest impact on improving design quality, followed by modelled representations. Sketching was also attributed with faster development of design solutions in comparison other with models. Römer *et al.* conclude that external representations are a vital factor in any design process, but the use of them may depend on the ease of their generation.

In another experimental study, Jones and Schkade (1995) examined the suggestion that alternative representations, despite being informationally equivalent, may still differ in the cognitive demands and benefits placed on the designer. As such, designers may translate a problem representation into a form that is more compatible with their preferences. These authors examined the representations used by systems analysts whilst performing system design modifications. Designers were presented with one of two diagrammatic problem representations, a flow chart or an input-output process table, considered approximately equivalent in the information that could be represented. Representations depicted the current system and were accompanied by new design specifications. Experimental groups received either form of external representation whilst a control group received both. They found a significant amount of participants translated the representation from the format it was presented in into a differing, preferred format. More specifically, more than half of those given a table translated the representation to a flow chart but few given a flow chart translated their representation. In addition, the majority of the control group favoured a flowchart representation. Jones and Schkade concluded that designers are not always bound by problem representations, as some designers will choose to translate a representation into a more compatible format should the cognitive cost of doing so not outweigh the benefits.

These few studies indicate that external representations are a good way of supporting various cognitive design activities. Expert designers may have multiple ways of representing a problem and various uses for these representations, whereas novice designers may not spontaneously produce an appropriate representation.

Transfer of knowledge between design problems

The transfer of knowledge between design problems may be an effective approach to design (Chrysikou & Weisberg, 2005). Two related aspects of knowledge transfer have been studied in relation to design, analogising and fixation. Analogising is the process of transferring knowledge from prior experience and mapping it onto new problems. Ball, Ormerod and Morley (2004) examined whether experts and novices differ in their use of analogies. These authors distinguished between schema-driven and case-driven analogies. Schema-driven analogising is the automatic identification of experiential knowledge of relevance. Case-driven analogising refers to the identification of physical aspects of prior problem solutions that can be directly mapped onto the solution of a current problem, proposed to be triggered when noticeable surface similarities exist between problems. Whereas case-driven analogising results from slow and effortful analysis of prior problems, schema-driven analogising is quick and relatively effortless but may stem from familiarity with design processes. Ball *et al.* proposed that experts would be more inclined to adopt schema-driven analogising than novices and tested this by contrasting the performance of student and expert engineers whilst designing an automated car rental facility. The frequency of schema-driven and case-driven analogising indicated that expert designers engaged in more analogising than novices, with schema-driven analogising more prevalent than case-driven analogising. Novices showed the opposite pattern of analogising. Ball *et al.* proposed that a novices' application of case-based analogising is a crucial step in knowledge schematisation that may eventually lead to expert levels of design skill. Nevertheless, even experts with considerable experience may engage in case-driven analogising when a design problem is unfamiliar or unusual.

Fixation is a form of negative transfer whereby a feature of a previous design is reproduced in a current design. Purcell, Williams, Gero and Colbron (1993) suggest that the provision of pictorial examples during design may produce design fixation. These authors compared the performance of student designers when developing a measuring product. Designers were provided with either a basic, or a complex, pictorial example. Purcell *et al.* reported that those who were provided with a complex example reproduced more of the designs flaws than those provided with a simple example. They concluded that design fixation only occurs under a specific set of conditions. Chrysikou and Weisberg (2005) also examined whether the presence of design examples led to fixation. They examined psychology students performing design tasks. A fixation group were provided with an example design picture. The same pictures were given to a de-fixation group along with instructions to avoid using problematic aspects of these designs. When designing a bike rack, the fixation group produced designs considered more physically similar to the problematic example design. However, designs produced by the de-fixation group were of lower similarity to the example design than that of a control group. When designing a spill-proof cup, the fixation group reproduced the most problematic features from the example design. Chrysikou and Weisberg concluded that the inclusion of a pictorial design example can produce fixation effects, even when flawed aspects of the design are highlighted. However, if designers are instructed to avoid using problematic features of the example designs fixation effects are eliminated. Both Purcell *et al.* (1993), and Chrysikou and Weisberg (2005) demonstrate that designers who are fixated may produce less innovative and less creative designs.

Conclusions

A number of differently themed design studies have been reviewed, each adding a little to what is known about design. Many of these studies touch upon some of the cognitive features (goals, constraints, representations, alternative & solutions) of design described by Smith and Browne (1993). For instance, studies on design strategy appear to concentrate on how well designers identify and work towards design ***goals*** (Ball *et al.*, 1997; Liikkanen & Perttula, 2008), and indicate that a top-down, breadth-first strategy may produce beneficial outcomes (i.e., Lee *et al.*, 2003). ***External representations***, including sketches, aid the design process in complex design situations (Carroll *et al.*, 1980; Goldschmidt & Smolkov, 2006; Sachse *et al.*, 2001), although not all external representations may be suited to a design problem (Jones & Schkade, 1995). In regards to ***alternatives*** and ***solutions***, Kruger and Cross (2006) identified a ‘solution driven strategy’ whilst Atman *et al.* (2004) report that experts produce more alternatives than novices and better quality solutions. Carroll *et al.* (1980) are the only study reviewed to explicitly examine the implementation of design ***constraints*** in a quantifiable way. Presently, very little empirical research focuses on designers’ resolution of design constraints (although an exception is Visser’s (2004) work on the travelling salesman problem). This may be an artefact of ill-structured design, as little is known about the applicable constraints at the outset of such design tasks. Here, any constraints identified are likely the result of individual problem solving strategies. As such, the systematic manipulation of constraints in ill-structured design tasks is not possible. Indeed, a related feature of the studies reviewed here is that most refer to ill-structured design. In most experimental tasks goals were open-ended and constraints were not specified. The publication of empirical evidence concerning human designers undertaking semi-structured design is somewhat lacking. Again, Carroll *et al.* (1980) appear to be an exception. These researchers provided their designers with a list of explicit constraints and provided a more detailed design goal. This allowed a quantifiable measure of design performance in terms of the number of constraints satisfied. It is presently felt that paradigms similar to those used by Carroll *et al.* may be a suitable platform for a closer inspection of the role of external constraints in producing variability in design performance in semi-structured design contexts.

A further notable feature of studies reviewed is the predominance of laboratory-based studies. None observed designers in applied settings, although many did endeavour to encourage ecological validity by constructing realistic design tasks. The realities of design in applied settings may make the systematic manipulation of design features unproductive. Pahl *et al.* (2007) point out that many individuals may be involved throughout the various stages of design, and as such, group dynamics may complicate the isolation and identification of effective strategies or helpful cognitive aides. Another potential complication is the variation in design time spans. In addition, similar design tasks are unlikely to be repeated by multiple designers. This poses problems to researchers wishing to study designers undergoing comparable design projects. The conclusion drawn from this is that laboratory studies may be the most suitable way of studying design under controlled conditions.

1.4 Summary and overview of research objectives

The preceding sections have attempted to identify what is meant by the term design problem solving, to outline various stages of design, and to describe the level of structure within design tasks. A case was made for the notion that design activities, whilst inherently complex, are not always ill-structured. Indeed, some semblance of structure must be imposed upon a design problem in order for solutions to be derived (Fernandes & Simon, 1999; Goel & Pirolli, 1992; Restrepo & Christiaans, 2004; Schön, 1987). In addition, some current design research themes have been discussed and an avenue for potentially fruitful empirical investigation has been formed. It appears that little empirical research has looked at how human designers deal with variation in external constraints in design contexts. A preliminary search for research on design constraints produces many results. However upon closer inspection, many of the results refer to attempts to model constraint satisfaction using computerised algorithms (i.e., Mullineux, 2011; Schaerf, 1999), attempts to propose constraint classification schemes or taxonomies (i.e., Ullman & D'Ambrosio, 1995), or, the identification of undocumented constraint types in a particular context (i.e., Manz, Brunner & Wullschleger, 2006). What is lacking is empirical research of the effects of systematic variation in constraints on human design performance, accompanied by an appropriate sized sample to support statistical analysis.

This gap in the design literature is most likely due to the lack of specification in many ill-structured design tasks. Creative forms of design are characterised by the presence of few explicit constraints. Here constraints dynamically emerge as a result of continued design efforts, via the collection of more design-relevant information. However, a continuum of design structure exists (Fernandes & Simon, 1999), accommodating design tasks with an increased level of structure. Semi-structured design activities have improved structure due to the presence of more specific design goals and the specification of external constraints. Within these contexts, a closer investigation of controlled constraint variation is possible. The research thread pursued in the present thesis shall endeavour to utilise such contexts as a preliminary attempt to address the gap in the literature. More specifically, the experimental work that follows focuses on the role of constraint variation in constraint satisfaction performance. Two research questions are posed;

1. What forms of variation in external constraints lead to variation in design efficiency? (Experiments 1-5)
2. Can training methods be used to improve design performance? (Experiments 6-8)

The only question yet to be addressed is why attempt to empirically study design? Despite the difficulty in empirically investigating aspects of design such as constraints, research aimed at isolating aspects of design performance variation is a worthy pursuit. Indeed, design is a pervasive activity that permeates both formal disciplines and day to day activities (i.e., route-planning; Visser, 2004). The consequences of bad design are, at the best inconvenient, but in some situations can be costly and even dangerous. As such, endeavouring to accumulate knowledge that may shed light on aspects of design difficulty may have promising applications. Indeed, Ball *et al.* (1997) state, "the more we know about how designers design, the more we should be able to counteract ineffective design strategies by means of education, training, and computer-based support" (p 248).

CHAPTER TWO

Psychological underpinnings

2.1 Psychological theories

There are a number of theories of problem solving. Few have been extended or adapted to encompass design problem solving. However, the most applicable and widely referenced theory of human problem solving, Newell and Simon's (1972) problem space theory, has been extended to design contexts. The following section aims to outline the main features of problem space theory and discuss its application to the examination of constraint variation in constraint satisfaction contexts. Following from this, other principal theories of problem solving that are applicable to the first research thread (What forms of variation in external constraints also lead to variation in design efficiency?) are briefly addressed.

Problem space theory. Problem space theory (Newell & Simon, 1972) stipulates that problem solving activities are bounded within a problem space. Here a problem solver must transform the problem from its current state, through a series of intermediate states to the desired goal state by applying a number of operators. The problem space contains all of the intermediate state spaces possible. Therefore, problem solving is essentially a search and navigation through this problem space. Efficient problem solving is denoted by discovering the most direct route through the problem space. Intertwined with problem space theory is Simon's (1979) information processing (SIP) theory. SIP theory suggests that problem solving is essentially an information processing task. As such, problem solving success depends heavily upon the amount of task relevant information. Should a larger problem space abound, more task relevant information needs to be encoded, weighed and utilised. At the same time, the maintenance of task goals, partial solutions and progress monitoring may become a much more taxing process. As such, the cognitive load entailed in searching larger problem spaces should be expected to increase accordingly.

The application of problem space theory to design problem solving has received some criticism (Goel, 1994; Goel & Pirolli, 1992; Guindon, 1990; Purao, Rossi & Bush, 2002). Criticism is usually based on the notion that problem space theory was developed to account for well-structured problem solving. In well-structured problem solving all relevant task information is presented at the outset. Here boundaries of the problem space can be estimated, possible states are known, and a direct pathway can be surmised from the application of sound logic and known operators. However, should ill-structured design be the focus, then the lack of inherent structure is mirrored by the lack of clarity in the corresponding design problem space. Indeed there is some suggestion that multiple problem spaces may be utilised by designers in the initial stages of ill-structured design (Dorst & Cross, 2001). In contrast, this criticism of problem space theory is not so applicable within semi-structured design problem solving contexts. Here, a problem space is more closely defined. The parameters within which problem search can take place is known and the problem rules are known. As such, the problem state space is well-specified. What remains unclear in such semi-structured design spaces is the optimisation of legal operators and the derivation of a pathway to a suitable solution (note the absence of one, optimal solution). As such, problem space theory is considered to be relevant to the research that follows.

Since its conception, various extensions and amendments have suggested further properties of problem spaces. Simon and Lea (1974) proposed an extension regarding the inclusion of a rule or operator space that exists alongside the state space. They suggest that as the rule space expands, complexity in regards to the applicable legal operators increases. This increased complexity in the rule space also increases the complexity when traversing the state space. Indeed, both Simon and Lea (1974) and Kotovsky, Hayes and Simon (1985) suggest that as the number of operators increase, problem spaces becomes larger and more complex to traverse. This has implications for variation in design constraints. Whilst SIP theory would indicate that additional constraints increase the information processing load, an expansion in the rule space and rule space complexity (Simon & Lea, 1974; Kotovsky *et al.*, 1985) initiated via the introduction of more constraints, should also increase the complexity of navigating through the problem space. Reductions in design efficiency could result. This theme shall be empirically investigated in Experiments 1 and 2.

Relational complexity theory & interactivity theory. Other variation within a design rule space (Simon & Lea, 1974) may have further implications for the complexity of a design task. In semi-structured constraint satisfaction contexts, both goals and constraints are specified at the outset of a task. Therefore, when considering the transition from one problem state to another, a subset of the constraints outlined will be applicable. Therefore, the variability in terms of the qualitatively differing types of constraint that need to be considered when performing a transformation may also be a source of performance variation, regardless of the overall size of the rule (or constraint) space.

Sweller and colleagues (Paas, Renkl & Sweller, 2004; Sweller, 1994; Sweller, Chandler, Tierney & Cooper, 1990; van Merriëboer & Sweller, 2005) notion of element interactivity suggests that, in learning contexts, having to integrate information from different sources in order to perform a cognitive task may be a source of difficulty. Whilst not previously applied to design contexts, Halford and colleagues (Halford, Baker, McCredden & Bain 2005; Halford, Cowan & Andrews, 2007; Halford *et al.*, 1998) similar notions of relational complexity were formulated to be applicable to all cognitive activity, regardless of context. Relational complexity theory, building upon notions of information processing and parallel processing limitations (Schneider & Detweiler, 1987), proposes that decisions that require the processing of multiple pieces of interrelated but qualitatively distinct information are more complex. In relation to constraint satisfaction design, should multiple interrelated constraints be applicable to a proposed design move, then accurate evaluation of that design move may be problematic. As such, the successful transformation of one design problem state into another, closer to a satisfactory goal state, should be more difficult. Should this be a recurring issue within a design task, reductions in design efficiency could result. This theme shall be empirically investigated in Experiments 3 and 4.

Differing forms of constraint processing and their match with the external representation. A final source of constraint variation that may affect the complexity of a design task may be the nature of certain external constraints. It may be the case that some types of constraint are easier to implement than others. They may be less complex to process, or the nature of the processing required may be better supported by the external representation provided. Cognitive fit theory (Shaft & Vessey, 2006; Vessey, 1991; Vessey & Gellata, 1991) suggests that when problem information is well-emphasised or well-represented within an external representation, then a greater degree of cognitive fit is evident. When a good match is achieved, problem solvers can use similar cognitive strategies to process both task information, here external constraints, and also to process representational information. Indeed, the previous review of literature concerning external representations implies that the format and suitability of the external representation is crucial in supporting efficient design activities (Carroll *et al.*, 1980; Jones & Schkade, 1995). In design contexts, an implication may be that qualitatively differing forms of constraint processing may be differentially supported by the same external representation. For instance, should a design task contain spatial constraints then a suitable spatial representation would enable a good degree of cognitive fit. However, should temporal constraints be applicable, a spatial representation may not provide a good degree of cognitive fit. Whilst present research efforts do not examine variation in the external representation, the degree of cognitive fit engendered by variation in the nature of constraint processing, with a set representation is examined more closely in Experiment 5.

The rest of the experimental work will focus on training methods. As this constitutes a qualitative and theoretical shift in the examination of constraint satisfaction design, relevant theories shall be addressed later in the thesis.

2.2 Experimental paradigms

Two experimental paradigms were developed for present purposes. Each was selected and developed on the basis that no specialised knowledge would be required. In addition, the contextual familiarity should result in no need to train participants as to the nature of the task and also means that no elaborate introductory cover story need be constructed. In addition, practical considerations such as the length of the design process, and the ability to carry out data collection in a single laboratory session factored in task selection. One constraint satisfaction paradigm entailed the spatial arrangement of a hypothetical office layout (adapted from Carroll *et al.*, 1980). The other involved the creation of educational or academic timetables.

Both tasks contained some comparable or equivalent task features. Firstly, each contained a number of design elements. Each design element was a necessary feature of the completed design. Therefore, incorporation of each design element was a pre-requisite of any completed design. In the spatial office arrangement, design elements were employee offices. In the timetabling paradigm, design elements were the classes that required a slot in the schedule. Secondly, each contained a number of explicit design constraints. Constraints stipulated restrictions on how design elements should be incorporated into a design. These constraints were categorised according to a two-tier hierarchy. Low-level specific constraints stipulated a restriction on a particular design element (or elements). High-level general constraints stipulated a rule for implementing a subset of specific constraints. For instance, given the office design paradigm a general constraint may stipulate that someone high in status requires a more prestigious office location. Applicable specific constraints may indicate that one employee may be higher in status than another. An example may be, “Employee B is higher in status than employee D”. Both experimental paradigms are discussed in detail in the following sections.

2.2.1 Timetabling design task

Timetabling occurs in many practical settings. In educational contexts timetables must be carefully organised and implemented. In workplace settings work activities are scheduled around organisational resources, constraints and deadlines. Likewise, healthcare systems implement procedures to provide timely and appropriate care. These selected examples highlight the importance of timetabling for productivity and efficiency in a number of applied settings. As such, timetabling principles should be familiar to most individuals, and extensive task training should not be required. Design elements within the timetabling paradigm developed were classes that needed to be scheduled into a timetable. General constraints were presented as timetabling rules. The seven rules, developed to appear ecologically valid given the context, were as follows:

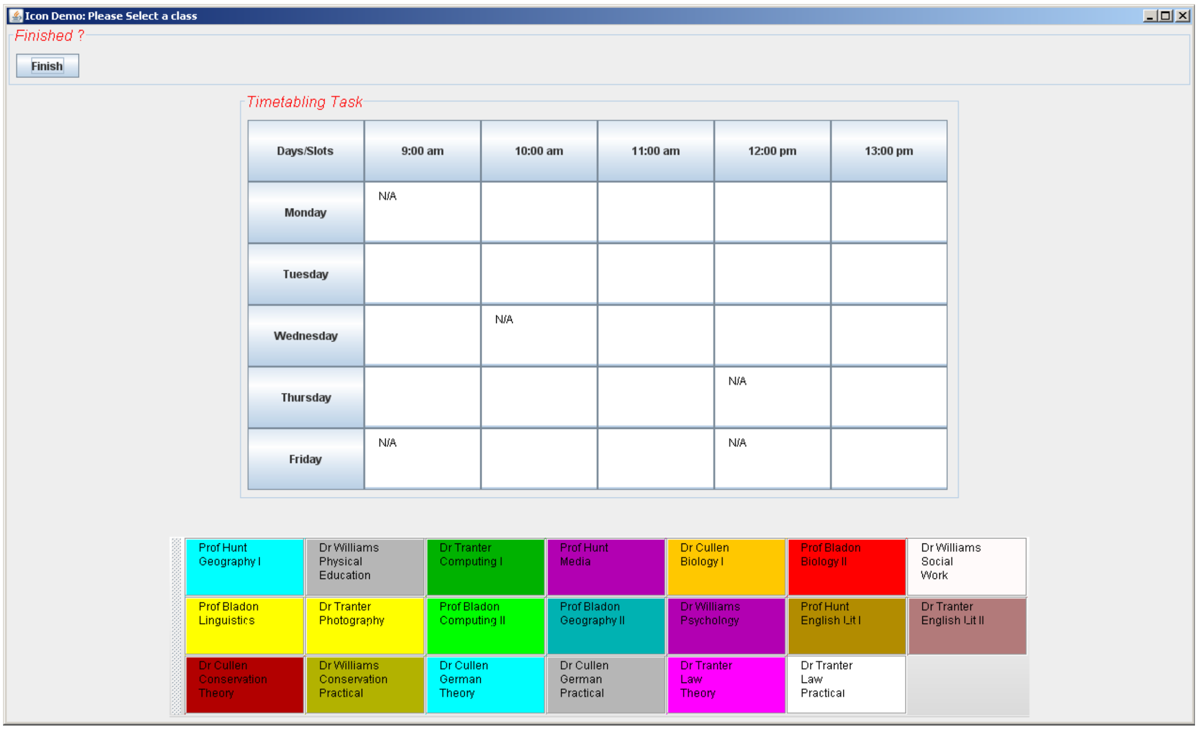
1) Classes must be scheduled in chronological order (i.e., Biology I must precede Biology II)
2) Classes must be scheduled so that theory classes precede practical classes (i.e., Music Theory must precede Music Practical)
3) The same teacher cannot be scheduled into consecutive timeslots if the classes are taking place in different locations
4) Teachers should not be scheduled for periods when they are not available
5) Rooms should not be scheduled for periods when they are not available
6) The number of students per class should not exceed the capacity of the room allocated

7) The required class facilities must be met by the room allocated to that class

Specific constraints were comprised of more exact restrictions derived from the application of relevant timetabling rules to a class (design element). For instance, should the hypothetical class *History I* have 80 students, the specific restriction derived from rule 6 is that *History I* should be scheduled to occur in a room that can accommodate this number of students. Should a *History II* be present, timetabling rule 1 (above) would result in the specific constraint that *History II* must appear later in the timetable than *History I*. A Java platform was developed that allowed a number of timetabling features, including any applicable general constraints, to be programmed and compiled into a timetabling task. Programmable features included the following:

- The timetable: The number of days and timeslots in which classes could be scheduled into were set.
- Timetabling rules (general constraints).
- Teachers: Teachers could be named and the periods in which they were available could be detailed.
- Classes: Classes were labelled. A teacher could be assigned to the class. The number of students and various required facilities could also be assigned to the class. In addition, classes could be linked to other related classes (i.e., English I, English II).
- Rooms: Class rooms could be labelled. Classes were also assigned a capacity and facilities. The times when the room was available for the scheduling of classes could also be designated.

Once compiled, tasks were presented via two computer monitors. One monitor depicted the timetable outline and a number of class tiles located towards the bottom of the screen (see Figure 2.1). Participants were able to drag class tiles into desired timeslots. Once a class had been dropped into a timeslot, participants were prompted to select a room and location in which to conduct the class via selection from a drop down menu (see Figure 2.2).



Days/Slots	9:00 am	10:00 am	11:00 am	12:00 pm	13:00 pm
Monday	N/A				
Tuesday					
Wednesday		N/A			
Thursday				N/A	
Friday	N/A			N/A	

Prof Hunt Geography I	Dr Williams Physical Education	Dr Tranter Computing I	Prof Hunt Media	Dr Cullen Biology I	Prof Bladon Biology II	Dr Williams Social Work
Prof Bladon Linguistics	Dr Tranter Photography	Prof Bladon Computing II	Prof Bladon Geography II	Dr Williams Psychology	Prof Hunt English Lit I	Dr Tranter English Lit II
Dr Cullen Conservation Theory	Dr Williams Conservation Practical	Dr Cullen German Theory	Dr Cullen German Practical	Dr Tranter Law Theory	Dr Tranter Law Practical	

Figure 2.1. Example of an interactive timetabling screen

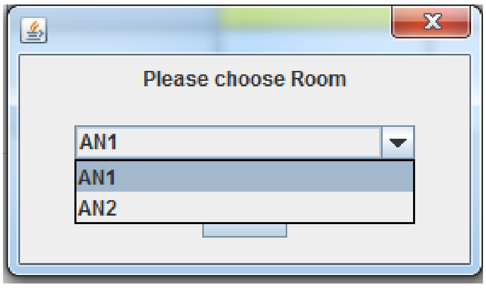
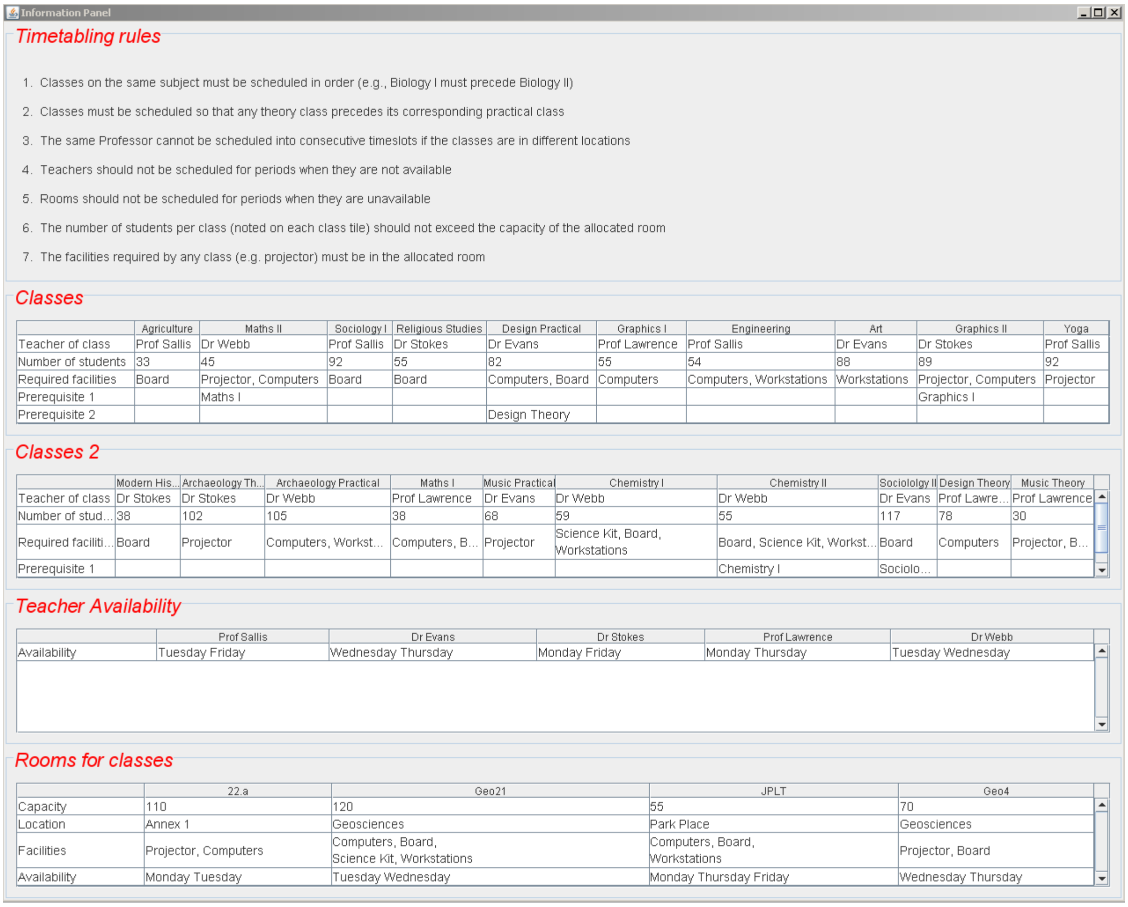


Figure 2.2. Example room selection menu

Another computer monitor was used to display other timetabling information. Figure 2.3 depicts an example information screen. At the top of this screen the timetabling rules are displayed. Beneath this, information regarding each class to be scheduled is shown. Here participants can view the class capacity and facilities required, as well as noting who teaches the class and whether there are any class prerequisites. Towards the bottom of the screen, information concerning teacher availability is displayed. Finally, at the bottom of the screen the available rooms were detailed. Here room availability was detailed along with room capacity and room facilities.



Timetabling rules

1. Classes on the same subject must be scheduled in order (e.g., Biology I must precede Biology II)
2. Classes must be scheduled so that any theory class precedes its corresponding practical class
3. The same Professor cannot be scheduled into consecutive timeslots if the classes are in different locations
4. Teachers should not be scheduled for periods when they are not available
5. Rooms should not be scheduled for periods when they are unavailable
6. The number of students per class (noted on each class tile) should not exceed the capacity of the allocated room
7. The facilities required by any class (e.g. projector) must be in the allocated room

Classes

	Agriculture	Maths II	Sociology I	Religious Studies	Design Practical	Graphics I	Engineering	Art	Graphics II	Yoga
Teacher of class	Prof Sallis	Dr Webb	Prof Sallis	Dr Stokes	Dr Evans	Prof Lawrence	Prof Sallis	Dr Evans	Dr Stokes	Prof Sallis
Number of students	33	45	92	55	82	55	54	88	89	92
Required facilities	Board	Projector, Computers	Board	Board	Computers, Board	Computers	Computers, Workstations	Workstations	Projector, Computers	Projector
Prerequisite 1		Maths I							Graphics I	
Prerequisite 2					Design Theory					

Classes 2

	Modern His...	Archaeology Th...	Archaeology Practical	Maths I	Music Practical	Chemistry I	Chemistry II	Sociololgy II	Design Theory	Music Theory
Teacher of class	Dr Stokes	Dr Stokes	Dr Webb	Prof Lawrence	Dr Evans	Dr Webb	Dr Webb	Dr Evans	Prof Lawre...	Prof Lawrence
Number of stud...	38	102	105	38	68	59	55	117	78	30
Required faciliti...	Board	Projector	Computers, Workst...	Computers, B...	Projector	Science Kit, Board, Workstations	Board, Science Kit, Workst...	Board	Computers	Projector, B...
Prerequisite 1							Chemistry I	Sociolo...		

Teacher Availability

	Prof Sallis	Dr Evans	Dr Stokes	Prof Lawrence	Dr Webb
Availability	Tuesday Friday	Wednesday Thursday	Monday Friday	Monday Thursday	Tuesday Wednesday

Rooms for classes

	22.a	Geo21	JPLT	Geo4
Capacity	110	120	55	70
Location	Annex 1	Geosciences	Park Place	Geosciences
Facilities	Projector, Computers	Computers, Board, Science Kit, Workstations	Computers, Board, Workstations	Projector, Board
Availability	Monday Tuesday	Tuesday Wednesday	Monday Thursday Friday	Wednesday Thursday

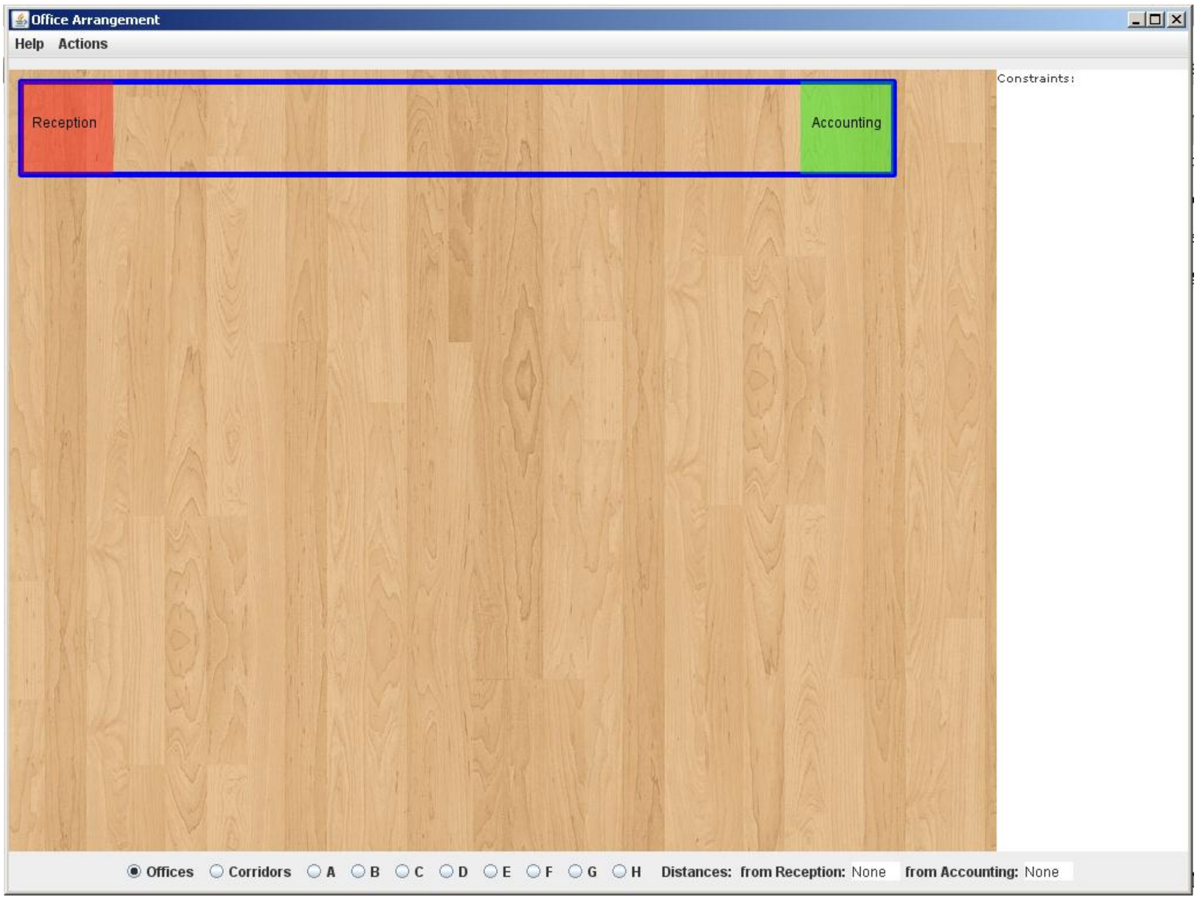
Figure 2.3. Example of a timetabling information screen

A number of dependent measures of timetabling efficiency could be extracted from any timetable design. Product measures include the number of classes scheduled successfully (those with no constraint violations), the number of constraint violations in the end design, and the time taken to complete a timetable design. Process measures include the number of design moves (defined as a class placement, class deletion or the rescheduling of a class) and the number of errors made throughout the timetabling process.

2.2.2 Office layout design task

Another form of constraint satisfaction that may occur in many practical settings is spatial arrangement design. Spatial arrangement design has been undertaken in nearly all functional spaces, both public and private, although public areas may be subject to more strict design rules. Some examples include manufacturing production lines and catering kitchens. Similarly to timetabling design contexts, the design of spatial layouts should not be completely unfamiliar to most individuals and no elaborate task cover story should be required. The present office design paradigm was adapted from one used by Carroll *et al.* (1980). Design elements here were the employees' offices. As in the timetabling task, the general constraints here were rules for the successful implementation of lower-level specific constraints, introduced via paper instructions studied before the presentation of the design task. The specific constraints, that appear upon commencement of each task, are binary in nature, as they specify a restriction on the location of one employees office relative to that of another employee. For example, one of Carroll *et al.*'s (1980) original arrangement rules was that employees who use a particular work area more often should have an office positioned closer to that area. A specific constraint related to this rule might be that employee B uses the accounting area more often than employee F.

All office design tasks were programmed in Java. Each task allowed participants to view a number of specific constraints and a set of seven or eight employees to be arranged into office spaces. Differing sets of employees could be selected when necessary. Also, differing sets of specific constraints could be uploaded to a task via a text file. Movable interface components allowed participants to construct corridors and offices, and also to assign employees to offices within the workspace. On first presentation, each office task program contained a representation of an office floor plan from a bird's eye view as depicted in Figure 2.4. This representation included a main corridor for the office complex, with a reception and accounting department at either end. The office space was contained within a set space, indicative of boundary walls. Dependent measures of design efficiency within this paradigm include the product measures; number of specific constraints satisfied and the time taken to complete the design. The process measure, number of design moves can be calculated via review of design videos.



Distances: from Reception: None from Accounting: None

Figure 2.4. Example of office design task upon task commencement

CHAPTER THREE

Number of constraints

3.1 Do more constraints lead to greater design difficulty?

In order to successfully transverse the problem space and formulate a solution, information from various sources needs to be assimilated into a meaningful problem representation. Within both experimental paradigms outlined in the previous chapter, much of the relevant information is provided via a task goal and a number of external constraints (both specific and general constraints). Indeed, in constraint satisfaction paradigms, external constraints may make up the bulk of all information to be processed. As such, variation in the number of external specific design constraints should be expected to impact upon design cognition and subsequent design performance.

Presently, little empirical research attests to the impact of variation in constraint numbers on human design performance. At the ill-structured end of the design spectrum, where goals are vague, and constraints are not explicit at the outset, the dynamic and changeable nature of the task does not lend itself to the systematic manipulation of constraints. Despite the lack of evidence, there is some suggestion that further constraints serve to reduce the number of viable design possibilities, as they may be beneficial in setting parameters, narrowing down the problem space and restricting design search to more viable design pathways (Besnard & Lawrie, 2002; Visser, 2006). Design tasks with a greater degree of structure, such as constraint satisfaction problems, allow more strategic variation of constraints whilst retaining some important features of design tasks (dynamic transformations, no set pathway to solution and no known best solution). In such contexts, it is unclear whether more constraints act to limit the problem space in a helpful manner. Indeed, as the problem space is already defined, additional constraints may constitute an increase in task complexity and also an increase in cognitive workload. The following discussion aims to outline a number of theoretical reasons for speculating as to why the presence of more specific constraints may induce deterioration in performance within constraint satisfaction design.

Newell and Simon's (1972) influential problem space theory details human problem solving as a search or navigation through possible problem states. Operators (rules or constraints) act to transform one problem state to another. Further extensions to problem space theory distinguish between a state space outlining all possible problem configurations, and a rule space outlining the constraints restricting the transformation from one problem state to the next (Burns & Vollmeyer, 2002; Simon & Lea, 1974; Zhang & Norman, 1994). Changes in the rule space necessarily impact upon search within the state space (Simon & Lea, 1974). Kotovsky *et al.* (1985) theorised that the larger the problem space associated with a task, the more difficult a task will be. Here it is proposed that the greater the number of constraints, and therefore the larger the rule space, the more complex navigation of the state space becomes. In the relatively more researched field of traditional, structured problem solving, there is evidence to attest to the notion that larger constraint/rule spaces are a source of problem difficulty. An example is the work of Davies (2003) who found that problem solving on the Tower of Hanoi (ToH) became more complex and required more careful planning when the number of discs and corresponding constraints increased.

Simon's (1979) information processing theory may offer further insight into why greater numbers of constraints may lead to greater difficulty and performance deterioration within constraint satisfaction problems. This approach implies that much of the difficulty involved in any task is a result of the amount, and complexity, of information to be processed. In unfamiliar tasks, designers have to rely heavily on working memory to process all relevant information. Larger numbers of specific constraints will impose greater load on working memory. As working memory is known to have a limited capacity (Baddeley, 1992;

Miller, 1956), at some point information to be processed will surpass working memory capacity. At this point a designer may attempt to expend additional cognitive effort to maintain design standards by encoding design information (Newell & Simon, 1972). Alternatively, designers may be unable or unwilling to deploy memory resources to aid performance. In this case, deteriorations in various design performance measures could be expected. However, there is reason to suggest that designers will not choose to encode design information. Recent research in human-computer interaction indicates that a more cognitively-intensive, memory-based strategy is unlikely to be adopted in problem situations when a computer continually displays all relevant problem information. In these instances, a display-based perceptual strategy is usually adopted as this requires the least cognitive effort on the part of the problem solver (Anderson, 1996; Gray, Sims, Fu & Schoelles, 2006; Morgan, Patrick, Waldron *et al.*, 2009; Waldron, Patrick & Duggan, 2010). However, this strategy is likely to come at the cost of reduced memory of task progress (Waldron, Patrick, Morgan & King, 2007). In constraint satisfaction situations, such as those outlined previously, this may translate into a reduced ability to monitor design progress and may ultimately result in unmet or forgotten constraints. As such, a greater number of specific constraints accompanied with greater working memory load should lead to reduced design efficiency.

In summary, both problem space theory (Newell & Simon, 1972), it's extensions (Burns & Vollmeyer, 2002; Kotovsky *et al.*, 1985; Simon & Lea, 1974; Zhang & Norman, 1994) and Simon's (1979) information processing theory imply that additional specific constraints applied to a constraint satisfaction problem will strain mental resources. The following experiments aim to establish whether increasing the number of constraints taxes cognitive resources to such a point so as to affect design performance.

3.2 Experiment 1

The previous discussion of human information processing (Simon, 1979) and problem solving (Newell & Simon, 1972; Kotovsky *et al.*, 1985) concluded that the presence of a greater number of specific constraints, in a constraint satisfaction context, should increase task complexity and subsequently lead to performance deterioration. This prediction is now examined using the timetabling paradigm, described in the previous chapter (Section 2.2.1).

In the experiment that follows, participants are required to complete three timetable design tasks. Tasks vary in the number of specific constraints applied. Performance measures obtained include product measures such as the number of classes successfully timetabled, the errors (constraint violations) in the end design and the time taken to complete the timetable. Process performance measures will entail the number of errors made throughout the design process and the number of design moves utilised. It is expected that increasing the number of specific constraints applicable within a task will lead to deterioration on all performance measures. Therefore, tasks with more specific constraints should evidence fewer successfully placed classes, more errors (constraint violations both at the end of the design process, and throughout the design process), and also longer design completion times and the utilisation of a greater number of design moves.

3.2.1 Method

Participants

Sixty psychology undergraduates, aged between 18 and 21 years, with an average age of 18.97 years (SD = .76), took part in the experiment in return for course credit. The sample consisted of 55 females and five males. None had any experience with the experimental task.

Materials

Experimental materials consisted of three timetabling tasks (see section 2.2.1), displayed via two interfaces (a timetabling screen and an information screen; see Figures 2.1 & 2.3) presented on two adjacent computer monitors. The timetabling screen contained an empty five day (monday - friday), five timeslot (9am-1pm), timetable. Towards the bottom of this screen, a resource window displayed the 20 classes to be scheduled into free timeslots. The information screen displayed further information concerning the attendees and requirements of each class, the availability of classrooms with various facilities and the availability of the teachers. This screen also displayed the seven timetabling rules detailed in Section 2.1. Tasks varied in the number of general constraints (here the applicable timetabling rules) and also in the number of derivable specific constraints, in order to produce a Low, a Medium, and a High constraint task. Table 3.1 displays the breakdown of constraint types across task.

Table 3.1. Breakdown of general and specific constraints across timetabling tasks

	Low constraint task	**Medium constraint task**	**High constraint task**
No. of applicable general constraints	3	5	7
Exact general constraints applicable	Rules 2, 4, 7	Rules 1, 2, 4, 5, 6	All rules (1, 2, 3, 4, 5, 6, 7)
Specific constraints †	48	67	97

† - as calculated by the maximum number of specific constraint violations that may be incurred in each task

Design

A within-subject design was used such that each participant performed all three timetable tasks. The tasks differed in the number of general and specific constraints applicable. As such, the independent variable, with three levels, was the number of constraints within the task; Low (3 general constraints producing 48 specific constraints), Medium (5 general constraints producing 67 specific constraints), and High (7 general constraints producing 97 specific constraints). The presentation order of tasks was counterbalanced across participants. Dependent measures were the number of successful class placements (classes not incurring any constraint violations) out of the maximum 20, time taken to complete the timetable design (seconds), and the total number of constraint violations in the end design. Process measures include the number of constraint violations incurred and the number of design moves used throughout the design process.

Procedure

Participants were given general instructions describing the nature of the timetabling task environment, the nature of the information on the two screens available during each timetabling task (including the possible rules constraining class placement) and how to drag and drop classes into timetable slots. These instructions were followed by a practice task in order to familiarise participants with the interactive elements of the interface. This involved copying a mini-timetable of 12 classes by dragging and dropping classes into the correct timetable slots. Instructions then explained that participants were required to schedule 20 classes, all an hour in duration, into 5 days (Monday to Friday, between 9am & 1pm), and that they should do so as quickly and accurately as possible whilst bearing in mind that not all timetabling rules might be applicable. Following this, participants completed the three experimental timetabling tasks according to the order they had been assigned. No time limit was given to complete the timetabling tasks.

3.2.2 Results

Thirty outliers were present within the dataset, constituting 3.33% of all data points. Outliers were identified by converting data points into z-scores and removing data points with a z-score outside of the ±3.28 range. The process of converting data into z-scores was repeated on the remaining data points in order to identify any further outliers. After three iterations, no additional outliers were identified. All outliers were then replaced with grand means in order to avoid a reduction in sample (Field, 2009). Table 3.2 displays the mean scores for all experimental performance measures. Means suggest that as the number of constraints increase, performance on product measures deteriorate. A similar trend was evident for the process performance measures with the exception of the number of successful class placements. Nevertheless, within this measure, those with the greatest number of constraint rules obtained the lowest scores. A number of within-subject one-way analyses of variance (ANOVAs) were conducted in order to determine if the number of constraints significantly impacted on design performance measures.

Table 3.2. The effect of number of constraints on timetabling design performance

	Measures of performance		**Low constraints**	**Medium constraints**	**High constraints**
Product measures	Successful class placements	Mean	13.67	14.34	12.10
		SD	5.34	4.07	4.06
	End violations	Mean	7.32	6.55	10.44
		SD	6.50	4.71	5.64
	Design time (seconds)	Mean	562.61	687.49	757.27
		SD	127.48	238.11	223.22
Process measures	Total constraint violations	Mean	10.46	12.40	17.72
		SD	8.39	8.76	8.66
	Design moves	Mean	26.08	28.13	28.97
		SD	5.17	6.13	7.82

Product measures

In regards to the number of classes successfully placed within a timetable there was a significant effect of constraint numbers, *F* (2, 118) = 7.57, MSE = 10.43, $p < .001$, $f = .63$. βonferroni comparisons revealed that significantly fewer classes were correctly placed in the high constraint task than in the medium ($p < .001$) or the low constraint task ($p < .05$), with no significant differences between the latter two tasks ($p = .94$).

The number of constraint violations in the end design also indicated a significant effect of constraint numbers, *F* (2, 118) = 13.85, MSE = 18.44, $p < .001$, $f = .48$. βonferroni comparisons indicated more constraint violations for high constraint tasks in comparison with both the low and medium constraint tasks (*p*s < .001 in both instances). No difference was apparent between the low and medium constraint tasks (*p* = 1.00).

An effect of number of constraints on design completion time was also found, *F* (2, 118) = 16.30, MSE = 35796.28, *p* < .001, *f* = .52. βonferroni post-hoc comparisons indicate that low constraint tasks were completed significantly quicker than both medium (*p* < .01) and high constraint (*p* < .001) tasks. The latter tasks did not significantly differ on completion time (*p* = .14).

Process measures

In respect to the measures obtained during the design process, there was a significant effect of constraint numbers on the number of constraint violations incurred throughout designing, *F* (2, 118) = 16.89, MSE = 50.21, *p* < .001, *f* = .54. βonferroni comparisons revealed that the high constraint task incurred more errors throughout the design process than both the low and the medium constraint tasks (*p*s < .001 in both instances). No apparent differences were found between the low and medium constraint task (*p* = .29).

Finally, the number of design moves utilised throughout the design process was also affected by the number of task constraints, *F* (2, 118) = 3.97, MSE = 33.52, *p* < .05, *f* = .26. βonferroni comparisons indicated that fewer moves were needed to complete a low constraint design task in comparison to high constraint task (*p* < .05). Moves on a medium constraint task did not significantly differ from either the low (*p* = .14) or the high constraint task (*p* = 1.00).

3.2.3 Discussion

The present results indicate that when designing a timetable, increasing the number of explicit constraints, here general constraints in the form of timetabling rules, and the specific restrictions that can be derived from them, has an effect on design efficiency. This was manifested in deterioration of both product and process measures as hypothesised. In respect to product measures, greater numbers of explicit constraints resulted in the use of more design time, in the increased number of constraints violated in the end design, and also in fewer successful class placements. The number of successful class placements, arguably the most important indicator of design effectiveness, did not evidence a consistent linear trend of deterioration as the constraint numbers increased. Nevertheless, performance in the high constraint task was worse than the other experimental tasks. In respect to process measures, the high constraint task produced significantly more errors throughout the design process than other experimental tasks. In addition, the number of design moves utilised was significantly higher in the high constraints task in comparison to the low, but not the medium constraints task. Overall, performance in the high constraint task was consistently worse on all dependent measures.

The present results support experimental predictions and can be explained by the notion that an enlargement of the constraint space (Simon & Lea, 1974) may increase the complexity of the overall problem space (Zhang & Norman, 1994). The results are also consistent with notions of greater numbers of constraints leading to increased information processing demands, and therefore also increased cognitive load when incorporating more timetabling rules into the search through problem space (Simon, 1979). In regards to the suggestion that greater numbers of design constraints may act to more effectively direct design efforts (Besnard & Lawrie, 2002; Visser, 2006), it appears that whilst this may be the case in more ill-structured, creative design, it may not be the case for more structured forms of design such as constraint satisfaction tasks. In these instances, the extent of the design space is not unknown. Therefore, further restrictions imposed by increasing the number of constraints may act to restrict the number of viable design options but, at the same time, may not prove helpful in narrowing the design search or reducing the problem space.

A particular feature of the present experimental design, which warrants highlighting, was that throughout all timetabling tasks all seven general constraints (timetabling rules) were displayed. This was implemented in order to preserve perceptual equality between the three tasks, as removing irrelevant timetabling rules would have led to noticeable differences in task layout (by altering the layout of the information screen, see Figure 2.3). This may have had the unintentional effect of making the low and medium constraint tasks more difficult, as participants still had to consider all timetabling rules and then identify those that were relevant to each task. This acts as a more stringent test of the experimental hypothesis as the presence of all timetabling rules may have suppressed some of the between task differences. Had only the task relevant general constraints been displayed, then the differences in design performance observed between conditions may have been greater.

Another notable feature of the present design is that the differing number of general constraints applied per task did not result in equivalent increases in specific constraints. The three general constraints in the low constraint task produced 48 specific constraints, the five general constraints in the medium task produced 67 specific constraints, and all of the timetabling rules in the high constraint task produced 98 constraints. This differential increases in constraint numbers is likely to have resulted in a differential increase in task complexity (as more general constraints are applied). As such, the smaller difference in constraint numbers between the low and the medium constraint tasks may have made these tasks more alike in comparison to the high constraint task and may explain why, for the majority of dependent measures obtained, no significant differences between the low and medium constraint tasks were evident. Nevertheless, an exception to the trend, a significant difference in the time taken to complete a low versus a medium constraint design task, is indicative that these tasks were not necessarily equivalent in complexity. It may be the case that the additional time taken to complete a medium constraint task (in comparison to a low constraint task) may have allowed participants to maintain the level of performance on other dependent measures. This notion is speculative and cannot be confirmed without replicating the design in which a suitable time limit is imposed on all tasks.

A few experimental limitations should be acknowledged. Firstly, the present experiment needs to be interpreted in the light of the timetabling rules (the general constraints) applied to each task. Just one version of the low and medium constraint task was used, as there was a limited number of timetabling rules that could be designed out of a task whilst still having all rules displayed on the information screen (note, all timetabling rules were displayed during all tasks). Therefore, it may be the case that idiosyncrasies surrounding certain rules may have contributed to some of the performance differences between the tasks. One such idiosyncrasy may be that some of the general timetabling constraints may refer to the relative scheduling of two classes (timetabling rules 1-3), rather than stipulating restrictions on one class (rules 4-7). This limitation is difficult to overcome in a within-subjects design. Ideally, provision should be made for a between-subjects replication to incorporate two differing versions of the low and medium constraint tasks (with differing subsets of the timetabling rules). A further limitation is that two aspects of constraint numbers, the number of general and the number of specific constraints co-varied. The present results and conclusions do not distinguish between the differing forms of constraint variation. Untangling these timetable design effects, although potentially beneficial, would not be straightforward and may require a revision of the current experimental paradigm. A final limitation was the sample. Participants were psychology students without any specific training in design. Therefore, participants would have had experience of the execution of timetables and may have some notion of effective principles for constructing them. Nevertheless, participants were not considered experts so the generalisability of results to designers with more extensive design experience may not be appropriate. Nevertheless, the findings reported here allow insight into what may constitute an important aspect of design difficulty for designers without formal training. It is further expected that the performance decrements found in the present study may be experienced in other constraint satisfaction tasks. These may include, but may not be limited to, manufacturing process scheduling, spatial layout design, and also computer programming. Further empirical research should examine this suggestion.

To summarise, the results from the present experiment indicate that there may be a threshold beyond which increasing the number of explicit design constraints, via changes in timetabling rules (general constraints), and the specific constraints which can be derived from them, leads to a more complex problem space. The search or navigation through these more complex problem spaces exert greater cognitive load upon designers, and lead to performance decrements. In this instance, when all timetabling rules were applicable to a task, performance was worse on all experimental measures. Further research should aim to establish whether similar effects are found in other constraint satisfaction design contexts. The experiment that follows aims to replicate some of the effects found here in a different constraint satisfaction context, namely spatial office layout design.

3.3 Experiment 2

Experiment 1 indicated that having more constraints in a timetabling task reduced efficiency as indicated by performance deterioration on a number of dependent measures. The present experiment aims to investigate whether increases in constraints in an office layout task results in similar design performance decrements. As previously outlined, theories of information processing (Simon, 1979) and problem space (Burns & Vollmeyer, 2002; Newell & Simon, 1972; Simon & Lea, 1974; Zhang & Norman, 1994) indicate that having more information to incorporate, increases task complexity. This is expected to impact on performance regardless of the context of the constraint satisfaction task. As such, experimental predictions are that more constraints should result in deterioration in design performance.

A subsidiary aim of the present experiment is to overcome some of the methodological issues surrounding Experiment 1. The use of differing timetabling rules, constituting different general constraints, raised the possibility that idiosyncrasies between differing rule sets may have affected results. In addition, number of constraints (specific constraints) was indistinguishable from the number of constraint types (general constraints) in that both co-varied together. Such limitations affected the confidence with which conclusions were drawn. Presently, a between-subjects design with specific constraints explicit at the outset of the design task proffers more experimental control and greater ease of interpretation. The present study also offers an opportunity to test whether some of the effects reported in Experiment 1 generalise to another constraint satisfaction paradigm, namely office design.

In the experiment that follows, groups of participants completed either a low, medium or high constraint office design task, whereby varying numbers of specific constraints were evenly divided amongst three general constraints, here rules restricting the positioning of employee offices. Performance measures obtained include the proportion of constraints satisfied in the end design, time taken to complete a design, and the proportion of time relative to each constraint satisfied. It is expected that having a greater number of design constraints will increase the complexity of the problem space, increase greater cognitive load and result in deteriorations in design performance. As such, an office design task with more constraints should result in a lower design scores, longer design completion times, more time devoted to each constraint satisfied and more design moves.

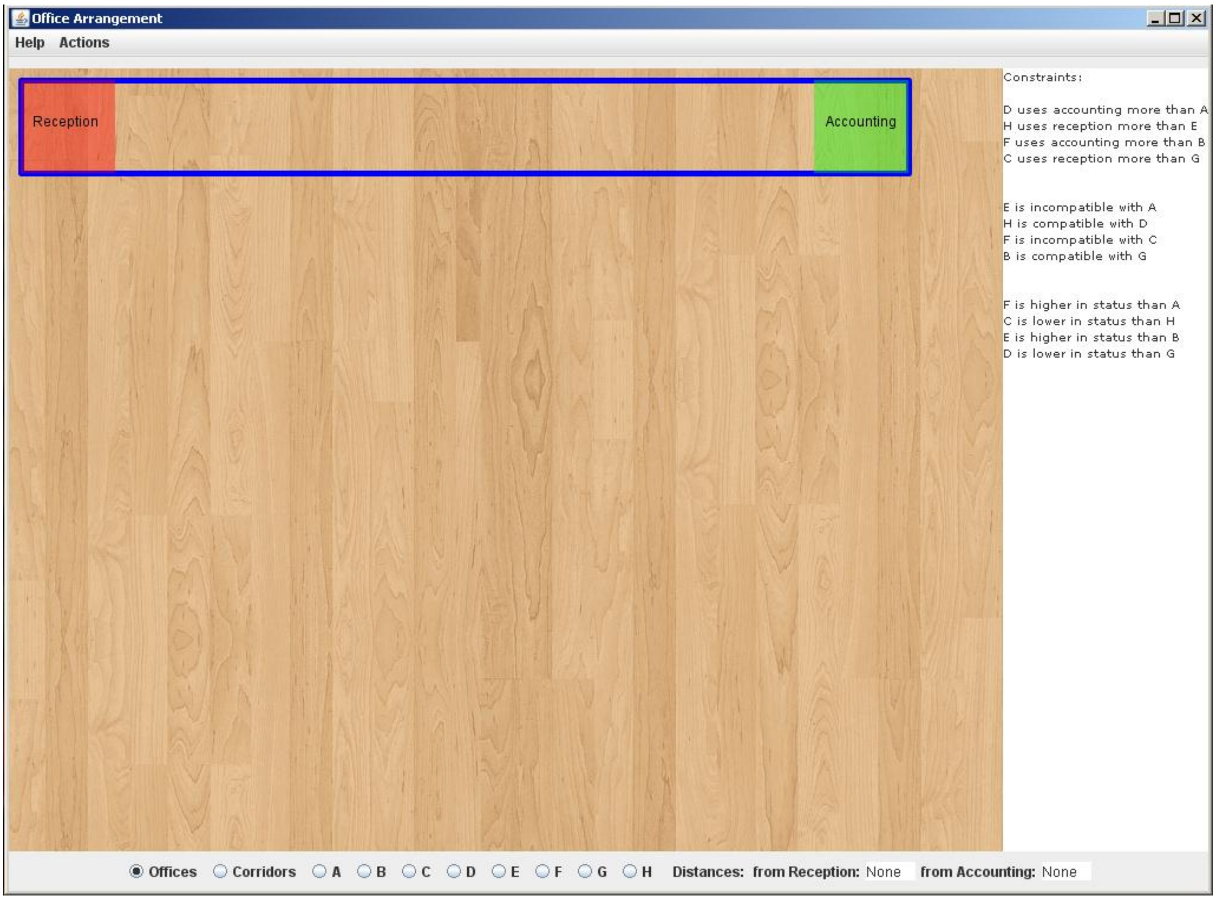
3.3.1 Method

Participants

Forty-five psychology undergraduates, aged between 18 and 26, with an average age of 19.00 (SD = 1.38) took part in the experiment in return for course credit. The sample consisted of 43 females and two males. None had experience with the experimental task. Participants were randomly allocated to one of the three experimental conditions.

Materials

Experimental materials consisted of three computerised office layout tasks adapted from Carroll *et al.* (1980) as described in Section 2.2.2. On first presentation, each office task program contained a representation of an office floor plan from a bird's eye view. This representation included a main corridor for the office complex, with a reception area at the west end and an accounting department at the east (see Figure 3.1).



Constraints:

D uses accounting more than A
H uses reception more than E
F uses accounting more than B
C uses reception more than G

E is incompatible with A
H is compatible with D
F is incompatible with C
B is compatible with G

F is higher in status than A
C is lower in status than H
E is higher in status than B
D is lower in status than G

● Offices ○ Corridors ○ A ○ B ○ C ○ D ○ E ○ F ○ G ○ H Distances: from Reception: None from Accounting: None

Figure 3.1. Initial screen on a low constraint office layout design task

All tasks were programmed in Java and were designed to allow participants to view a set of either 12, 24 or 36 specific constraints in the form of binary relationships between fictional employees (labelled A - H) that appeared once the task commenced. Each task encompassed three general constraints. These were:

- Compatibility; employees who are compatible should have adjacent offices on the same corridor but employees who are not compatible should not. An example of a specific constraint derived from this is, “Employee B is incompatible with employee G”.
- Status; employees who are higher in status should have an office closer to the central corridor. A corresponding specific constraint may be, “Employee A is higher in status than employee F”.
- Work area; should an employee use a work area more often than another, they should have an office positioned closer to that area. A corresponding specific constraint might be, “Employee G uses the reception area more than employee B”.

Repetition in binary constraints between the same two employees never occurred within the same general constraint or was avoided when possible (that is, in the low and medium constraint tasks). Within each task, movable interface components allowed the construction of offices and corridors and the relocation of employees in the workspace. CamStudio was used to record on screen actions whilst participants produced their office designs.

Design

A between-subjects design with one independent variable, number of constraints was used. This variable had three levels; low (12 specific constraints), medium (24 specific constraints), and high (36 specific constraints). Specific constraints were divided equally amongst the three general constraints (see Table 3.3). Dependent variables measured include the output measures; proportion of task constraints satisfied (in order to control for the differing overall number of constraints between experimental conditions), time taken to complete the office design (seconds), and the length of time taken to satisfy each constraint (seconds). The process measure, number of design moves utilised, is also examined.

Table 3.3. The number of specific constraints applicable to each design element

	Low constraint Task	**Medium constraint Task**	**High constraint task**
General constraint 1	1	2	3
General constraint 2	1	2	3
General constraint 3	1	2	3
Total specific constraints	3	6	9

Procedure

All participants were familiarised with the office design interface and instructed how to generate design features by copying, and then altering, a mini office arrangement containing four employee offices from paper instructions. Participants were then asked to read through instructions detailing the constraints they would be implementing and instructed that they should complete the task as quickly and accurately as they could. CamStudio recording began as soon as experimental constraints were displayed. No finish time was set. Participants terminated performance themselves.

3.3.2 Results

Two participants failed to understand task instructions and complete the task correctly so their data was deleted from subsequent analysis. Table 3.4 displays the mean scores on each dependent measure as a function of number of task constraints. Mean scores suggest that those with fewer task constraints generally satisfied a higher proportion of those task constraints and that they completed their designs quicker. However, they spent proportionally more time satisfying each constraint.

A number of between-subjects, one-way ANOVAs were carried out in order to establish whether differences were significant. Firstly, a significant effect of number of specific constraints on the proportion of task constraints satisfied was found, *F* (2, 40) = 9.66, MSE = 76.97, *p* < .001, *f* = .70. βonferroni post-hoc comparisons revealed a significantly higher proportion of constraints were satisfied in the low constraint task in comparison to the high constraint task (*p* < .001) but not the medium constraint task (*p* = .62). Those performing the medium constraint task satisfied a significantly higher proportion of their constraints than those performing the high constraint task (*p* < .01).

Table 3.4. The effect of number of specific constraints on office design performance

Measures of performance		**Low constraints (n = 13)**	**Medium constraints (n = 15)**	**High constraints (n = 15)**
Proportion of constraints satisfied	Mean	.82	.78	.68
	SD	.09	.11	.06
Design completion time (seconds)	Mean	656.69	833.53	1026.40
	SD	168.54	199.21	223.51
Time taken per constraint satisfied	Mean	66.34	44.72	42.11
	SD	14.02	12.19	9.20
Number of design moves	Mean	43.77	47.40	63.33
	SD	15.21	22.71	23.35

A significant effect of the number of specific constraints was also found when examining time taking to complete a design, *F* (2, 40) = 12.06, MSE = 39896.80, *p* < .001, *f* = .78. βonferroni comparisons indicated that those tackling the high constraint task took significantly longer to complete their designs than those tackling the low and or the medium constraint tasks (*p*s < .001, & < .05 respectively). Design time between the low and the medium constraint tasks did not differ significantly (*p* = .10). In addition, as overall design times varied a proportional time measure, time divided by the number of constraints satisfied, was examined to give a crude indication of the rate of design progress. A significant effect of the number of constraints was found on time taken to satisfy each constraint, F (2, 40) = 17.13, MSE = 140.60, $p < .001$, $f = .92$. Further βonferroni comparisons revealed that those performing the low constraint task took significantly longer, proportionally, per constraint satisfied, than participants in both the medium and high constraint conditions (ps < .001 in both cases). Time per constraint did not differ significantly between the medium and high constraint tasks ($p = 1.00$).

Finally, a significant effect of the number of specific constraints was found on the process measure, number of design moves used (F (2, 40) = 3.54, MSE = 440.68, $p <.05$, $f = .42$). βonferroni comparisons indicated that the difference between the number of design moves utilised in high and low constraint conditions approached significance ($p = .06$). No further differences were found.

3.3.3 Discussion

Results imply an important effect of number of specific office design constraints on multiple measures of design performance. Those undertaking a high constraint task satisfied a significantly smaller proportion of their task constraints and took longer to complete their design than those faced with a medium or low constraint task (who did not differ significantly). In addition, participants in the high condition used more design moves to complete their design than those in the low constraint condition. Whilst these results do not indicate a consistent linear increase in design difficulty, a threshold effect is apparent. It appears to be the case that having to process 36 specific constraints taxes the mental resources of designer's to the extent that comparative performance is worse than in all of the other experimental tasks. In contrast, the doubling of the number of specific constraints between the low and the medium constraints task did not appear to constitute a great enough increment to produce significant deterioration in performance. These results provide support for the notion that the presence of a large number of specific constraints leads to greater task complexity. The need to process and integrate additional information into an already complex problem space results in greater levels of cognitive load (Simon, 1979) and ultimately, lower standards of design performance.

Further results concerning the amount of time devoted to satisfying each task constraint do not support experimental predictions. The amount of time devoted to satisfying each task constraint was calculated in order to provide a crude indication of the rate of progress. Contrary to expectations, those completing an office design with just 12 constraints took longer to satisfy each constraint than other groups. Whilst it could be expected that the fewer constraints imposed in the 12 condition would make satisfying each constraint easier, due to a smaller and less complex problem space (Newell & Simon, 1972), coupled with a lower amount of information to process (Simon, 1979), this measure indicated that rate of progress was actually slower. Whilst seemingly problematic, this result could be interpreted in a number of ways. It may be the case that having just 12 constraints corresponds to an easy task where cognitive load is relative low. As such, it may be that participants felt they had a good chance of improving the design further, coupled with the spare cognitive processing capacity to do so. This may have led participants to work on the constraints displayed for longer (proportional to number of constraints in each task). A related explanation may be that participants completing tasks with just 12 constraints may have not felt time pressures as keenly as participants completing other experimental tasks. This may have caused there rate of progress to slow down. Alternatively, a methodological limitation may be responsible whereby participants continued designing in an effort to seem to have done enough work/participation in return for the course credit on offer. Participants were advised that their participation slot would be up to 30 minutes. Whilst experimental instructions urged them to complete their task as quickly and accurately as possible, participants were notified, upon booking an appointment to participate, that this may take up to 30 minutes. No definite conclusions can be made as to why rate of progress appeared slower for participants faced with fewer task constraints. Further replications could involve analysing design videos in order to examine why rate of progress may appear slower in tasks with fewer constraints.

Some further methodological considerations need to be acknowledged. Firstly, just as in Experiment 1, the current sample were psychology students. Whilst this may be considered a homogenous sample, it is not expected to greatly affect the results as processing capacity should not differ greatly from a more diverse sample. Secondly, all participants were experimentally naive and had no formal design training. However, it is expected that participants would have some experience, or knowledge of, spatial layouts arrangement. Therefore, the effects found presently may not generalise to expert designers. In addition to the above considerations, there were some methodological advantages of the present design worth highlighting. In Experiment 1 the explicit constraints, timetabling rules, were not necessarily applicable to every design element (class). Also, the differing numbers of timetabling constraints necessitated that both the number and the nature of the timetabling constraints varied between tasks. In the present experiment, a between-subjects design allowed all three office layout rules to be applied in every task. This allowed control over the number of both general and specific constraints, and also afforded the construction of design tasks whereby each general constraint applied to each design element.

In summary, the present results offer further support for the notion that there may be a threshold beyond which increasing the number of explicit design constraints, leads to substantial increases in problem space complexity. This results in deterioration in important measures of performance efficiency, here the proportion of constraints satisfied and the time taken to complete a design. The time spent satisfying each constraint successfully implemented did not follow a trend for deterioration in performance with an increasing number of specific task constraints. Reasons for this were not clear-cut and no firm conclusions could be made. Further research could focus more closely on investigating whether strategic differences exist between those undertaking a task with just 12, rather than 36 constraints.

3.4 Conclusions

Both Experiment 1 and 2 demonstrate that the presence of more explicit design constraints result in deterioration in design performance, as assessed by various measures of design efficiency. Both Experiments 1 and 2 compared and contrasted performance in three tasks with differing levels of explicit constraints. Whilst strict linear deteriorations were not apparent with the increasing numbers of constraints, there was considerable evidence to suggest that there was a threshold effect whereby completing a constraint satisfaction task with the highest numbers of constraints, resulted in many performance disadvantages.

Theoretical explanations of results revolve around notions of enlarged rule spaces (Burns & Vollmeyer, 2002; Simon & Lea, 1974), and increased problem space complexity (Newell & Simon, 1972; Zhang & Norman, 1994). Alone, the integration of more information into a designer’s task representation would increase the cognitive load of the designer, but when also coupled with greater complexity when searching within that problem space, greater numbers of explicit constraints appear to tax designers’ cognitive resources enough to produce deterioration in performance when compared to similar tasks with fewer constraints. These results, whilst not groundbreaking, are an initial endeavour to fill in the literature gap concerning human performance on constraint satisfaction design tasks. Conclusions here confirm similar effects found in more structured problem solving contexts such as Davies (2003) work with the Tower of Hanoi, where the presence of additional task constraints resulted in longer task completion times and more task errors.

Results contradict notions that additional constraints in design contexts may act to narrow the problem space and more effectively direct design search (Besnard & Lawrie, 2002; Visser, 2006). It is presently suggested that these actions are perhaps more applicable to more ill-structured, creative design, where the problem space is not bounded at the beginning of the design process. In constraint satisfaction paradigms, the extent of the problem space is much clearer from the task specifications, the external representation format and the explicit constraints supplied. Here, further constraints may still act to limit design options but may not narrow down the various design options effectively.

In regards to the experimental paradigms used, the office layout task appeared to offer more experimental control over variables of interest than the timetabling task. The effects of the number and nature of timetabling constraints were not clearly distinguishable as the number of general and specific constraints co-varied. The office task allowed more controlled and systematic variation in lower level, specific constraints. Another pragmatic difference between the office design and timetabling paradigms is the nature of the constraints. The office task involved binary constraints, all of which specified an interrelation between two design elements. Whilst the timetabling task incorporated some constraints indicating a relation between two design elements (i.e., Biology I must precede Biology II), other constraints referred to differing scheduling features such as location or teacher availability. As suggested in the discussion of Experiment 1, it is possible that variation in the types of constraints may be a source of constraint satisfaction difficulty. Likewise, a possible source of constraint difficulty in Experiment 2 may have been the growing number of interrelations between design elements that increased with every additional binary constraint. In a low constraint office task, each employee would have fewer interrelations with their fellow employees than in a high constraint office design task. Further research would therefore benefit from examining whether differences in constraint types, specifically general constraint types, and their interrelations with design elements, affects design performance. It may be the case that some particular general constraints are harder to successfully implement. Alternatively, having more general constraints, specifying a growing numbers of qualitatively different interrelations among design elements, may constitute a greater cognitive load. The following chapter aims to investigate these suppositions.

CHAPTER FOUR

Variation in the number of qualitatively different constraint types

---

4.1 Types of general constraints and design elements

Newell and Simon’s (1972) problem space theory, and later extensions of this theory (Lea & Simon, 1974; Zhang & Norman, 1994) indicate that the presence of greater amounts of information (such as specific constraints) and task rules (here general constraints) lead to greater task complexity. Correspondingly, Chapter three concluded that a greater number of constraints, be they lower level, specific constraints between pairs of employees in the office task, or more specific and general constraints in the timetabling task, leads to greater design complexity. The previous chapter also introduced the notion that variation in the differing types of constraints (the differing general constraints), may also have an impact on design complexity. Therefore, the present chapter examines whether design complexity and subsequent design performance is affected by such variation. More specifically, the following experiments will introduce variation in the number of differing types of constraint, and their interactions with design elements, whilst keeping the overall number of specific constraints constant. Relevant theories are now considered.

In the field of instructional design, the notion of cognitive complexity as a result of variation in element interactivity has been proposed. Sweller and colleagues (Paas *et al.*, 2004; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005) suggest that it may not be the number of information items in any material being studied that is a source of cognitive complexity, but whether these items can conceivably and effectively be considered in isolation. In instances where element interactivity is high and information from a number of differing sources needs to be assimilated, greater cognitive load will be placed on the learner. For instance when learning technical information, element interactivity will be higher when trying to incorporate information from a range of sources, compared to learning from worked examples where this information is already integrated. Whilst problem solving performance and learning are not always separate or distinguishable processes (Anderson, 1993; Schnotz & Kürschner, 2007; Sweller *et al.*, 1990, van Merriërboer & Sweller, 2005), the notion of interactivity, albeit phrased somewhat differently, has been proposed as a source of cognitive complexity in all human activity. Indeed, memory processing capacity has been deemed to extend to the parallel processing of around four items (Schneider & Detweiler, 1987) indicating that all cognitive activity may be limited by whether items need to be considered in parallel rather than serially. Similarly, relational complexity theory (Halford *et al.,* 2005; Halford *et al.*, 2007; Halford *et al.*, 1998) suggests that processing capacity is constrained not just by processing limitations but also by the complexity of the items to be processed. Halford *et al.* (1998) indicate that complexity increases with the number of interrelated aspects (related sources of variation) that need to be processed in parallel. They further propose that as relational complexity increases, so will cognitive load. These authors speculate that the need to process an aspect with more than four interrelated aspects may overwhelm working memory capacity and should result in deterioration in performance.

In the context of constraint satisfaction design, both Sweller and colleagues' interactivity theory of cognitive load (Paas, *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005), and Halford and colleagues' theory of relational complexity (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998) may provide insight on how variation in the differing types of constraints applicable to a design element effects the ease and efficiency with which that element is incorporated into a design. When more constraint types interact with, or are applicable to each design element, the section of the problem space that the designer must negotiate when integrating a design element becomes more qualitatively complex. For instance, given an office design task, employee A may be subject to three specific constraints. These specific constraints may all be of the same constraint type (i.e., status, “A is higher in status than C…lower in status than F… lower in status than H”), or could fall under multiple general constraints (i.e., status, “A is higher in status than C”, compatibility, “A is compatible with F”, and work area, “A uses the reception area more often than H”). It is apparent that the latter example requires the integration of more qualitatively different dimensions of information when selecting a position for employee A. An outcome of this is that cognitive processing capacity may be taxed, or indeed exceeded, and design performance may deteriorate. This notion is consistent with the extensions of Newell and Simon’s (1972) problem space theory that imply that an expansion in the problem rule space are a source of task complexity (Simon & Lea, 1974; Zhang & Norman, 1994). Therefore, an increase in the amount of qualitatively differing types of constraint (presently more general constraints), without an increase in the number of constraints (here specific constraints), would still constitute an expansion in rule space. An outcome of such an expansion should be increased task complexity and performance deteriorations may occur. The rest of this section will now discuss preliminary evidence of such effects in related disciplines.

There is a lack of empirical evidence investigating whether variation in the qualitatively differing types of constraints affect human performance in constraint satisfaction contexts. However, within the cognitive load literature, Sweller *et al.* (1990) found that during the learning of technical material, worked examples that integrate relevant diagrams and text produced better learning outcomes in high school children than conventional methods with the same relevant information spread over a variety of sources. These authors indicated that having to attend to and integrate various sources of relevant information, perceptually spread out over different modalities and different sources, was a cause of increased cognitive load that could lead to deterioration in performance. Similarly, various studies of relational complexity have indicated that having a greater number of informational items to consider when reasoning results in performance deterioration. For instance, Birney and Halford (2002) presented psychology students with suppositional reasoning problems containing either four or five relational aspects. They found that participants reasoning with five related items made more errors and had longer response times than those reasoning with just four items. A later study by Halford *et al.* (2005) found that problem solving performance in adults required to match a graphical depiction of numeric results to a verbal description of those results deteriorated as the number of item interrelations increased.

Within the field of computer science, attempts to automate the constraint satisfaction procedure have highlighted some principles, or heuristics, that researchers believe to be both important and helpful. Burke and Petrovic (2002), and Burke, Petrovic and Qu (2006) suggest that an effective heuristic in timetabling is that the design element with the largest number of requirements should be placed first. Likewise, Schaerf (1999) suggests that the most urgent component, that with the greatest amount of restrictions and the fewest design options should be placed first. Both of these principles imply that having a greater number of constraints hinging upon a design element will increase the chances that incorporation of that element will be problematic. This intuitively encompasses notions of element interactivity and relational complexity within a constraint satisfaction context.

In summary, the theories and evidence reviewed here imply that having a greater number of differing constraints interacting with each design element results in greater relational complexity and greater cognitive load. Within constraint satisfaction tasks, this should mean that having to consider more general constraints when incorporating each design element, will be a source of difficulty and should lead to performance deteriorations. Using the office design task, this hypothesis will be examined.

4.2 Experiment 3

The previous section reviewed theories of element interactivity (Paas *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005), relational complexity (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998) and problem rule space expansion (Simon & Lea, 1974; Zhang & Norman, 1994). These complementary theories indicate that when more sources of information (here, differing types of constraint) are applicable to more design elements, the complexity and subsequent difficulty in incorporating these elements will increase. The present experiment examines this prediction using the office layout task.

Participants completed two office design tasks. Both tasks required the arrangement of eight employees (design elements), with the number of specific constraints per employee held constant. However, between the two tasks the number of types of constraint applicable to each employee (or each design element) varied. In one task, labelled the high variability task, three differing constraint dimensions were applicable to each design element. In the other task, labelled the low variability task, whilst three different types of constraint were present within the task, only one or two of these were applicable to any one design element. It is expected that when performing the high variability task, designers will be subjected to greater task complexity and increased cognitive load and task performance will be worse in comparison to that in the low variability task. As such, the high variability task should result in deterioration on all performance measures.

4.2.1 Method

Participants

Forty students, aged between 18 and 25, with an average age of 19.2 years (SD 1.51) participated in Experiment 3 in return for course credit. The sample contained 36 females and four males. None had experience with the experimental task.

Materials

Experimental materials consisted of four office layout tasks adapted from Carroll *et al.* (1980). Each task, programmed in Java, was designed to allow participants to view a set of 18 binary constraints (specific constraints) between fictional employees labelled either A-H, or S-Z. Tables 4.1 and 4.2 display the division of general and specific constraints according to each design element. Each task entailed three general constraints stipulating rules for specific constraint implementation. An equal number of specific constraints pertained to each general constraint. Across tasks, the number of specific constraints per design element was constant (half with 3 & half with 6 constraints).

Table 4.1. The number of general and specific constraints applicable to each design element in a low variability task

Design elements:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Specific constraints pertaining to general constraint 1	3	3				3		3
Specific constraints pertaining to general constraint 2		3	3	3			3	
Specific constraints pertaining to general constraint 3				3	3	3		3
Total specific constraints per design element	3	6	3	6	3	6	3	6

N.B. design element labelling interchangeable (A-H, or S-Z).

Table 4.2. The number of general and specific constraints applicable to each design element in a high variability task

Design elements:	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
Specific constraints pertaining to general constraint 1	1	2	1	2	1	2	1	2
Specific constraints pertaining to general constraint 2	1	2	1	2	1	2	1	2
Specific constraints pertaining to general constraint 3	1	2	1	2	1	2	1	2
Total specific constraints per design element	3	6	3	6	3	6	3	6

N.B. design element labelling interchangeable (A-H, or S-Z).

Two versions of the low and high variability tasks were developed, requiring two sets of general constraints to prevent any constraint-specific practice effects. The general constraints used in Experiment 2, those concerning compatibility, status and work area, made up rule set A. Further constraint rules making up rule set B were as follows:

- Car park location constraints, for example, “Employee A uses the east car park more than employee C”, indicating that employee A should have an office situated closer to the east facing side of the building.
- Lunch arrangement constraints, for example, “Employee F goes out for lunch more than employee H”, indicating that employee F should have to walk past fewer offices when exiting the corridor on which their office is situated.
- Noise level constraints, for example, “Employee G requires a quiet office more than employee D”, indicating that employee G should have an office further away from the noisy, main corridor.

As in Experiment 2, each task had movable interface components allowing the construction of offices and corridors and the relocation of employees in the workspace. On initial presentation, each task program contained a representation of an office floor plan as in Experiment 2. CamStudio was used to record on screen actions whilst participants produced their office designs.

Design

A mixed-factor design was used. The independent variable of interest, the level of design element variability was within-subjects. This variable had two levels; high variability, whereby all three differing constraint types were applicable to each design element, and low variability, whereby just one or two of the differing constraint types were applicable (see Tables 4.1 & 4.2). The between-subjects variable was presentation order. This was fully counterbalanced, as half of the participants completed the low variability task first, and the other half completed the high variability task first. Furthermore, to avoid practice effects, participants were exposed to different constraint rule sets for each task completed. For example, should a participant have performed a low variability task with rule set A, they would then complete a high variability task with rule set B. The dependent product measures collected were the number of constraints satisfied in the participants' end design and the time taken to complete the design. A further process measure, number of design moves utilised throughout the design process was also collected.

Procedure

Participants were initially familiarised with the office design interface and instructed how to generate design features by copying, and then altering, a mini office arrangement containing four employee offices from paper instructions. During each of the counterbalanced experimental tasks the following procedure was used. Participants were given paper instructions to read detailing the constraint rules they would be implementing on the following task and instructed that they should complete the task as quickly and accurately as they could. CamStudio recording began as soon as experimental constraints were displayed. No finish time was set. Participants terminated performance when they felt they had satisfied as many constraints as they could.

4.2.2 Results

Participants with multiple statistical outliers were deleted leaving a sample of 36. As a precautionary measure, the differing sets of general constraints were compared on the high variability task in order to check for any differences that may have arose due to differing rule sets. A paired-samples t-test revealed no significant effect of rule set in respects to the number of task constraints satisfied (t (34) = 1.00, p = .32), arguably the most important indicator of design efficiency. Similarly, non-significant results were obtained for task time (t (34) = .43, p = .67) and design moves (t (34) = .83, p = .41)

Table 4.3 displays the means for all dependent measures as a function of variability in the number of differing constraint types applicable to each design element. Trends among the means displayed here indicate that performance in the low variability task was more efficient on all dependent measures than performance in the high variability task.

Table 4.3. The effect of general constraint variability on performance in an office layout design task

Performance measures		**Low variability task (n = 18)**	**High variability task (n = 18)**
Number of constraints satisfied	Mean	13.28	12.22
	SD	2.93	2.66
Design completion time (seconds)	Mean	580.97	668.69
	SD	176.84	180.12
Number of design moves	Mean	32.36	41.22
	SD	11.26	15.12

Paired-sample t-tests revealed a significant difference between these tasks on the number of constraints satisfied (*t* (35) = 2.14, *p* < .05, *d* = .38), the design completion times (*t* (35) = -2.90, *p* < .01, *d* = .44) and the number of design moves (*t* (35) = -3.73, *p* < .001, *d* = .66). However, due to counterbalancing of experimental tasks, a number of ANOVAs are now reported to ensure additional design measures did not adversely impact upon scores on dependent measures.

Two-way ANOVAs reconfirmed the effects of task variability on the number of constraints satisfied (*F* (1, 34) = 4.63, MSE = 4.33, *p* < .05, *f* = .37), on time taken to complete a design (*F* (1, 34) = 8.35, MSE = 16592.07, *p* < .01, *f* = .50), and also on the number of design moves utilised (*F* (1, 34) = 16.91, MSE = 83.60, *p* < .001, *f* = .70. In respects to constraints satisfied, there was no effect of presentation order (*F* (1, 34) = .04, MSE = 5.81, *p* = .84) and no interaction between presentation order and task variability (*F* (1, 34) = 1.55, MSE = 4.33, *p* = .22). Likewise, for design time, there was no significant effect of presentation order (*F* (1, 34) = 3.61, MSE = 22008.78, *p* = .07) and no interaction (*F* (1, 34) = .64, MSE = 16592.07, *p* = .43). In respects to the number of design moves utilised, there was no significant effect of presentation order (*F* (1, 34) = 2.60, MSE = 121.27, *p* = .12) but there was a significant interaction between task variability and the presentation order of these tasks on the number of design moves utilised (*F* (1, 34) = 8.56, MSE = 83.60, $p < .01$, $f = .50$), as those who performed a high variability task first used fewer design moves on this task than those who performed the high variability task second (see Figure 4.1).

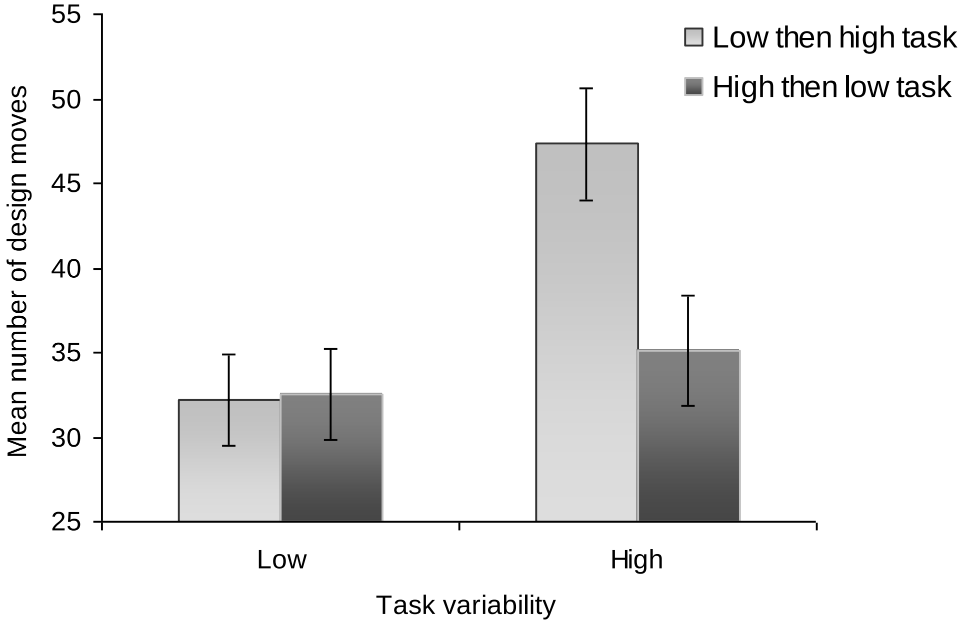


Figure 4.1. The effect of task variability level and task presentation order on the number of design moves utilised. Error bars are +/- 1 standard error.

4.2.3 Discussion

The present findings indicate that increasing the variability of qualitatively different constraint types applicable to each design element results in fewer constraints satisfied, longer design completion times and the use of more design moves. These results imply that increases in the level of variability in general constraints surrounding each design element (despite a constant number of specific constraints) is a source of task complexity and subsequent task difficulty, consistent with experimental predictions. The results are consistent with the implications of various theories of cognitive problem solving discussed in Section

4.1. Problem spaces (Newell & Simon, 1972; Simon & Lea, 1974) may expand in complexity due to an increase in the number of qualitatively differing types of constraint, without the need for an expansion in the number of specific constraints imposed. In effect, a distinction is made between a qualitative enlargement of a rule space as opposed to a quantitative one. As such, when considering the next design move to implement, the section of the problem space relevant to a design element is qualitatively more complex when undertaking a high variability task.

Results are also consistent with implications of Sweller and colleagues' (Paas *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005) interactivity theory of cognitive load, and Halford and colleagues' (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998) theory of relational complexity. Both theories suggest that in situations where multiple pieces of related information need to be considered or appraised in parallel, task complexity will increase. In interactivity theory related information may refer to the differing sources of informational input (i.e., combining information from diagrams with textual explanations; Sweller *et al.*¸ 1990). In relational complexity theory, related information could refer to qualitatively different relationships. The commonality is that increases in the qualitatively different, but relevant information items is a source of cognitive complexity. Within the present constraint satisfaction context, qualitative increases in the number of constraint types applicable to each design element constitutes increased interactivity, or increased relational complexity, and leads to greater complexity when searching for efficient design moves when developing a design.

Additional findings indicate that there were no direct effects of task presentation order on any of the design performance measures. There was, however, an interaction between task variability and task presentation order when examining the number of design moves utilised. Participants who encountered a low variability task first used more design moves on the subsequent high variability task than those who completed the high variability task first. This asymmetric transfer, between tasks may be indicative of different strategies. Perhaps when completing a low variability task first, the relative ease of the task may mean that participants formulate a strategy whereby they consider each constraint, or general constraint, separately. Whilst this may prove a successful strategy for a low variability task, however, should participants attempt to transfer this strategy to the high variability task, the strategy should prove less efficient. This may result in the need to use additional moves in an attempt to satisfy more constraints.

A methodological consideration should be acknowledged as it may affect the interpretation of present results. Whilst the overall number of specific and general constraints were equivalent for both experimental tasks, the allocation of design elements among general constraints differed between tasks. Whilst it was intended that a differing subset of constraints types be applicable to each design element in the low variability task, an artefact of this was that only four design elements were subjected to each general constraint (see empty cells in Table 4.1), but all eight design elements were subjected to each constraint type in the high variability task (see Table 4.2). This may be a confounding form of interactivity, or relational complexity, that may contribute to the pattern of results obtained. Nevertheless, it still constitutes a form of variability among general constraints and design elements.

In conclusion, the present experiment indicates that variation in the number of differing constraint types applicable to design elements has an effect on design efficiency in a constraint satisfaction paradigm. In particular, the greater the variability between design elements and the various general constraint types present within a task, the greater the difficulty in incorporating all of those design elements and constraints into a satisfactory end design. The variability between the number of constraint types and design elements introduced presently could be described as the number of constraint types per design element, or the number of design elements per constraint type. This distinction may be important. As such, the following experiment aims to provide converging evidence that variation and interrelations between the number of qualitatively different constraint types per design element (but not in the number of design elements per constraint type) affects design performance.

4.3 Experiment 4

The results of Experiment 3 supported predictions derived from problem solving theories concerning increased element interactivity (Paas *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005) and greater relational complexity (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998) as sources of task difficulty. It also provided some support for the proposition that navigation of the part of the problem space devoted to each design element becomes more complex in line with increasing the applicable number of differing constraint types (despite a constant amount of specific constraints). The present experiment aims to further investigate the notion that increasing the number of differing constraint types applicable to each design element increases task complexity and should result in deterioration in design performance. Further, the present experiment aims to overcome a limitation of the previous experiment whereby the effects of varying both the number of differing constraint types per design element, and the number of design elements per constraint type, were intractable. This will be attempted by varying the number of constraint types applicable to each design element whilst keeping the number of design elements falling under each different constraint type constant.

A between-subjects design was used. One group of participants completed an office design task with three qualitatively different types of constraint, another group completed a task with six constraint types. In both instances, the constraint types stipulated were applicable to all eight design elements and the quantitative number of constraints, the specific constraints remained constant. Design performance measures collected included the number of constraints satisfied and design completion times. As in Experiment 3, deterioration in both product measures is expected as a result of the presence of a greater number of constraint types interacting with each design element.

4.3.1 Method

Participants

Thirty students, aged between 18 and 24, with an average age of 19.93 years (SD 1.35) took part in Experiment 4 in return for payment. Nineteen were Psychology students, 11 others studied a variety of disciplines. The sample contained 27 females and three males. None had any experience with the experimental task. Participants were randomly allocated to one of two experimental conditions.

Materials

The two office tasks developed for the purposes of the present experiment had the same representational format as those used in Experiment 3. Each office task, programmed in Java, contained 24 specific constraints and eight employees labelled A-H. Office tasks varied in the number of differing general constraints. The low variability tasks contained three general constraints, whereas the high variability task contained six general constraints. Within tasks, each design element was subject to six specific constraints distributed evenly among the differing general constraints. Table 4.4 displays descriptions of the general constraint types used, as well as the distribution of general constraint types and specific constraints among design elements for both experimental tasks. CamStudio was used to record on screen actions whilst participants produced their office designs.

Design

A between-subjects design was used in order to reduce the number of general constraint types required, as developing new constraints raises the possibility that idiosyncrasies may result in some general constraints being harder to implement than others. The independent variable, level of variability between the differing constraint types and design elements, had two levels; low (3 general constraints) and high (6 general constraints). An equivalent number of specific task constraints were divided evenly among general task constraints (see Table 4.4). Dependent measures were the number of constraints satisfied in each participants' end design and time taken to complete the design.

Table 4.4. The division of general and specific constraints per design element for low and high variability tasks

		Low variability task	**High variability task**
General constraint 1	Status: An employee who is higher in status should have an office positioned closer to the central corridor	2 specific constraints	1 specific constraint
General constraint 2	Compatibility: Employees who are compatible should be positioned in adjacent offices, employees who are incompatible should not	2 specific constraints	1 specific constraint
General constraint 3	Work area: Employees who use a particular work area more often than another should have an office positioned closer to that area.	2 specific constraints	1 specific constraint
General constraint 4	Noise: Employees who require a quieter office should have an office positioned further away from the central corridor		1 specific constraint
General constraint 5	Work team: Employees who work in a particular team should have an office adjacent to another person from that work team.		1 specific constraint
General constraint 6	Car park location: Employees who use a particular car park more often than another employee should have their office positioned so that they are closer to the relevant exit.		1 specific constraint

Procedure

The procedure and instructions were very similar to that of Experiment 3 with the exception that participants completed only one experimental design task. The same practise activities were used. Again, there was no time limit for task completion.

4.3.2 Results

No screening issues were found. As a precautionary measure, a paired samples *t*-test was used to compare general constraints 1-3 and 4-6 within the high variability task. Results were not significant (*t* (14) = 1.01, *p* = .33). This was interpreted as an indication that differences in the difficulty of the differing sets of general constraint types did not impact upon results.

Table 4.5. The effect of general constraint variability on performance measures in an office layout design task

Performance measures		Low variability task	High variability task
Number of constraints satisfied	Mean	17.67	13.87
	SD	2.53	2.62
Design completion time (seconds)	Mean	761.33	733.87
	SD	178.74	221.22

Table 4.5 displays the means on both dependent product measures as a function of the level of variability in general constraint types. Means indicate that participants performing a low variability task satisfied approximately four more specific design constraints than those completing the high variability task. An independent samples t-test revealed that this difference was significant (*t* (28) = 4.05, *p* < .001, *d* = 1.53). In contrast to Experiment 3, mean time (seconds) taken to complete an office design was slightly higher in the low constraint type task. However, these design times between the low and high variability tasks did not differ significantly (*t* (28) = .37, *p* = .71).

4.3.3 Discussion

Results indicated participants undertaking a low variability office design task, satisfied more specific constraints than those completing a high variability task. This result supports experimental predictions as derived from the implications of Sweller and colleagues' theory of interactivity and cognitive load (Paas *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005), and also Halford and colleagues' similar theory of relational complexity (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998). As such, having a greater number of qualitatively different constraint types to consider when attempting to position design elements appears to be an important source of task complexity and subsequent design difficulty.

This result also provides further support to the notion that variation in the qualitative complexity of the problem space surrounding each separate design element is a source of task difficulty. Simon and Lea's (1974) extension of problem space theory (Newell & Simon, 1972) suggested that the task problem space was made up of a state space defining all possible task states and a rule space defining the operators and restrictions applicable within that task. Simon and Lea also proposed that the size of these spaces and the interaction between these spaces determine task complexity. The present experiment, and the proceeding one, indicate that variation in the complexity of the general constraint space, or the rule space, surrounding each design element, is also a source of performance variation. Here the increase in complexity is achieved via the introduction of a greater number of qualitatively differing rule dimensions rather than via an increase in the overall number of constraints applicable.

Further results concerning time taken to complete a design task did not support experimental predictions. It could be the case that had participants completing the high variability task invested more time, they may have been able to improve upon the number of design constraints that they satisfied. However, in order to do so, time efficiency would deteriorate. The lack of a significant difference in design completion times between experimental tasks is difficult to interpret. Payne and Duggan (2011) suggest that in uncertain problem solving environments, when faced with a problem of considerable complexity, there are several reasons why someone may terminate problem solving efforts whilst not having reached a satisfactory solution. One such reason is a frustration effect. Another is an inability to monitor whether forward progress is being made, or the evaluation that further efforts will not act to improve upon a problem solution. Indeed, uncertain stop-criteria are a typical feature of design tasks with some researchers speculating that the designer will need to select their own stop criteria (Goel & Pirrolli, 1992). It could be speculated that the added complexity of the high variability task may lead to any of the personal stopping criteria outlined by Payne and Duggan (2011).

As in Experiment 3, the present experiment has a methodological consideration that should be acknowledged. Whilst the present design overcomes the limitation of having differing numbers of design elements per general constraint type, it also has another form of variation that may affect the interpretation of the results. Here, design elements in the low variability task had two specific constraints falling under each of the three differing constraint types, whereas elements in the high variability task had just one specific constraint pertaining to each of the six general constraint types. An artefact of this was that the high variability task contained double the number of general constraints per task than the low variability task. It appears that both attempts to systematically introduce variability between the number of differing constraint types applicable to each design element introduces another, unintentional form of variation.

In summary, present results offer support to the proposal that variability in the number of qualitatively different constraint types applicable to each design element, is a source of performance variation in constraint satisfaction tasks. Whilst no effect of altering the variability between design elements and the number of applicable constraint types was found for design completion times, the number of explicit constraints satisfied decreased as variability increased.

4.4 Conclusions

Experiment 3 and 4 offer converging support for the notion that design element variability, as instigated via the number of qualitatively differing constraint types applicable to each design element, is an important determinant of task complexity and subsequent task difficulty. Increasing variability has clear and consistent effects on the number of constraints satisfied. However, mixed effects were found in relation to design completion times, with an effect found in Experiment 3 but none in Experiment 4. For the greater part, results support theoretical implications derived from the theory of element interactivity (Paas *et al.*, 2004; Sweller, 1988; Sweller, 1994; Sweller *et al.*, 1990; van Merriërboer & Sweller, 2005) and relational complexity theory (Halford *et al.*, 2005; Halford *et al.*, 2007; Halford *et al.*, 1998). Both imply that increasing the number of items that need to be considered in parallel will increasingly tax and then maybe also exceed designers' mental resources. Results also support extensions of problem space theory concerning rule spaces. Simon and Lea (1974) suggest that a rule space will interact with a state space to determine problem complexity. As such the size of the rule space is a determinant of task complexity (Simon & Lea, 1974; Zhang & Norman, 1994). The distinction made presently is that an increase in the size of the rule space does not necessarily refer to the inclusion of more items of information; this would constitute a quantitative enlargement, as examined in the previous chapter. Instead qualitative enlargements, via the presence of more information dimensions, are a cause of cognitive complexity and performance difficulties.

In relation to experimental design, both experiments had related methodological considerations. It appears that interactivity, or relational complexity within the office design paradigm is a delicate issue. Variation in one form of interactions necessitates another form of variation. For instance, in Experiment 3, variability between design elements and the number of differing constraint types relevant could be defined as the number of constraint types acting upon each design element, or, the number of design elements falling under each constraint type. In Experiment 4, the number of specific constraints per general constraint type varied. Untangling some of these forms of variation may be beyond the scope of the office task paradigm.

An aspect that has not been discussed as yet is whether the nature of the constraint types, rather than the number of them, is a further source of design complexity and performance variation. It may be the case that certain constraint types are harder to implement than others. Comparisons of the differing sets of general constraint types in Experiments 3 and 4 did not uncover any differences in the ease of applying the sets of general constraints utilised so far. However, it is possible that the nature of certain constraints may still play a role in determining design difficulty. For instance, a difference may exist in the ease of satisfying a constraint that specifies a relationship between two employees and a physical reference point, as opposed to the ease of satisfying a constraint stipulating the relationship between two employees in reference to those who surround them. It could be the case that constraints containing reference points that are also design elements (with requirements and restrictions of their own), may be harder to implement. As such, there may be a differential ease with which certain constraint types interact with, and translate onto the external representation. The following chapter shall focus on whether such variation in the nature of constraints, and their interplay with the set external office representation impacts upon design performance.

CHAPTER FIVE

Constraints and cognitive fit

5.1 Cognitive fit between differing constraints and the external representation

There is growing interest in the interplay between problem information (here constraints), a problem solution and a problem representation (Scaife & Rogers, 1996; Shaft & Vessey, 2006). Whilst the preceding experiments have explored variation in constraint satisfaction performance as a result of variation in the quantitative or qualitative number of constraints, the present discussion focuses on whether qualitative differences in the processing of certain categories of constraints, and their interplay with the external problem representation, may be a source of task complexity. The discussion that follows outlines how external representations may differentially support design cognition, and introduces cognitive fit theory (Shaft & Vessey, 2006; Vessey; 1991; Vessey & Gellata, 1991) as a theoretical framework accounting for why an external representation may be better suited to supporting a certain type of information, or constraint, processing. Following this, differences in constraint processing within the office design paradigm are outlined in respects to implications derived from cognitive fit theory.

Problem-solving is inexplicably bound to concepts of problem representation (Dym, 1992). Indeed, there is some suggestion that the structure of the external representation plays a crucial role in directing cognitive efforts and determining cognitive efficiency (Carroll *et al.*, 1980; Visser, 2006; Zhang, 1997; Zhang, 2000; Zhang & Norman, 1994). Likewise, there is also evidence suggesting that differing forms of external representation may differ in the suitability in which they support cognitive processes (Jones & Schkade, 1995; Moreno, Ozogul & Reisslein, 2011; Novick, 2001). Cognitive fit theory formalises and explains what underlies variation in external representation suitability. Cognitive fit theory stipulates that when problem information is emphasised, that is, well represented within the representational or presentational format, a greater degree of fit is achieved (Shaft & Vessey, 2006; Vessey, 1991; Vessey & Gellata, 1991). When the degree of cognitive fit is high, problem solvers can use the same cognitive strategy to process both presentation and task information, leading to more effective problem solving (Kwon, Lee & Mustapha, 2011; Vessey, 1991). Put simply, the presence of a greater degree of cognitive fit affords a degree of computational offloading (Gibson, 1979; Gero & Kannengiesser, 2012) that may relieve the cognitive load experienced by the problem solver. In contrast, when the degree of cognitive fit between task information and the external representation is low, problem solvers may need to instigate effortful cognitive transformations in order to assimilate problem information satisfactorily. This will require greater cognitive effort, and may tax processing limitations leading to poorer problem solving performance.

Given an already complex problem solving context such as constraint satisfaction design, variation in the level of cognitive fit between the processing requirements of differing constraints and the external representation may be a crucial determinant of task difficulty and subsequent performance efficiency. There is some preliminary evidence to attest to the differential ease with which certain constraints may be satisfied within constraint satisfaction contexts. Both Carroll *et al.* (1980) and Visser (2004) report that spatial constraints may be easier to satisfy than temporal constraints. Carroll *et al.* also touch upon ideas of cognitive fit when proposing that spatial constraints lend themselves more easily to external spatial representation whilst temporal constraints appear harder to represent externally. Carroll *et al.* further propose that sourcing an appropriate external representation may alleviate some of the difficulty in satisfying temporal constraints. The rest of this introduction is now devoted to discussing underlying causes of variation in cognitive fit in the present office design paradigm. Firstly, further distinctions, or constraint categories, are proposed. Afterwards, the nature of their interplay with the external representation and the subsequent cognitive fit is explored.

In the office paradigm, all external constraints pertain to a spatial arrangement task and are essentially spatial in nature. However, there is a distinction to be made between constraints that reference a fixed point (an anchor point) in the problem representation, or constraints that simply reference other design elements. An example of a specific constraint referencing a fixed point (forthwith referred to as fixed constraints) within the design problem representation is, “Employee B uses the reception area more than employee G”, with the corresponding general constraint stipulating that B should be positioned closer to this fixed point. Other task constraints relate to the relative positioning of a design element in relation to another element, or in relation to how many neighbouring elements it should have (forthwith referred to as non-fixed constraints). Examples of non-fixed constraints include, “Employee H requires a quieter office than employee C”, indicating that employee H should be adjacent to fewer occupied offices, and, “Employee D is compatible with employee F” indicating that they should be positioned in adjacent offices.

It is this distinction between fixed and non-fixed constraints that may be crucial in determining the level of cognitive fit between task information and the external representation. When attempting to satisfy a fixed constraint, an unmovable, physical anchor point is available for use when monitoring the implementation of that constraint. As such, this physical anchor serves as an external memory point, highlighting problem information and affording a degree of computational offloading (Gibson, 1979; Gero & Kannengiesser, 2012). Here a good degree of cognitive fit is achieved. However, the monitoring and implementation of non-fixed constraints is not afforded any similar cognitive support within the present external office representation. Here designers need to actively track the relative positioning of design elements and perhaps also actively monitor any changes to the neighbouring design elements should any changes be implemented. Designers will need to source, or if not yet assimilated into a design imagine, the positioning of other relevant design elements. An example office design is now used to demonstrate some of these differing processing requirements.

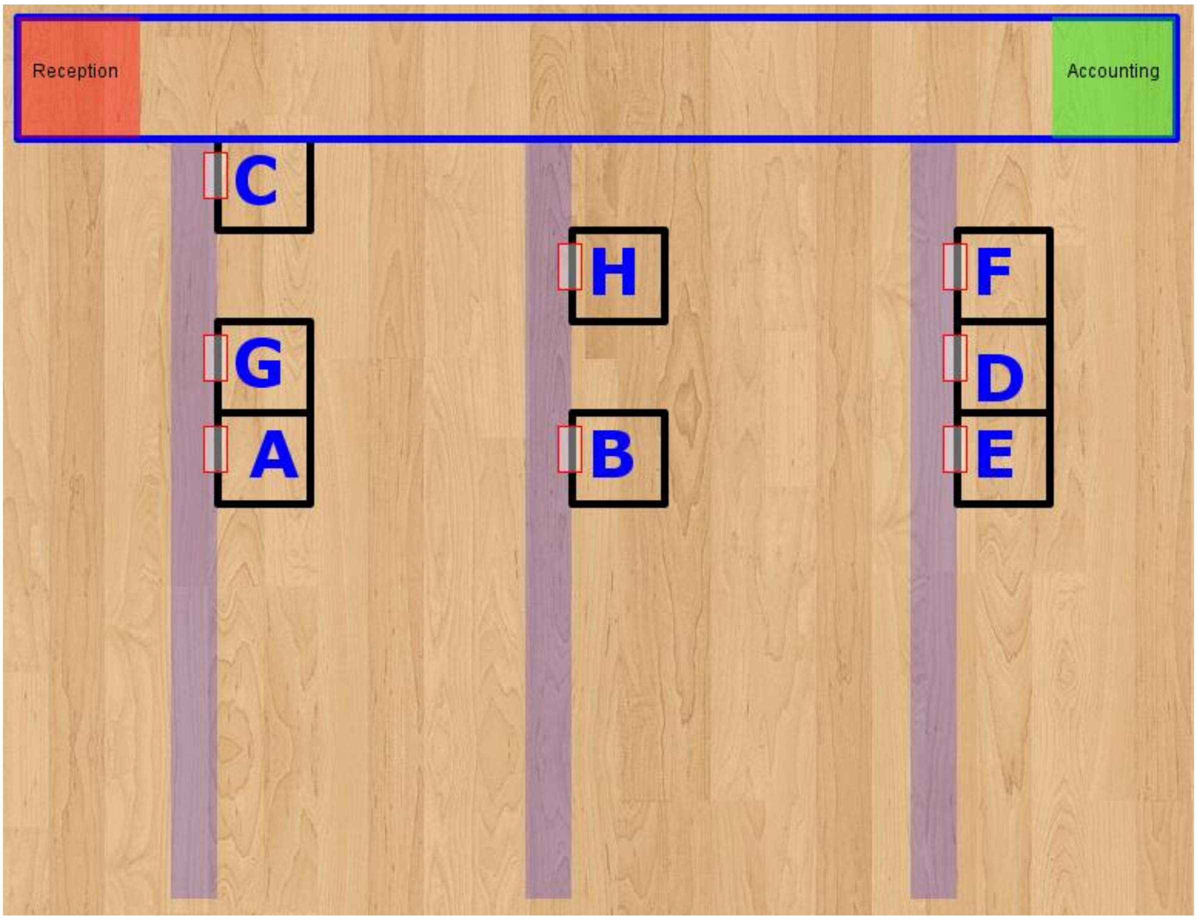


Figure 5.1. Example office design

Table 5.1. Examples of fixed and non-fixed constraints applicable to employee F

Fixed constraints	*Non-fixed constraints*
1. F requires a quieter office than C and should be positioned further from the main corridor than C.	4. C requires a quieter office than F and should be positioned next to fewer adjacent offices than F.
2. H is higher in status than F and should be positioned closer to the main corridor than F.	5. F goes out for lunch more often than H, therefore F should pass fewer other employee offices when exiting their corridor in comparison to H.
3. G uses the reception area more than F and should be positioned closer to this area.	6. G is incompatible with B, therefore these employees should not be adjacent.

N.B. Adjacent defined as in office directly above or below another so that they share an office wall

Figure 5.1 displays an external representation with an office design containing eight design elements. Consider the example fixed and non-fixed constraints, pertaining to employee F detailed in Table 5.1. In both lists only one constraint is currently satisfied. However, this is much easier to detect within the fixed constraints. In addition, should a designer attempt to implement more constraints satisfactorily, sourcing design moves that would satisfy additional fixed constraints is simpler. Here a designer need only track the two relevant employees and assess their distance from a fixed point that is perceptually easy to pick out in the representation. In contrast, when attempting to satisfy another non-fixed constraint two employees and their neighbours must be considered, a more complex process with little computational support from the external representation. Having to implement a larger number of non-fixed constraints should serve only to compound these difficulties. As such, the greater the proportion of non-fixed constraints, the harder it should be to successfully implement, monitor or evaluate design progress.

5.2 Experiment 5

This experiment aimed to assess whether increasing the proportion of fixed constraints within a task produces greater levels of cognitive fit between the constraints stipulated and the external representation. In order to test this, four groups of participants performed differing office design layout tasks. Each task contained a different proportion of fixed and non-fixed constraints, all with equivalent numbers of specific constraints and general constraints. It was expected that the greater the proportion of fixed constraints, the greater the degree of cognitive fit (Shaft & Vessey, 2006; Vessey, 1991; Vessey & Gellata, 1991). This greater cognitive fit should afford a degree of computational offloading (Gibson, 1979; Gero & Kannengiesser, 2012) not available when processing non-fixed constraints. The outcome of improved cognitive fit should be better design performance.

5.2.1 Method

Participants

Forty-five psychology students, aged between 18 and 24 (mean 19.07 years, SD 1.17) took part in Experiment 5 in return for course credit. The sample consisted of six males and 39 females. None had any experience with the experimental task. Participants were randomly assigned to one of four experimental conditions.

Materials

Six office design tasks were developed for the purposes of the present experiment. All were programme in Java with the same presentational format as that used in previous office design tasks. Each office task allowed the participants to view a set of 27 specific constraints, relating to eight design elements (employees labelled A-H), that were divided equally amongst three general constraints. The number of fixed and non-fixed constraints varied. Table 5.2 displays the combinations of fixed and non-fixed general constraints used in each experimental task. Cam Studio was used to record participants on screen actions throughout each.

Design

A between-subjects design was used with one independent variable, proportion of fixed and non-fixed general constraints; all fixed, 2/3 fixed, 1/3 fixed & No fixed. As a precautionary measure, tasks comprising both fixed and non-fixed constraints had two versions (labelled A & B). In these tasks, the general constraint type (fixed or non-fixed) in the minority was varied in order to reduce the possibility that the idiosyncrasy of this qualitative change may effect dependent measures. Once again, the dependent measures were the product measures number of constraints satisfied and time taken to complete the design.

Table 5.2. General constraints used in each experimental task

	All fixed constraint	**2/3 fixed constraints (version A)**	**1/3 fixed constraints (version A)**	**No fixed constraints**
General constraint 1	**Fixed** - Work area; employees using a work area more often than another should have an office positioned closer to that area	**Fixed** - Work area; employees using a work area more often than another should have an office positioned closer to that area	**Fixed**[2] - Work area; employees using a work area more often than another should have an office positioned closer to that area	**Non-fixed** - Noise; an employee who requires a quieter office than another employee should have fewer adjacent neighbours
General constraint 2	**Fixed** - Status; employees who are higher in status should have an office closer to the central corridor.	**Fixed** - Status; employees who are higher in status should have an office closer to the central corridor.	**Non-fixed** - Noise; an employee who requires a quieter office than another employee should have fewer adjacent neighbours	**Non-fixed** - Compatibility; employees who are compatible should have adjacent offices, employees who are incompatible should not.
General constraint 3	**Fixed** - Car park location; employees who use a car park more often than another should have an office closer to that side of the building.	**Non-fixed**[1] - Compatibility; employees who are compatible should have adjacent offices, employees who are incompatible should not.	**Non-fixed** - Compatibility; employees who are compatible should have adjacent offices, employees who are incompatible should not.	**Non-fixed** - Lunch arrangement ; an employee who goes out for lunch more often than another walk past fewer other offices when exiting their corridor

[1] Version B of this condition substituted this non-fixed general constraint for the non-fixed constraint, noise (an employee who requires a quieter office than another employee should have fewer adjacent neighbours).

[2] Version B of this condition substituted this fixed general constraint for the fixed constraint, status (employees who are higher in status should have an office closer to the central corridor).

Procedure

The procedure and instructions were very similar to that of Experiment 4. Participants completed only one experimental design task and the same practise activities were used. Again, there was no time limit for task completion.

5.2.2 Results

One participant was excluded as they failed to follow experimental instructions. High levels of kurtosis were found within the 1/3 fixed constraint condition. However, transformations were not applied as they may alter the interpretation of the data (Tabachnick & Fidell, 2007). In order to check whether there were any idiosyncratic differences as a result of the differing versions of general and specific constraints within the 2/3 and 1/3 fixed constraints conditions, independent samples *t*-tests were conducted. For the 2/3 fixed constraint condition, there was no difference of task version (A or B) on the number of constraints satisfied (*t* (9) = .81, *p* = .44) or design completion time (*t* (9) = -.84, *p* = .42). For the 1/3 fixed constraint condition, there was also no effect of task version on the number of constraints satisfied (*t* (9) = -1.79, *p* = .12) or design time (*t* (9) = 1.56, *p* = .15).

Table 5.3. The effect of varying proportions of fixed constraints on office design performance measures

		All fixed constraints (n =12)	**2/ 3 fixed constraints (n =11)**	**1/3 fixed constraints (n = 10)**	**No fixed constraints (n = 11)**
Number of constraints satisfied	Mean	19.33	17.64	13.60	11.00
	SD	2.53	2.16	2.07	3.38
Design completion time (seconds)	Mean	1231.17	1093.09	1163.40	1088.64
	SD	323.88	366.93	282.87	291.48

Means for both dependent measures are displayed in Table 5.3. These indicate that participants completing an office design task entailing only fixed constraints satisfied the highest number of design constraints. In contrast, those undertaking an office design task entailing only non-fixed constraints satisfied the fewest task constraints. Participants completing tasks entailing combinations of these constraints (2/3 & 1/3 fixed constraint conditions) showed intermediate performance in terms of the number of constraints satisfied (see Figure 5.2)

A one-way, between-subjects ANOVA revealed a significant effect of the proportion of fixed constraints on the number of constraints satisfied (*F* (3, 40) = 24.00, MSE = 6.74, *p* < .001, *f* = 1.34). βonferroni comparisons indicated that participants in the All fixed condition satisfied more constraints than both the 1/3 fixed condition and the no fixed constraint condition (*p*s < .001) but not significantly more than the 2/3 fixed constraint condition (*p* = .75). In addition, participants in the 2/3 fixed constraint condition satisfied more constraints than those in the 1/3 fixed constraint condition (*p* < .01) and the no fixed constraint condition (*p* <001). Performance between the 1/3 fixed constraint condition, and the no fixed constraint condition did not significantly differ (*p* = .16). A further linear analysis was conducted. A significant linear effect was found (*F* (1, 40) = 70.51, MSE = 6.74, *p* < .001). Deviation from this linear effect was not significant (*F* (1, 40) = .75, MSE = 6.74, *p* = .48). As indicated in Table 5.3, there was no apparent trend in the design completion times. A one-way between-subjects ANOVA confirmed that there was no significant differences between experimental conditions for design completion times (*F* (3, 40) = .52, MSE = 101749.54, *p* = .68).

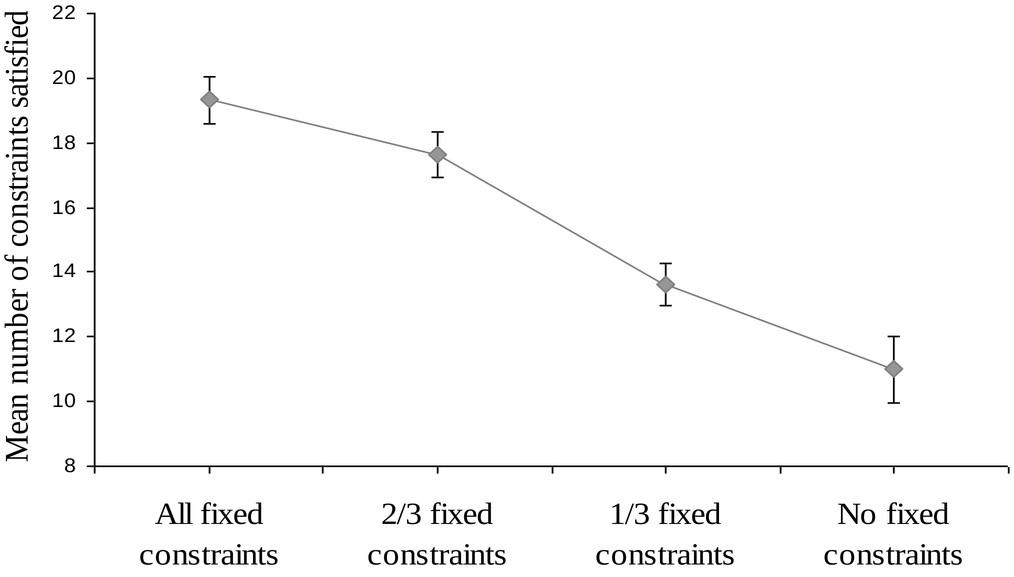


Figure 5.2. The effect of the proportion of fixed constraints on the number of design constraints satisfied. Error bars are +/- 1 standard error.

5.2.3 Discussion and conclusions

The present results indicate that varying the proportion of fixed and non-fixed constraints, both quintessentially spatial constraints, has an effect on the number of constraints satisfied, arguably the most important indicator of design efficiency. A linear effect of greater fixed constraints was established, with no significant deviation from this trend detected. These results support the idea that there exists differences in the ease of satisfying differing types of spatial constraint (i.e., Carroll *et al.*, 1980). Further, evidence that the interplay between these differing forms of constraint and the external representation is provided. This interplay is explained by cognitive fit theory (Kwon *et al.*, 2011; Shaft & Vessey, 2006; Vessey, 1991; Vessey & Gellata, 1991). When problem information is emphasised within the external representation, the representation will afford a degree of computational offloading (Gibson, 1979; Gero & Kannengiesser, 2012), as the same processing strategies, revolving around a fixed representational reference point, can be used to process constraint information and assimilate constraints into the external representation. Presently, good cognitive fit is evident between fixed spatial constraints. The lack of an equivalent representational affordance, or match, results in reduced cognitive fit when processing non-fixed constraints. Here, designers are likely to experience substantially more cognitive load. Put simply, within the present experiment, design constraints are more readily implemented when cognitive activity can be anchored around an easily recognised feature within the external solution representation.

An alternative explanation of the results concerning the number of constraints satisfied is possible. It may be the case that having a greater proportion of non-fixed constraints results in greater task complexity for other reasons. In order to implement a non-fixed constraint, at least some of the constraint-relevant design elements must already be positioned within the external office representation. This is also true of fixed constraints, but to a lesser extent as typically there are fewer design elements to consider when implementing a fixed constraint. Nevertheless, this may be another way in which reduced cognitive fit is demonstrated, as a certain level of transformational cognitive processing is needed before a non-fixed constraint can be implemented (Kwon *et al.*, 2011). This may have resulted in more design moves being needed to cope for the dynamically changing patterns of implemented constraints in tasks with greater proportions of non-fixed constraints. No firm conclusions can be made. Transcribing of design videos could be attempted to find out if this suggestion is valid.

Results concerning time taken to complete a design task did not support experimental predictions, as there was no clear ascending pattern of results in line with decreasing proportions of fixed constraints. As in Experiment 4, these results are difficult to interpret. Possible variation in task completion times may be due to frustration effects, differences in the belief that further progress can be made or difficulty to monitor further design moves (Payne & Duggan, 2011). What is apparent is that stop criteria may vary between individuals both between, and within conditions.

There are a couple of methodological considerations that should be acknowledged as they may affect the interpretation of present findings. Firstly, the designers used here were all psychology undergraduates. None had any formal design training or any previous experience with the experimental task. As such, it may be the case that an increased importance was assigned to the presentational format as a basis for task exploration and subsequent cognitive activity (cf., Scaife & Rogers, 1996). The presentational format used here was an external representation of the office space, with some fixed features. Should the novelty of the task make these fixed features more salient, then differences in cognitive fit between the external representation and the differing subtypes of constraints may become exaggerated. Visser (2004) reported that despite initial difficulty incorporating temporal (as opposed to spatial) constraints in a travelling salesman paradigm, this abated somewhat as the designers gained more task experience. It could be the case that differences in the ease of implementing non-fixed constraints may diminish with more experience. As such, some caution should be advised when generalising the present results to more experienced designers.

In summary, the present experiment indicates that performing an office design task containing greater proportions of non-fixed spatial constraints leads to deterioration in the number of constraints satisfied, an important determinant of performance efficiency. This is attributed to a reduced degree of cognitive fit between the task constraints and the representational format provided, as the representation affords cognitive benefits only when processing fixed spatial constraints.

FOREWORD TO FOLLOWING EXPERIMENTAL CHAPTERS

Attention so far has remained focused upon performance variation as a result of systematic changes in explicit design constraints. The remaining experimental work is now devoted to examining whether training interventions may prove beneficial in improving design performance in constraint satisfaction contexts. In particular, the following chapters will focus on practice and also on metacognitive training interventions. The following discussion provides a brief introduction to the training literature.

Training refers to a systematic approach to induce learning and development (Aguinis & Kraiger, 2009; Goldstein & Ford, 2002). The overall goal of training is to enhance individuals' knowledge, skills and abilities (Holladay & Quiñones, 2003; Yamnill & McLean, 2001). Within occupational and organisational training domains, training has been accredited with individual, team and organisational performance benefits (Goldstein & Ford, 2002; Patrick, 1992). Presently, focus is placed on the training and acquisition of constraint satisfaction skills on an individual basis. At this level of analysis, documented benefits of training interventions include the acquisition of declarative, procedural and strategic knowledge (Aguinis & Kraiger, 2009). However, not all training schemes are guaranteed to produce such benefits. The effectiveness of training may vary depending on the design and delivery of the training, and also depending on the complexity and requirements of the task being learned (Aguinis & Kraiger, 2009; Arthur, Bennett, Edens & Bell, 2003; Salas & Cannon-Bowers, 2001; Tannenbaum & Yukl, 1992). Salas and Cannon-Bowers's (2001) narrative review of the training literature concluded that the most effective training strategies present relevant information to be learned, create opportunities for practice, and provide feedback.

Further important distinctions can be made between certain types of feedback (Patrick, 1992). Intrinsic feedback is the information that a problem solver can gleam simply from simple task exposure. The level of intrinsic feedback may vary according to the task. For instance, in motor learning tasks such as darts, with increasing levels of practice learners would be able to judge what adjustments may be needed, and change their strategy accordingly in order to hit the target (Annett, 1991). However, in complex cognitive tasks, the utility of environmental feedback may be unclear. Here, the intrinsic feedback gained from task exposure may be hard to process. In contexts such as constraint satisfaction, intrinsic feedback may be of use in familiarising designers with the application of task rules. Whether intrinsic feedback would be of further benefit is unknown. Extrinsic feedback is the provision of task relevant information above what may be gained from simply practicing a task. Here, additional information on task strategies and progress is provided in order to hone cognitive processes. There are typically two sources of extrinsic feedback: that provided by a trainer, or that elicited from the problem solver. Such feedback may be quickly assimilated and utilised. Whilst the use of intrinsic feedback accumulated via practice alone may produce slow learning effects, practice with extrinsic feedback achieves more rapid improvements through initiating cognitive changes (Annett, 1991; Lussier & Shadrick, 2006).

Whilst a combination of the training design features suggested by Salas and Cannon-Bowers (2001) should lead to the optimal design performance, instructional designs may focus more closely on particular elements. For instance, practice is both simple and easy to administer. Lussier and Shadrick (2006) suggest a key feature of practice is that it may help performance to speed up. However, an extensive amount of training may be required in order to achieve this. Nevertheless, practice may allow individuals to assimilate intrinsic feedback and spontaneously improve. As such, the possibility that practice may lead to performance benefits without the need to engage in more effortful training interventions aimed at changing behaviour cannot be ruled out. Chapter 6 examines whether the intrinsic feedback available via practice, and variability in the amount of such practice, can lead to spontaneous improvements in constraint satisfaction design.

Following on from the examination of practice and intrinsic feedback, Chapter 7 examines the utility of additional, extrinsic feedback. Extrinsic feedback comes in many forms and varieties. Examining the utility of all forms of extrinsic constraint satisfaction feedback would be a mammoth undertaking and is beyond the scope of this thesis. Presently, a couple of select interventions are examined. Feedback involving self-explanation, a source of self-generated metacognitive feedback, is gaining increasing attention as a method of inducing performance improvements on cognitive tasks. Shin *et al.* (2003) suggest that solving ill-structured problems may require metacognitive activities such as self-justification. Given the little empirical evidence concerning constraint satisfaction in humans, it is not surprising that metacognition has yet to be examined as a potential constraint satisfaction aid. Chapter 7 aims to address this gap in the literature.

A final notable feature of training interventions is the method by which performance improvements are measured. The effectiveness of training is typically assessed via comparative performance on transfer tests (Schmidt & Bjork, 1992) relative to a control condition (and relative to any other training conditions of interest). These transfer tasks typically fall into two categories; near transfer and far transfer. In a near transfer task, task features are similar to that used in the training phase, whereas in far transfer tasks task features differ to some degree (Holladay & Quiñones, 2003; Yamnill & McLean, 2001). The degree of similarity between training and transfer tasks may be based upon multiple task characteristics, including the breadth of domain knowledge, the performance context and the functionality of the task solution (Barnett & Ceci, 2002). Positive transfer on a near transfer task may indicate a task-specific form of skill acquisition or learning, whereas, positive transfer on a far transfer task may be indicative of the acquisition of more general skills and strategies (Gray & Orasanu, 1987; Healy, Wohldmann, Sutton & Bourne, 2006). Whilst, task-specific learning may be relatively easy to achieve, these benefits are typically outweighed by the acquisition of general skills. When general transfer is achieved, flexibility in the application of learnt skills is demonstrated.

CHAPTER SIX

Practice, practice schedules and intrinsic feedback

6.1 Practice as training

Theories of skill acquisition have outlined the various mechanisms and processes through which skilled performance may be encouraged. Arguably the most influential theory, Anderson's (1982, 1987, 1993, 1996) theory of skill acquisition, proposes that skill is underlined by the accumulation of declarative and procedural knowledge, that is, knowing how to apply task knowledge in a timely manner in order to achieve a desired goal. Such activities should lead to rule familiarisation and schema development (Paas, 1992; Pashler, Johnston & Ruthruff, 2001). An integral part of accumulating procedural and declarative knowledge is engaging in task practice.

The experiment detailed in this chapter focuses solely on the impact of intrinsic feedback, accumulated solely through simple task practice, on constraint satisfaction performance. In addition, the effects of repeated practice trials are also scrutinised. Intrinsic feedback refers to the task-relevant learning that occurs via simple exposure to a task (Patrick, 1992). Intrinsic feedback can encompass environmental feedback and improved memory for task features. Such feedback has been sufficient enough to produce some performance benefits in motor tasks, however, improvements in other forms of skill acquisition is usually only seen once extensive practice has been undertaken. (i.e., Annett, 1991; Newell & Rosenbloom, 1981). Nevertheless, it is important to gauge whether the intrinsic feedback available through practice can lead to better design performance in constraint satisfaction contexts. Should any performance advantages be gained by practice alone, it may be the case that additional performance benefits are gained via engaging in repetitive practice trials.

Presently, there is a diminishing body of empirical evidence surrounding practice without further instruction in complex cognitive problem solving. Most of the recent literature documents training developments that incorporate some form of additional training instruction and are not relevant to the present discussion. The empirical studies that do document practice without extrinsic feedback have turned their attention to exploring how practice schedules may affect task performance. Evidence concerning practice schedules is now reviewed briefly.

There is some suggestion that increasing amounts of practice result in reduced cognitive load and fewer performance errors (Carlson, Sullivan & Sneider, 1989; Lussier & Shadrick, 2006; Wickens & Hollands, 1999). Shute, Gawlick and Gluck (1998) found that learners in a computer-based statistics course performed better on a performance test if they had engaged in more practice problems per topic tested. Likewise, Yeo and Neal (2004) found that participants given multiple practice trials on an air traffic control task showed increasingly higher performance. Whether such improved performance transferred to other tasks was not tested. Both of these studies indicate that repeated opportunities to practice have beneficial effects for learning in complex problem-solving domains. Newell and Rosembloom (1981) proposed a relationship between practice and performance such that the speed of skilled performance should be dependent upon the number of practice trials, performance on the first practice trial and the learning rate. Essentially, when performing repetitive tasks, performance should improve substantially after the first trial, but subsequent improvements will decrease exponentially until performance plateaus (Patrick, 1992). Intuitively, increasing the number of practice trials should result in improved performance, above and beyond that apparent after one practice trial. Whether this is true on constraint satisfaction design shall be the focus of the experiment that follows.

In order to assess whether there are any beneficial effects of the intrinsic feedback available via simple task exposure, or via greater levels of such practice trials, evaluation criteria need to be established. The potential beneficial effects of any skill acquisition acquired via practice or training interventions are typically measured via the amount of positive transfer between training and transfer tasks. Positive transfer between any training intervention and a transfer task often depends upon the similarity between the skills, the description of the task and contextual factors (Yamnill & McLean, 2001). Greater transfer is often accredited to greater similarity (Holding, 1965) although gauging similarity may be problematic. Yamnill and McLean (2001) distinguish between near and far post-training performance. Near transfer tasks require the application of trained/practiced skills to similar tasks. Here positive transfer may be expected to be greater. Far transfer tasks are dissimilar from the tasks encountered in the training stage. Barnett and Ceci (2002) argue that descriptions of near and far transfer vary widely in the training literature. They propose that near and far can refer to various features of a transfer task such as the knowledge domain/discipline, contextual factors or the timing of transfer tasks. Presently, a near transfer task is defined as a task with the same structure, carried out in the same context and within the same discipline. Far transfer is a task within the same discipline and context, but possessing a differing, enlarged task structure making the task more complex (see Experiment 1). If positive transfer is shown on a far transfer task (as would be indicated by more efficient performance for the practice conditions in contrast to a control condition), this may be taken as evidence that generic skills or heuristics, are acquired during practice (Schmidt & Bjork, 1992).

6.2 Experiment 6

The previous section indicated that the intrinsic feedback available via practice without further instruction may be of benefit should multiple practice trials be undertaken. Whilst there is a distinct lack of empirical evidence attesting to any cognitive benefits of a single task exposure, there is some indication that engaging in multiple practice trials may invoke memory traces and quicken performance (Yeo & Neal, 2004; Shute *et al.*, 1998).

In the present experiment, groups of participants were subjected to differing practice schedules (1 practice trial, 5 practice trials or no practice). Practice was accompanied by no further instruction or extrinsic feedback. Only intrinsic feedback was available. Practice tasks were followed by a near transfer task, whereby the same task structure (the same general and specific explicit constraints) was applicable. However, superficial differences were introduced by renaming the classes, teachers, rooms and facilities. In addition, all participants completed a far transfer task. Here, the task differed not only in superficial differences, but also in task structure (as induced by increasing the number of general constraints).

Given the complexity of the timetabling constraint satisfaction paradigm utilised here, it is expected that the intrinsic feedback available via repeated practice trials should result in improved performance in a near transfer task, with fewer errors shown by those participants undergoing multiple practice trials in comparison to participants with no prior experience and those undergoing only one timetabling practice trial. Due to the lack of evidence concerning the value of intrinsic feedback from single practice trials, differences between the 1 trial group and the control group are not hypothesised. Whether practice and intrinsic feedback will result in the acquisition of generic timetabling abilities, as would be indicated via more efficient design performance on the far transfer task in comparison to a control group, remains unclear. Issues surrounding the specificity of skills acquired during such practice activities may limit the transfer of these skills. However, should transfer be evident, a greater level of positive transfer would be expected from those undergoing five, as opposed to one, practice task.

6.2.1 Method

Participants

Forty-eight psychology undergraduates, aged between 18 and 25, with an average age of 20.00 years (SD 2.09) took part in the experiment in return for course credit. The sample consisted of 10 males and 38 females. None had any experience with the experimental task. Participants were randomly allocated to one of the three experimental conditions.

Materials

Experimental materials comprised of three experimental timetabling tasks: a practice task, a near transfer task and a far transfer task. These tasks contained 20 classes (design elements) to be scheduled into a five day timetable, each day containing four hour-long class timeslots (falling between 9am & 1pm). In addition, a small timetabling copying task and accompanying print-out of a completed timetable containing 12 classes, was used to familiarise participants with the interface used. Each timetabling task was programmed in Java, had accompanying instruction files, and was compiled into the differing running orders required for each experimental group. Two PC monitors were used to display the two experimental screens (an information screen displaying timetabling rules and availability and also a timetabling screen displaying an empty timetable and the classes to be scheduled, see Figures 2.1 & 2.3). Seven types of general constraint, presented as timetabling rules, were displayed in every timetabling task. These were:

1. Classes must be scheduled in chronological order (i.e., Biology I must precede Biology II)
2. Classes must be scheduled so that theory classes precede practical classes (i.e., Music Theory must precede Music Practical)
3. The same teacher cannot be scheduled into consecutive timeslots if the classes are taking place in different locations

4. Teachers should not be scheduled for periods when they are not available
5. Rooms should not be scheduled for periods when they are not available
6. The number of students per class should not exceed the capacity of the room allocated
7. The required class facilities must be met by the room allocated to that class

In the practice and the near transfer tasks only 5 general constraints were applicable. These were timetabling rules 1, 2, 4, 5, and 6. In addition, the same number of specific constraints, 67 (inferred from the potential number of task errors that could be made), were derivable from the timetabling rules stipulated. As such, practice and near transfer tasks were similar in underlying task structure. Surface differences were introduced by using different class subjects and by renaming all teachers, facilities and class room names. The far transfer task contained seven general constraints, namely all of the timetabling rules listed above, and 97 derivable specific constraints. Surface differences were also utilised. As such the far transfer task was relatively dissimilar both structurally and superficially from the practice and near transfer tasks. A PowerPoint mathematics task containing multiple numeric equations was also used. Each equation, presented for 20 seconds, required the multiplication, division, addition or subtraction of two numbers, each with two digits.

Design

A between-subjects design was used with three levels of the independent variable, practice schedule; five practice trials, one practice trial, or no practice (henceforth referred to as the 5 trial group, the 1 trial group and the control group respectively). The practice task was not accompanied by any extrinsic information concerning task progress. A transfer task with the same underlying task structure as the practice task acted as the near transfer task. A relatively more complex task, incorporating more general and specific constraints acted as the far transfer task. Positive transfer on a far transfer task should be harder to achieve as it indicates that some general task principles, or heuristics, have been acquired. Positive near transfer performance is indicative of a degree of task-specific learning and should be easier to achieve. Dependent measures included the number of classes placed that do not violate any of the applicable timetabling rules (referred to as successful class placements) and the number of classes not scheduled at all (referred to as omissions) in both the near and far transfer task.

All experimental tasks were given a time limit of 12 minutes. This was done to control for the large variation in design time seen in some of the previous experiments (i.e., Experiment 1), so that each practice trial was approximately equivalent in the amount of task experience garnered. This factor also led to the emergence of the new dependent variable, class omissions. Counterbalancing of the near and far transfer tasks was not deemed appropriate here. It was considered impractical to have participants perform a task of greater complexity and then switch back to a simpler task, as this may lead to asymmetrical transfer effects. A possible outcome of this may be that task presentation order becomes confounded with the nature of the transfer task. Nevertheless, it seems practical to assess the closer form of transfer, near transfer performance, before assessing far transfer performance. An irrelevant mathematics task was used in place of a practice trial for the control group to ensure these participants did not have an undue advantage due to lack of experimental fatigue.

Procedure

The control group were instructed on the nature of the mathematics task and then left to correctly record the answer to as many equations as they could in 12 minutes, with the time limit here matched to the maximum time allocated to each timetabling task to maintain an equivalent fatigue effect. Following this procedure, they were familiarised with the timetabling interface and instructed how to place classes by copying the small timetable from paper to the screen. Then control participants performed the near and far transfer task.

Both experimental practice groups were first familiarised with the experimental interface using the copying task. The 5 trial group then performed the practice task five times whilst the 1 trial group performed it only once. Following this, both groups performed the near and far transfer tasks. For all practice, near and far timetabling tasks, participants were given 12 minutes to schedule as many classes as they could, as accurately as possible.

6.2.2 Results

In order to assess whether practice, and amount of practice, had an effect on timetabling performance, it was important to assess whether dependent measure scores differed between the different practice schedules for each transfer task. However, first it is important to determine whether, during the first practice trial, those in the practice groups had approximately equivalent levels of performance as would be expected by random allocation of participants. Independent samples *t*-tests were performed on scores for both dependent performance measures on the first practice task undertaken. No significant difference in successful class placements between the 1 trial group (mean = 11.63, SD = 4.03) and the 5 trial group (mean = 12.43, SD = 4.19) was evident, *t* (30) = -.56, *p* = .58. A further *t*-test indicated no significant difference for omissions on the first practice task, *t* (30) = -1.39, *p* = .17 (mean = 3.43, SD = 3.88, and mean = 5.13, SD = 2.89, for the 1 trial group and 5 trial group respectively). Whilst not indicating equivalent performance on first exposure to the practice task, these results indicate that performance was comparable statistically. As such, any differences in transfer task performance may be attributed to differing practice schedules.

Table 6.1 displays the mean successful class placements and omissions for all experimental groups on both transfer tasks. Table 6.1 indicates that the mean number of successful class placements in the near transfer task was higher for those who had undertaken timetabling practice previously. No similar trend was found for the number of successful placements in the far transfer task. In regards to omissions, those in the control group made consistently greater omissions in both the near and far transfer tasks.

Table 6.1. The effect of practice schedule on near and far transfer tasks

			Control (n = 16)	**1 trial group (n = 16)**	**5 trial group (n = 16)**
Near transfer Task	Successful class placements	Mean	10.88	16.06	16.62
		SD	5.32	2.62	3.84
	Omissions	Mean	4.62	1.00	1.50
		SD	5.12	1.15	2.53
Far transfer Task	Successful class placements	Mean	13.56	12.94	14.00
		SD	5.82	3.5	2.73
	Omissions	Mean	2.75	1.69	1.31
		SD	4.55	2.41	2.18

In order to determine whether statistical differences exist, between-subjects one-way ANOVAs were carried out. In regards to the near transfer task, there was a significant effect of practice schedule on the number of successful class placements, F (2, 45) = 9.64, MSE = 16.68, $p < .001$, $f = .17$. βonferroni post-hoc tests revealed that the control group achieved significantly lower successful class placements than both the 1 trial group ($p < .01$) and the 5 trial group ($p < .001$). The two practice groups did not have significantly different scores (p = 1.00) here. In regards to class omissions for the near transfer task, there was also a significant effect of practice schedule, F (2, 45) = 5.45, MSE = 11.33, $p < .01$, $f = 93$. βonferroni post-hoc tests indicate that the control group made significantly more omissions than both the 1 trial group ($p < .05$) and the 5 trial group ($p < .05$). Again, practice groups did not differ significantly from each other (p = 1.00).

In respects to the more complex, far transfer task, there was no apparent effect of practice schedule on successful class placements, F (2, 45) = .25, MSE = 17.93, $p = .78$.

Neither was there an effect of practice schedule on the number of omissions, *F* (2, 45) = .85, MSE = 10.44, *p* = .43).

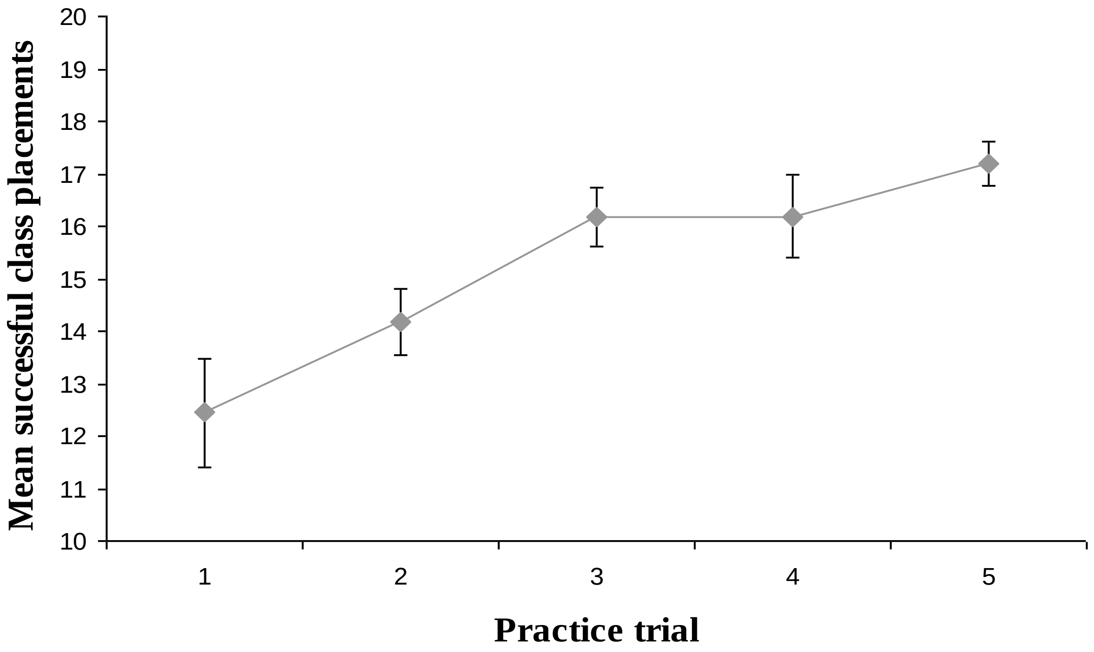


Figure 6.1. The effect of practice on successful class placements for the 5 trial group. Error bars are +/- 1 standard error.

Further analyses were carried out to assess whether performance improvements were evident across practice trials for those in the 5 trial group, as would be indicated by increasing number of successful class placements and fewer omissions. Figure 6.1 displays the mean successful class placements across practice trials in the 5 trial group. A trend towards increasing numbers of successful places is indicated. A within-subjects one-way ANOVA indicated a significant effect of practice trial on successful class placements, *F* (4, 15) = 12.76, MSE = 4.56, *p* < .001, *f* = .74. Further βonferroni post-hoc comparisons revealed that significantly more classes were successfully placed in practice trials 3, 4 and 5 in comparison to the first practice trial (*p*s < .01, < .05 & < .001 respectively). Successful placements in practice trial 2 were significantly lower than successful placements in practice trial, 3 (*p* < .01) and 5 (*p* < .001). Successful placements in practice trials 4 and 5 were not significantly different (*p* = 1.00). In regards to omissions, Figure 6.2 indicates that the mean number of omissions fell across the first three practice trials before assymptoting. A within-subjects one-way ANOVA indicated a significant effect of practice trial for omissions made, *F* (4, 15) = 14.79, MSE = 3.65, *p* < .001, *f* = .99. βonferroni post-hoc comparisons revealed that significantly more omissions were made in the first practice trial in comparison to the 3rd (*p* < .001), 4th (*p* < .01) and 5th (*p* < .001) practice trials. Omissions made in practice trial 2, 3 and 4 did not significantly differ from omissions in any subsequent practice trials.

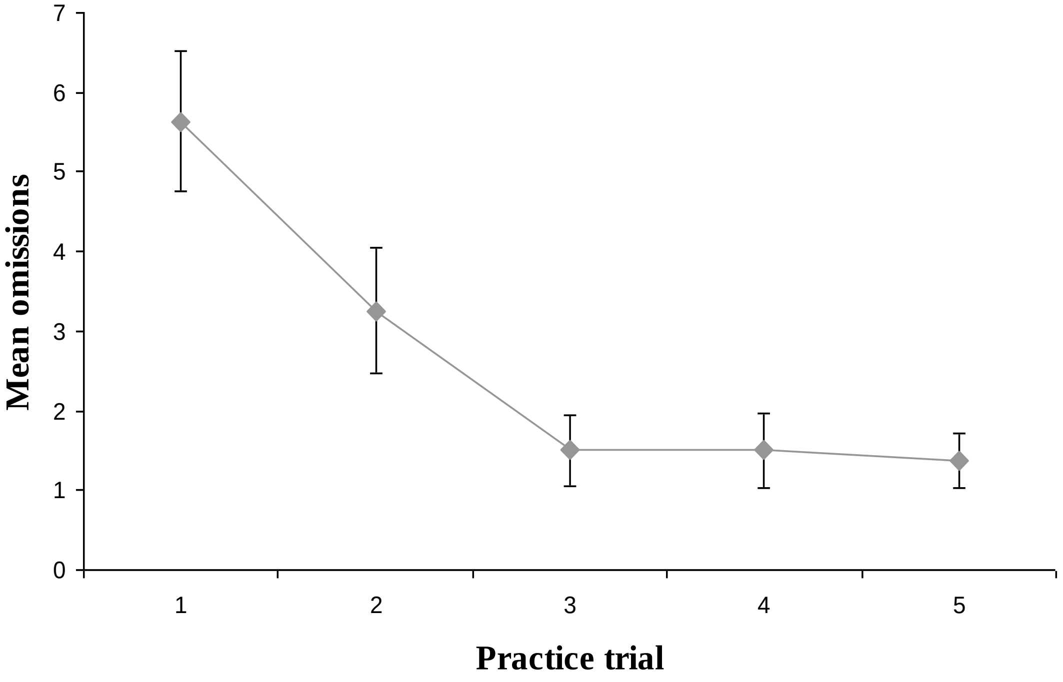


Figure 6.2. The effect of practice on omissions for the 5 trial group. Error bars are +/- 1 standard error.

6.2.3 Discussion and conclusions

Results indicate that undertaking any practice on a constraint satisfaction task, where only intrinsic feedback was available, led to improved performance on a near transfer task in comparison to a control group. This was true of both successful class placements and class omissions. Performance on a far transfer task did not significantly differ amongst the differing practice schedules for either dependent measure. These results suggest that transfer of skills acquired during practice stages may only occur when tasks are structurally similar to tasks performed in training (Holding, 1965; Yamhill & McLean, 2001). As such, a practice-task, transfer-task specificity is likely to be occurring. Subsequently, acquisition of generic timetabling skills or heuristics, applicable to far transfer tasks varying only in the incorporation of a more complex task structure, are not evident here (Schmidt & Bjork, 1992). Extrapolating from this, should a far transfer task with even greater task dissimilarity (Barnet & Ceci, 2002), such as an office design task that features both contextual and structural differences, little beneficial performance effects may be expected.

Whilst no solid conclusions can be made regarding the underlying mechanisms supported via practice, it is presently suggested that the process of engaging in practice may lead to reduced cognitive resource demands in the near transfer task resulting in performance benefits. Here, practice has enabled participants to familiarise themselves with the task rules (constraints) and the task environment. As such, participants need not expend cognitive resources gauging task requirements. This supposition is in line with the theoretical propositions of Paas (1992), and Paas *et al.* (1994) who suggest that practice may act to focus information on crucial task information.

Other experimental results indicate, contrary to expectations, that the number of practice trials undertaken (1 versus 5 trials) had little effect on any of the dependent performance measures. Therefore, in regards to performance in the near transfer task, simply being exposed to just one practice task was sufficient enough to bring about the significant improvements in performance in comparison to performance shown by the control group. However, there was some evidence indicating that those undertaking five practice trials showed improvements in performance as a function of subsequent practice trials. Within successive practice task trials, statistical improvements were seen across the first three trials for both successful class placements and omissions. Schmidt and Bjork (1992) suggest that blocked practice, such as that experienced by the 5 trial group, may lead to temporary performance improvements that disappear once test conditions change. It is presently proposed that the performance improvements seen across practice trials may be the result of enhanced memory for the specific task components. When test conditions alter, the enhanced memory for the practice task should afford no additional benefit when performing the near transfer task. This enhanced memory may constitute a further form of task-specific learning. It is possible that multiple practice trials may be of benefit in a constraint satisfaction context should it entail introducing more variability into the practice tasks (Holladay & Quiñones, 2003). Such variability should discourage specific task learning and may encourage the development of general, and more widely applicable, constraint satisfaction strategies.

There are a couple of methodological limitations within the present experiment that could potentially impact upon the conclusions drawn here and should therefore be acknowledged. Firstly, the lack of counterbalancing of the transfer tasks could be problematic. Counterbalancing was considered impractical given the increasing task complexity of the far transfer task, and the potential for asymmetrical transfer effects. However, the lack of counterbalancing may have raised the possibility of an effect of presentation order. It may be the case that performance decrements in the last task, the far transfer task, shown by both of the conditions undertaking practice trials, may be due to fatigue. However, performance for the transfer tasks indicated the reverse pattern of performance for the control group. As such, no consistent effects of presentation order can be concluded.

To summarise, it appears that within a timetabling constraint satisfaction task, exposure to a practice task, and the intrinsic feedback made available, leads to some improved design performance outcomes. Multiple practice trials may lead to better performance during the practice phase but does not appear to offer any additional performance increments in transfer tasks. However, performance in the near transfer task for those who had undertaken any practice, improved in comparison to those with no previous task exposure. This trend did not extend to the far transfer task where practice offered no advantage. Here, the control group, having simply been exposed to the near transfer task performed just as well as participants with more extended practice. It may be that the similarity between the practice and near transfer task, both incorporating the same set of applicable general constraints, underlies the performance improvements seen in the near, but not the far transfer task. Hence a level of task-specific learning is demonstrated. As such, it may be the case that short term practice interventions are not sufficient enough to produce beneficial far transfer effects. The following chapter aims to investigate whether metacognitive training interventions can offer advantages above those gleaned through practice. Here, in addition to intrinsic task feedback, designers will be prompted to evaluate their design strategies explicitly. In doing so, the extrinsic feedback generated should prompt the evaluation of cognitive strategies. This change in cognitive behaviour should result in performance benefits outweighing those seen by those produced via intrinsic feedback.

CHAPTER SEVEN

Metacognitive training

The previous chapter indicated that intrinsic feedback, provided by practice as a stand-alone training intervention, is of limited benefit within constraint satisfaction contexts. Lussier and Shaddick (2006) propose that the processes engaged in during practice are distinct from those instigated through practice plus training instructions. Training involves additional explicit instruction aimed at eliciting cognitive activity that isolates and improves upon problematic aspects of performance. Lussier and Shaddick further suggest that practice without additional training procedures, may not be an effective method of inducing performance improvements. This reflects the distinction that can be made between intrinsic feedback that is available in any practice activity, and the extrinsic feedback that arises due to explicit instructions to engage in additional performance enhancing processes (Patrick, 1992). In accordance, attention in the present chapter focuses on more strategic training interventions eliciting extrinsic feedback. Pashler, Capeda, Wixted and Rohrer (2005) suggest that providing feedback concerning progress towards a task goal during a practice or learning phase may improve performance above and beyond that shown via practice alone. The experiments that follow examine whether metacognitive interventions aimed at prompting participants to explicitly self-explain aspects of their performance may improve constraint satisfaction design.

7.1 Introduction to metacognitive training

Metacognition has received growing attention in the field of problem solving. In order to assess the potential impact metacognitive training interventions may have within design constraint satisfaction problems, it is first important to discuss what is meant by the term, metacognition.

Metacognition is the application of metacognitive knowledge, that is, knowledge about one's own cognitive processes, to a goal-directed activity (Ainsworth & Burcham, 2007; Veenman, Van Hout-Wolters & Afflerbach, 2006). Various forms of metacognitive activity are detailed in the psychological literature. These include but are not limited to, self-explanation, monitoring, self-observation, self-regulation, critical thinking, error management, reflection and evaluation (Ainsworth & Burcham, 2007; Helsdingen, van den Bosch, van Gog, & van Merriënboer, 2010; Keith & Frese, 2008; Meijer, Veenman, & van Hout-Wolters, 2006; Pintrich, Wolters & Baxter, 2000; Wetzstein & Hacker, 2004). There appears to be some conceptual overlap between some of these activities. Some of these constructs encompass a number of distinguishable metacognitive processes. Ainsworth and Burcham (2007) describe self-explanation as the process of generating additional knowledge beyond what is laid out in the task specification, a process involving a degree of monitoring, reflection and evaluation. Critical thinking is the process of testing for missing or conflicting information and evaluating options (Helsdingen *et al.*, 2010), which also entails monitoring and reflection. Likewise, error management cognition involves encouraging learning/performance improvements through encouraging errors and rectifying them (Keith & Frese, 2008), again requiring active monitoring and reflection. Indeed, various differing descriptions of metacognition have been offered since Flavell's (1979) original explanation of metacognition. Flavell (1979) distinguished between the metacognitive activities of planning, monitoring and evaluation that may be utilised before, during or after task performance. Despite differing and often overlapping terminology, it is apparent that metacognition is a multi-faceted construct (Veenman *et al.*, 2006) that can entail a mixture of processes that may be difficult to distinguish. As such, it is important, when referring to metacognitive training, to attempt to identify which of its many processes may be being undertaken.

Some preliminary research attests to the positive influence of metacognition in design contexts. Wetzstein and Hacker (2004) found that having designers explain their strategies by describing, justifying and evaluating them to a naive partner produced better quality designs than a control group. Here Wetzstein and Hacker (2004) provided one of the first examples of the beneficial effects of metacognition in design contexts.

Elsewhere, evidence of the positive effects of metacognition (in various forms) is well-documented, especially in learning and problem solving contexts. Some examples are now discussed. Ahlum-Heath and Vesta (1986) found that naive problem solvers forced to justify (design, explain and evaluate) their moves on Tower of Hanoi (ToH) problems did better on further, extended ToH tasks. Here problem solvers, having justified previous performance were less likely to make excess moves, a further sign of more efficient performance. Keith and Frese's (2008) meta-analysis of error management training (EMT) found that across the 24 studies surveyed the mean effect of EMT was significant in a positive direction (*d* = .44), with evidence of larger effect sizes for transfer performance rather than within task performance. Keith and Frese suggest that EMT promotes metacognitive activity as it encourages individuals to reflect upon the processes that lead to errors. As such the individual is encouraged to evaluate and revise strategies. Tajika, Nakatsu, Nozaki, Neumann and Maruno (2007) found that children solving mathematical word problems utilising self-explanation techniques, whereby they had to explain every problem solving step, outperformed control and self-learning groups (instructed to attempt to understand every problem step rather than explain it). Finally, Pennequin, Sorel, Nanty and Fontaine (2010) studied children undergoing metacognitive training for maths problems where they had to evaluate the importance of various problem solving strategies such as repeating calculations, or ranking information, when reaching solutions. Pennequin *et al.* found that those undergoing this form of metacognitive training had better awareness of their own skill use subsequently and achieved better problem-solving scores. They further found that this metacognitive training was especially beneficial for previously low maths achievers. Collectively, the evidence cited here indicates that metacognitive training exerts positive effects in multiple problem-solving and learning contexts. It also indicates that metacognitive training can have positive effects for individuals of all ages, particularly when expertise is lacking, as may well be the case in novel design tasks.

There is some suggestion that metacognitive skill(s) may not be domain specific but may be generic (Veenman *et al.*, 2006; Wetzstein & Hacker, 2004). Indeed, the small range of studies cited above indicate that metacognition may be beneficial across a range of tasks of differing complexity, from structured mathematical problem solving, to ill-structured design problems. Meijer *et al.* (2006) acknowledge that metacognition plays an important role in problem solving as it may aid in the successful deployment of appropriate strategies. Therefore, metacognitive training may be expected to also exert some beneficial effects when attempting to reach solutions in constraint satisfaction paradigms. The question left unanswered, is how metacognitive methods may lead to improvements in constraint satisfaction performance, and what methods are best for eliciting metacognitive activity.

Various explanations of the underlying processes through which metacognition produces beneficial effects have been proposed. Ainsworth and Burcham (2007) suggest that active construction of knowledge is vital for understanding. Metacognitive techniques may encourage such active engagement through internalising task principles by supporting the revision of relevant knowledge and by stimulating inference generation. Tajika *et al.* (2007) suggest that certain metacognitive instructions (self-explanation and self-examination) may act to determine problem misunderstandings and generate more relevant problem information. Wetzstein and Hacker (2004) offer three ways in which metacognition may lead to improved performance. It may act to induce a more systematic and analytic way of thinking that may result in an expansion of the problem space. It may direct attention to relevant problem features and evoke new problem representations. Finally, it may be a crucial feedback procedure that may motivate a problem solver to extend extra effort in reaching a solution. Alternatively, it may be a mixture of multiple beneficial effects producing performance improvements. Overall, in order to bring about positive effects, metacognition must somehow encourage the revision of faulty mental models (Ainsworth & Burcham, 2007; Keith & Frese, 2008). This is in line with action theory that stipulates that adequate mental models can only be acquired by actively engaging and dealing with relevant subject matter (Keith & Frese, 2008).

In order for any metacognitive training intervention to be successful within constraint satisfaction contexts, the evaluative processes evoked must help to encourage an accurate and comprehensive mental representation of task relevant information. Appropriate methodology is crucial. Veenman *et al.* (2006) propose that a fundamental principle for successful metacognitive instruction is that problem solvers should be instructed as to the usefulness of metacognitive processes to encourage them to exert the additional effort needed to benefit from metacognitive instruction. Also, prolonged training is advised to elicit maintained metacognitive activity. Various methods have been used to train and elicit metacognitive processes including questionnaires, interviews, think-aloud protocols, observations, stimulated recall, mental integration, questioning and also self-explanation (Pannequin *et al.*, 2010; Veenman *et al.*, 2006). To induce concurrent metacognition whilst undergoing some form of practice task, think-aloud, questioning and self-explanation methods appear most appropriate (Ericsson & Simon, 1998; Tajika *et al.*, 2007). Think-aloud processes involve the voicing of thoughts out loud without the need for any explanation of the content, a technique believed to be non-intrusive, needing no additional effort to produce (Ericsson & Simon, 1993, Hacker & Dunlosky, 2003). These techniques are used to gain insight into an individuals' natural thought processes as they perform a cognitive task. As such, think-aloud techniques should not be expected to spontaneously induce metacognitive activity unless accompanied by further instruction to self-explain. Neuman and Schwarz (1998) suggest that whilst metacognitive methods such as self-explanation and think-aloud may appear only marginally different, self-explanation involves the active and symbolic processing of information in working memory, whereas think-aloud is simply a passive commentary. More strategic self-explanation methods, such as questioning (originating from self or other), prompt the problem solver to evaluate crucial task rules and encourages critical evaluation of strategic progress in the search for solution. As the paradigms used here are heavily dependent on rule and constraint satisfaction, self-explanation should be the most appropriate way to elicit metacognition. Having participants self-explain the potential presence of any constraint violations should cause participants to reflect and evaluate upon their progress. Any errors uncovered should lead participants to re-evaluate the accuracy of their task representation and task strategy.

The experiments that follow investigate whether participants undergoing a metacognitive training intervention utilising prompted, self-explanation techniques, can outperform those with no such training (control groups) and those simply given practice with no further instruction. It is expected that the process of being required and prompted to self-explain, and the metacognitive content this will elicit, will encourage problem solvers to adopt a more analytic approach, to more accurately address conflicting information and identify errors, to internalise constraint rules more constructively, and to develop a more accurate understanding of the task (Ainsworth & Burcham, 2007; Keith & Frese, 2008; Tajika *et al.*, 2007; Wetzstein & Hacker, 2004). This is expected to result in improved performance in a transfer task in line with some of the empirical evidence cited above (i.e., Ahlum-Heath & Vesta, 1986; Keith & Frese, 2008).

7.2 Experiment 7

Experiment 6 indicated that simple practice, but not repetitive practice, leads to some improvements in performance on a structurally similar, near transfer task, but not a structurally dissimilar, far transfer task. It was concluded that practice provides only intrinsic feedback which may be of little utility in regards to the acquisition of flexible constraint satisfaction strategies. As such, without further training instruction, practice may not be most effective way to improve general constraint satisfaction performance. The present experiment now examines whether metacognitive training interventions, in the form of reflective self-explanation, where designers are required and prompted to explicitly monitor and evaluate task progress, may offer performance benefits above and beyond those gained from having engaged in practice alone.

As discussed in the previous section, metacognition is a multifaceted construct that may entail many processes aimed at improving performance in cognitive tasks. Metacognition may aid the active construction of task knowledge (Ainsworth & Burcham, 2007), may help to highlight any misconceptions (Tajika *et al.*, 2007), and is generally thought to induce more analytical, systematic and effective strategies (Wetzstein & Hacker, 2004). Growing bodies of research attest to the beneficial effects of metacognition in a variety of contexts, including ill-structured design (i.e., Wetzstein & Hacker, 2004). The question left unanswered is can metacognitive training interventions prove beneficial within constraint satisfaction design?

In the present experiment, three groups of participants were each presented with an office design task which was preceded by either an irrelevant task (control group), a practice task without further training instruction (practice group) or a practice task accompanied by metacognitive instructions to reflectively self-explain the utility of the previous design move (the reflective self-explanation group). More specifically, following each design move, participants were asked to consider which of the specific constraints stipulated were satisfied, any that may have been undone, and those yet to be considered. Such a metacognitive self-explanation training intervention should prompt participants to adopt a more analytic style, to proceed more strategically, and to implement explicit constraints more accurately. Self-explanation should also prompt participants to check for any faulty assumptions concerning the implementation of explicit constraints. As such, it is expected that performance on a subsequent transfer task, would be improved for the reflective self-explanation group in comparison to the performance shown by both the practice and control groups. The transfer task selected currently was an enlarged version of the training task. It differed by incorporating different general constraints but also in that it incorporated another design element and a greater number of specific constraints. In addition, designers were required to complete the transfer task as quickly and accurately as they could, whereas, training tasks were terminated after a set number of design moves. As such, the transfer task was dissimilar in a number of ways to the task used during training. Therefore, positive transfer here would be indicative of the acquisition of more general, widely applicable skills, considered a feature of far transfer (i.e., Barnett & Ceci, 2002).

7.2.1 Method

Participants

Seventy-two psychology undergraduates, aged between 18 and 31, with an average age of 19.78 (SD 1.95) took part in the experiment in return for course credit. The sample consisted of 14 males and 58 females. None had any experience with the experimental task. Participants were randomly allocated to one of the three experimental conditions.

Materials

Two office layout tasks adapted from Carroll *et al.* (1980) and programmed in Java were used. Both had the same representational format as previous office experiments. The task used for practice and training purposes enabled participants to view a set of 18 specific constraints, and seven employees (design elements). Constraints were evenly distributed among three general constraints. The particular constraint types used stipulated requirements of the positioning of employee offices based on car park location, lunch arrangements and noise levels, identical to general constraint set B used previously in Experiment 3. The office task used for the far transfer task contained 27 specific constraints, eight employees (design elements) and three differing general constraints concerning compatibility, status and the use of work areas, identical to constraint set A used in Experiment 3. CamStudio was used to record onscreen actions whilst participants produced their office designs. A Powerpoint maths presentation was created to act as an irrelevant task in place of practice or training trials for the control condition. Maths equations requiring the multiplication, subtraction, division or addition of two numbers ("40 + 73", "28 x 4", etc.) were formatted in such a way as to display each equation for 20 seconds.

Design

A between-subjects design with three levels of the independent variable, pre-task training, was used. A control group received no training in the task of interest, but performed an irrelevant maths task instead. Of the remaining two intervention conditions, a practice group received a brief practice task on a simplified office task with no explicit feedback provided. Finally, those in the metacognitive self-explanation group received the same practice task with additional instructions to explicitly and audibly monitor and evaluate changes in the constraints satisfied following each design move made. In particular, following each placement, or relocation of an employee's office, the experimenter instructed participants to consider which of the specific constraints stipulated were satisfied, any that may have been undone, and those yet to be considered. Participants voiced their explanations aloud. In doing so, they were required to check for any faulty assumptions concerning the implementation of explicit constraints and also evaluate the effectiveness of their present strategy.

All participants completed the far transfer task. The dependent variables of interest from this task were product measures, the number of constraints satisfied at the end of the design process and the design completion time (seconds). As a precautionary measure, participants' experience of the office task used in the experimental conditions involving practice or metacognition was kept approximately equivalent by limiting both tasks to 10 moves (as indicated by the placement, relocation or removal of a design element within the workspace). In addition, the time each participant in the practice and questioning group spent completing their training task, was yoked to the time each participant in the control group spent performing their irrelevant maths task. This was done to ensure a conservative estimate of potential fatigue effects from the first experimental task.

Procedure

Before the presentation of the first office design task participants were familiarised with the interface and instructed how to generate design features and the nature of the constraint types they may encounter. During the training task, those in the practice group were instructed to make ten moves (ten employee office placements). Those in the reflective self-explanation group were instructed to also make ten moves and assess the consequences of each move against the explicit constraints stipulated. In particular, participants were asked to consider which of the specific constraints stipulated were satisfied, any constraints that may have been undone, and those they had not yet considered. For the first five moves the experimenter actively prompted participants to engage in this process, and guided them through the process with further prompts and questioning whenever necessary. For the last five moves the experimenter simply prompted the self-explanation process and identified any constraints not categorised should the designer fail to acknowledge it in their explanation. The control group simply performed the maths task. Following on from training, all participants were instructed to try to complete the transfer task as quickly and accurately as possible. No time limit was set.

7.2.2 Results

Table 7.1 displays the mean scores on each dependent measure as a function of training condition. This table indicates that those in the control group took the longest to complete their design and also scored the lowest in terms of constraints satisfied. Those in the reflective self-explanation group had the highest average score for number of constraints satisfied and achieved this in the shortest average design time. Those in the practice group were intermediaries, falling between the control and the practice and questioning groups on both scores and completion times.

Table 7.1. The effect of training condition on number of constraints satisfied and design completion times

		Control (n = 24)	**Practice** (n = 24)	**Reflective self-explanation** (n = 24)
Number of constraints satisfied	Mean	17.75	18.87	20.42
	SD	2.40	2.53	1.35
Design completion time (seconds)	Mean	853.83	755.75	781.38
	SD	368.08	280.93	242.12

A one-way ANOVA revealed an effect of training condition on the number of constraints satisfied, *F* (2, 69) = 9.25, MSE = 4.65, $p < .001$, $f = .52$. Post-hoc βonferroni comparisons revealed that the control and practice group did not differ from each other significantly ($p = .23$). However, the reflective self-explanation group satisfied significantly more constraints than both the control ($p < .001$), and the practice ($p < .05$) groups.

A further one way ANOVA indicated no effect of training condition on design completion times (F (2, 69) = .68, MSE = 91009.47, $p = .51$). However, in response to the assertions of Norman and Bobrow's (1975) resource theory which proposes that as task complexity increases, speed-accuracy trade-offs should decrease, additional correlational analysis was undertaken. Should a task become more complex, progress towards task goals should be harder to monitor. This should act to increase variation in time taken to complete tasks. Birney and Halford (2002) propose that should a task prove to be more complex, the correlation between performance scores and times should diminish in comparison due to the greater variation in time. Therefore, correlations between time and constraints satisfied were calculated for the present experimental groups. Whilst there was a significant association between the number of constraints satisfied for the reflective self-explanation group ($r = .61$, $p < .01$), the correlation was not significant for those in the practice ($r = .19$, $p = .37$) or control group ($r = .19$, $p = .38$). This indicates that there was a clear pattern of increasing design times resulting in increasing numbers of constraints satisfied for the reflective self-explanation group only. In other groups, expending a greater amount of design time did not necessarily result in more constraints satisfied.

7.2.3 Discussion

Present results indicate a beneficial effect of undergoing a metacognitive self-explanation intervention, over and above the effects of simple practice in relation to the number of constraints satisfied. Whilst results indicated a trend for decreasing completion times as practice, and then practice combined with metacognitive self-explanation, was introduced, no significant effect was established. However, further examination of correlations between the number of constraints satisfied and design completion times indicated that these measures were only significantly associated in the reflective self-explanation condition. According to Birney and Halford (2002), correlations of greater magnitude, and greater significance, are found between time and score for tasks that are of lower complexity. Indeed, Norman and Bobrow (1975) suggest higher variability in task time may result from tasks with greater complexity. As such, speed-accuracy trade-offs, evident in many cognitive tasks, become less evident. As all participants completed the same transfer task, the positive and significant association between time and score for the metacognitive group is an indication that those undertaking practice with metacognitive self-explanation found the transfer task less complex than their counterparts in the practice and control conditions.

Results offer support to experimental predictions and are in line with previous evidence concerning the beneficial effects of metacognitive interventions (e.g., Ahlum-Heath & Vesta, 1986; Pennequin *et al.*, 2010; Wetzstein & Hacker, 2004). Further to this, present results offer an initial indication that metacognition may have a beneficial role to play within constraint satisfaction design. This builds upon the research of Wetzstein and Hacker (2004) who found metacognitive interventions to be beneficial in the related domain of ill-structured or creative design.

Why and how reflective self-explanation improves subsequent performance on a transfer task is not clear cut. Suggestions as to how metacognitive interventions exert beneficial effects vary. Some suggest metacognition may provoke the active construction of knowledge or encourage the internalising of task principles, vital for understanding (Ainsworth & Burcham, 2007). Others propose that metacognition may act to highlight problem misunderstandings and generate more relevant problem information (Tajika *et al.*, 2007). Alternatively, metacognition may direct attention to relevant problem features, may allow the generation of crucial task feedback, and may evoke new, more accurate, mental models (Wetzstein & Hacker, 2004). Whilst the present experiment cannot differentiate or disentangle all of these proposed effects, it can deduce that one benefit of metacognitive self-explanation training is that it acts to reduce task complexity. In order to do so, the metacognitive intervention must have acted in such a way so as to make participants focus their attention on crucial task features allowing them to be more systematic, as suggested by Wetzstein and Hacker (2004).

As in previous experiments (see Experiments 4 & 5) the lack of a difference in design time between experimental conditions is inconclusive. Whilst notions of frustration effects and the like (Payne & Duggan, 2011) may have applied to some designers, speculation of this nature does not support any firm conclusions. It may be the case that timing alone is not always a reliable indicator of performance. It may also be the case that should time limits have been imposed, exaggerations in the differences in constraints satisfied, between conditions, may have been found. Further studies would be needed to confirm this suggestion.

Presently, and akin to conclusions drawn in Experiment 6, it appears that practice alone is not sufficient to bring about beneficial performance increments in tasks that differ structurally from those used during practice. Constraint satisfaction, whilst containing a certain degree of structure compared to more creative and ill-structured design tasks, remains cognitively complex. The complexity may result in participants being unable, or not motivated, to utilise the intrinsic feedback available from engaging in practice without further instruction. Indeed, Beradi-Coletta, Buyer, Dominowski, and Rellinger (1995) suggest that individuals not in metacognitive treatments may rely on environmental feedback rather than self-monitor. Given the complexity of the current paradigm, and the dynamically changing task environment, environmental feedback and other potential sources of intrinsic feedback may prove hard to decipher and interpret in any beneficial way. What appears evident, due to the lack of differences between the control and practice group, is that trainees undergoing simple practice do not spontaneously improve.

A feature of the current experiment that should be acknowledged is the nature of the transfer task. Here, transfer was tested via performance on a task relatively dissimilar to the practice task. Both the foreword to the training experiments, and Experiment 6 discussed the features of transfer tasks that have been used to distinguish between near and far transfer (Barnett & Ceci, 2002). The transfer task used here was structurally different and also superficially different, as was the far transfer task in Experiment 6. As such, the reflective self-explanation techniques used presently produced a positive degree of far transfer. Whether positive transfer would also be shown in a structurally similar, near transfer task, is not clear from the present results.

The present experiment also has a couple of methodological considerations to be taken into account. As in previous experiments, the sample was psychology students, therefore results may not generalise well to experienced designers. Nevertheless, results highlight that metacognitive self-explanation may be of particular benefit within the training of novice designers. A further consideration is the use of verbalisation. Participants undergoing metacognitive training were asked to pause following each move during practice and explain their previous move in terms of changes in the satisfaction status of the specific constraints stipulated. It may the case that the act of verbalising, rather than the content elicited may have been what initiated the beneficial effects. In the introduction to this chapter, it was argued that self-evaluation techniques involve the active, symbolic processing of information in working memory, whereas think-aloud is simply a passive commentary (Neuman & Schwarz, 1998). However, others have argued that the act of verbalisation in experimental laboratory settings, where participants may expect such information to be scrutinised more closely, may be enough to alter cognitive processes (Bernadini, 2001) and induce more self-monitoring. As a practice accompanied by think-aloud condition was not included in the present experimental design, the possibility that the act of verbalisation itself, rather than the actual content elicited, may have led to some of the performance effects cannot be ruled out.

In conclusion, the present experiment indicates that metacognitive training, in the form of reflective self-explanation, can have a beneficial impact upon subsequent task performance in a constraint satisfaction context. Here, simply having participants pause after each design move and audibly evaluate how that move has aided their progress towards task goals, aids the transfer of skills acquired to a more complex task. One possibility not yet addressed is the timing of the metacognitive interventions. Whilst the retrospective, reflective intervention used presently exerted beneficial effects, whether a prospective metacognitive intervention would exert similar effects is unclear. It is expected that the present reflective intervention would prompt designers to adopt a degree of prospective planning, in order to produce a design move that they could favourably self-evaluate. However, a degree of prospective strategising, and the possible beneficial effects, can only be speculated. The experiment that follows aims to investigate more specifically whether prospective metacognitive methods can be advantageous in a constraint satisfaction context.

7.3 Experiment 8

The previous experiment indicated that retrospective self-explanation in training an office design task resulted in improved transfer of skills to a more complex task, in comparison to a simple practice and a control group. It was concluded that the process of engaging in reflective self-explanation resulted in effective metacognitive activity. Results indicated that this activity may have acted to reduce the cognitive complexity of the transfer task, either via the process of helping focus attention towards relevant problem information (i.e., Tajika *et al.,* 2007), and/or by invoking the development of a more efficient constraint satisfaction strategy (Wetzstein & Hacker, 2004). The present study aims to discover whether an explicit prospective metacognitive training, aimed at eliciting planning strategies and evaluating them aloud, may also prove beneficial when performing a subsequent transfer task. A subsidiary aim of the current experiment was to disentangle the effects of simple think-aloud verbalisations from metacognitive verbalisations, a limitation noted in the previous study.

Flavell (1979) distinguished between the metacognitive activities of planning, monitoring and evaluation that may be utilised before, during or after a task. These processes may not be mutually exclusive. Ideally they should be combined for more efficient metacognitive activity. However, the timing of such processes may vary throughout the problem-solving process. Intuitively, the act of planning aligns with a more prospective strategy whereas reflective evaluation would constitute a more retrospective strategy. This distinction raises the possibility that the timing of concurrent metacognitive self-explanation interventions may induce differing strategies. The metacognitive intervention in Experiment 7 was enacted after each design move and was reflective, and therefore retrospective, in nature. Should a metacognitive intervention prompt participants to review and revise their strategy before implementing a design move, a prospective strategy would be further encouraged, with attention focused on planning processes.

There is some evidence to attest to the beneficial effects of prospective metacognitive training interventions in the field of structured problem solving. Beradi-Coletta *et al.* (1995) conducted a couple of experiments whereby they interrupted problem solvers before they made a move in order to prompt metacognition. In their first study, groups of participants performed Tower of Hanoi (ToH) tasks. A metacognitive group received verbalisation instructions in the form of questions that had to be answered before each move (i.e., “How are you deciding which disk to move next? How are you deciding where to move the next disc? Do you know how good a move this is?”). Subsequent task performance was improved for this group in comparison to a think-aloud condition and control conditions. In a follow-up study, a silent metacognitive group were introduced. Here participants were asked to answer the same questions used previously in their heads. A six second delay was enforced before any move could be implemented to encourage metacognitive activity. Results indicated that despite the lack of vocalisation, the metacognitive group still evidenced improved performance over other conditions. The authors reported that performance for the silent metacognitive group was equivalent to that of the verbalised metacognition group produced in their first study.

As yet, no empirical evidence of prospective metacognitive interventions, aimed at encouraging planning, is available in the constraint satisfaction literature. It is possible that the complex nature of constraint satisfaction tasks may hinder any motivation to plan. Ormerod (2005) suggests that problems without a clearly defined structure (ill-structured, complex problems) do not lend themselves easily to planning due to the lack of a clear pathway to solution. There is also some suggestion that planning is not easily undertaken in an unfamiliar problem context (Delaney, Ericsson & Knowles, 2004). Within the present office design paradigm, planning may not occur due to the perception that it is computationally costly. Alternatively, attempts to plan may fail. Nevertheless, Ormerod (2005) implies that people are able to plan in complex situations. It may be the case that planning processes can become more comprehensive and more effective should problem solvers be forced to focus on planning and to make their planning strategy explicit. It is here that metacognitive interventions may prove beneficial and this is the focus of the following experiment.

One final feature of metacognitive studies remains unaddressed, the nature of verbalisations. Verbalisations can be elicited via simple think-aloud methods, or can entail verbalisations of metacognitive self-explanations. Ericsson and Simon (1998) concluded from a review of 30 studies that verbalisations elicited via think-aloud methods without any further training instruction do not systematically affect the thought processes of participants. However, no control for verbalising was incorporated into Experiment 7. Here, it may have been the case that the act of verbalising, rather than the process of verbalising metacognitive content, resulted in a change in cognitive processes (Smagorinsky, 1998). Such unintentional effects may be especially true of situations where participants perceive they are interacting with an experimenter (Bernadini, 2001), as would be likely in a laboratory situation, where a participant may feel their efforts are under scrutiny. As such, it is important that the effects of verbalisation and verbalisation plus metacognitive content, here centred on planning processes, be incorporated into Experiment 8.

In the present experiment, four groups of designers were assigned to differing training schedules. Participants either underwent no practice (the control group), practice without further instruction, practice plus concurrent think-aloud procedures (henceforth referred to as the verbalisation group), or practice plus a metacognitive intervention aimed at eliciting prospective planning via prompted self-explanation (referred to as the metacognitive planning group). The metacognitive intervention required participants to state their next proposed move, to consider whether a better move exists, and finally to justify what move they were implementing next. Dependent measures obtained were the number of constraints satisfied and the time taken to complete the design. It was expected that those in the metacognitive planning group would show the most favourable performance in a transfer task.

7.3.1 Method

Participants

Sixty-five psychology undergraduates, aged between 18 and 26, with an average age of 19.08 (SD 1.48) took part in the experiment in return for course credit. The sample consisted of five males and 60 females. None had any experience with the experimental task. Participants were randomly allocated to one of the four experimental conditions.

Materials

Materials used were identical to those used in Experiment 7.

Design

A between-subjects design with four levels of the independent variable, pre-task training was used. Training levels were no practice (controls), practice without further instruction, practice with verbalisation, and practice with metacognitive planning. The control group received no training in the task of interest, but instead performed an irrelevant maths task. Of the remaining three intervention conditions, a practice group undertook practice on a simplified office task (as used in Experiment 7). A verbalisation group performed this practice task whilst concurrently performing think aloud procedures. Finally, those in the metacognitive planning group received the same practice task with additional instructions to pause before each move and justify or adjust their proposed moved in order to maximise the number of constraints satisfied. In particular, before each design move, participants were required to state their proposed move, to consider whether they could propose a better move that satisfied more constraints, and then to justify the move they had decided to implement. Here, as in Experiment 7, experimenter prompted self-explanation techniques are used to elicit metacognitive activity. All participants completed the far transfer task used in Experiment 7.

The dependent variables obtained were the number of constraints satisfied at the end of the design process and the design completion time. As in Experiment 7, participants' experience of the task in intervention/training conditions was kept approximately equivalent by limiting both tasks to 10 moves (as indicated by the placement, relocation or removal of a design element within the workspace). In addition, the time each participant in the practice and metacognitive planning group spent completing their training task was yoked to the time each participant in the control group spent performing their irrelevant maths task. This was done to ensure a conservative estimate of potential fatigue effects from the first experimental task.

Procedure

Participants were familiarised with the interface and instructed how to generate design features. During the training task, those in the practice group were instructed to make ten design moves. Those in the verbalisation group were instructed to think aloud as they made 10 moves. Ericcson and Simon’s (1993) think aloud protocol procedure was followed. This involved administering think aloud practice activities until participants were comfortable with verbalising. An example activity was to think aloud whilst describing the number of windows in ones home. Those in the practice and metacognitive planning group were instructed to also make ten moves but were required to pause before each one to answer questions aimed at eliciting and evaluating their planning strategy. These participants were required to explain their choice of move and potential alternative move, then justify which move they chose to implement. The control group simply performed the maths task. Afterwards, all participants were instructed to complete the extended transfer task as quickly and accurately as possible. No time limit was set.

7.3.2 Results

One participant failed to complete the task and was excluded from analysis. Table 7.2 displays the means and standard deviations for all dependent variables. More constraints were satisfied in conditions where there was prior exposure to the office design paradigm. In addition, participants undergoing the metacognitive planning intervention satisfied the most design constraints. Design time evidenced a high level of variation. However, design time was quickest in the metacognitive planning group.

A one-way between-subject ANOVA revealed a significant effect of training condition on the number of constraints satisfied, F (3, 60) = 5.66, MSE = 4.45, $p < .01$, f = .53. βonferroni post-hoc comparisons revealed differences between the metacognitive planning condition and all other experimental conditions (ps < .05, < .01, & < .01 for the control, practice and verbalisation group respectively). There were no significant differences among any other conditions. A further one-way ANOVA indicated no significant effect of time taken to complete the design, *F* (3, 60) = 1.29, MSE = 152449.22, *p* = .29.

Table 7.2. The effect of training condition on number of constraints satisfied and design completion times

		Control (n = 15)	**Practice (n = 16)**	**Verbalisation (n = 17)**	**Metacognitive planning (n = 16)**
Number of constraints satisfied	Mean	17.93	18.00	18.59	20.62
	SD	1.98	2.56	2.15	1.63
Design completion time (seconds)	Mean	827.80	927.25	992.65	743.13
	SD	283.61	451.97	439.42	352.83

As in Experiment 7, correlations between design time and the number of constraints satisfied were calculated. Should a task prove to be more complex, the greater likeliness of speed-accuracy trade-offs, and the subsequent increase in design time variation, should result in reduced correlations (Birney & Halford, 2002; Normon & Bobrow, 1975). The correlation between time and constraints satisfied was strongest for the practice and metacognitive planning condition, *r* = .49, *p* = .06, however this result did not quite reach statistical significance. Correlations for the other experimental conditions were, *r* = .01, *p* = .96 for the verbalisation condition, *r* = -.36, *p* = .17, for the practice condition, and *r* = .22, *p* = .43, for the control condition.

Finally, as a crude indicator that participants in the metacognitive planning condition were engaging in the metacognitive training activities, the number of alternative moves proposed, as a result of the question, “can you propose a better move that satisfies more constraints?”, and the number of alternative moves implemented were examined. On average, participants proposed 1.53 alternative moves (SD 1.23). Of the ten training moves made by each participant, .88 (SD .99) were alternative moves that were implemented. Simply put, just over half of the alternative improved moves that were proposed were implemented. Whilst standard deviations indicate a lot of variation, these results indicate that the metacognitive intervention did impact upon design decisions.

7.3.3 Discussion

Results imply that undertaking a metacognitive training activity, aimed at prompting more effective planning via self-explanation techniques, led to a greater number of constraints satisfied in a transfer task than all other training schedules. Performance among the control, practice and verbalisation conditions did not significantly differ. This result offers further support to the proposal that specific forms of metacognition training can be beneficial in constraint satisfaction settings.

The mechanisms proposed to underlie the cognitive advantages offered via metacognition have varied in the literature. It may be that metacognition here has acted to highlight misunderstandings and generate more task relevant behaviour (Tajika *et al.*, 2007). Metacognition may aid the construction of knowledge and the internalisation of crucial task principles (Ainsworth & Burcham, 2007). Alternatively metacognition may elicit a more analytic approach where feedback concerning task understanding is sought (Wetzstein & Hacker, 2004). Whilst these results cannot conclusively distinguish between these mechanisms, the results do attest to the usefulness of metacognition for eliciting efficient planning strategies that may not be automatically undertaken by those performing complex tasks such as constraint satisfaction (Beradi-Colleta *et al.*, 1995). Planning is acknowledged to be difficult in ill-structured (Ormerod, 2005) and unfamiliar (Delaney *et al.*, 2004) tasks. Nevertheless, the present metacognitive intervention prompted planning behaviour. It did so by forcing participants to consider the ramifications of a proposed design move, to consider that an alternative more efficient move may exist, to propose such a move should one be found, and then to justify the move implemented. As a result a number of design moves were revised and positive transfer was evident.

As in Experiments 6 and 7, the practice group evidenced no beneficial effects of task-relevant experience on performance on a dissimilar, expanded transfer task. As a reminder, the transfer task used presently, whilst similar in context and administered in close temporal proximity, was also structurally different. This particular dissimilarity (Barnett & Ceci, 2002), along with the increased level of specific constraints (in comparison to the training/practice task, cf. Experiment 2), constitute a more complex, far transfer task. There is some suggestion that far transfer is harder to achieve than near transfer (Barnett & Ceci, 2002; Healy *et al*., 2006). As such, it may be speculated that beneficial effects of this metacognitive training intervention would produce similar, or perhaps greater, beneficial effects in a near transfer task where the underlying task structure and task complexity are equivalent to that used during training. No conclusions on this subject are possible until the merits of transfer to a near transfer task are tested empirically.

The present results overcome a limitation of the previous experiment. Here, a simple think-aloud verbalisation group was also examined, based on the notion that the act of verbalising, rather than the metacognitive content elicited, may be enough to produce the cognitive changes underlying task performance (Smagorinsky, 1998). Presently, despite the laboratory setting that could cause the perception that an interaction is taking place between a designer and experimenter (which Bernardini (2001) believes may spontaneously prompt metacognitive explanations), and despite the use of student participants who may be sensitive to monitoring behaviour in experimental settings (Ericsson & Simon, 1998), there was no beneficial effect of think-aloud. As such, it may be concluded that the improved performance of the metacognitive training group is due to the metacognitive content of the verbalisations elicited rather than the act of verbalising. Ericsson and Simon (1993) argue that metacognitive interventions require the active symbolic processing of information in working memory, in effect the introduction of a secondary processing level. On the other hand, simply verbalising is a passive activity of commenting on the information in working memory. It is unlikely to be a coherent or comprehensive account of the depths of processing engaged in, rather a simple commentary on momentary thoughts (Ericsson & Simon, 1998). Cooper, Sandi-Urena and Stevens (2007) distinguish between metacognition that is necessary for task understanding and cognition that is necessary for task performance. As there is evidence that individuals may not spontaneously engage in metacognitive activities (Gama, 2004), it seems likely that the think-aloud procedures used presently tapped into performance cognition rather than metacognitive task understanding.

Once again, no significant effect of training condition on design time was found. Design times varied considerably. It may be that the lack of a stringent criterion to attain before terminating design attempts may lead to the variation. As such, the decision to stop designing could have been based on multiple task aspects such as motivation to continue, frustration effects, the reduced ability to monitor progress, or the belief that progress has peaked (Payne & Duggan, 2011). Akin to Experiment 7, it is speculated that imposing a time restriction may exacerbate some of the differences in the number of constraints satisfied between conditions. Further empirical research would be needed to make conclusions regarding this supposition.

In addition to the non-significant effect of training schedule on design time, no significant correlations were found between the number of constraints satisfied and design times. However, the association between these measures approached significance for the metacognitive planning group. Birney and Halford (2002) suggest that positive correlations of a greater magnitude are indicative that the task was experienced as less complex. It may be insinuated, cautiously, that the metacognitive training undertaken by the metacognitive group acted to reduce the complexity of a subsequent transfer task.

As in previous experiments, the sample may be a limitation. Here designers may not be considered naive, as all will have had some experience of layout design. However none were expected to have any formal design experience. As such, these results may not generalise to expert designers. There is some suggestion that domain experts can be distinguished from those with less experience based on their verbalisations, as these experts tend to spontaneously produce verbalisations with metacognitive content (Eteläpelto, 1993; Veenman *et al.*, 2006). Should the present study be replicated using a sample of more experienced designers, the effects found may be expected to be smaller in magnitude. Further research would be needed to verify this proposition.

In conclusion, the present experiment has indicated that metacognitive activity concerning prospective performance can prompt planning activity in unfamiliar and complex constraint satisfaction tasks. Engaging in such self-explanation processes produces benefits in performance on a subsequent, and more complex, transfer task. Comparable effects are not seen in those conditions in which participants experience practice or think-aloud training methods. Whilst the exact mechanisms by which metacognition exerts effects are not known with any certainty, various processes such as focusing attention on crucial task features and inducing a more systematic approach have been proposed (i.e., Veenman *et al.*, 2006; Wetzstein & Hacker, 2004). It appears that the ease of application, and the generic applicability of metacognitive techniques to multiple tasks, makes it a prime candidate for further exploration in constraint satisfaction domains.

7.4 Conclusions

Experiments 7 and 8 have presented converging support for the notion that metacognition can play a beneficial role in improving constraint satisfaction performance. In particular, Experiment 7 indicated that reflective self-explanation, a retrospective intervention, whereby participants are forced to explicitly evaluate their progress so far, results in more constraints satisfied on a subsequent transfer task. Experiment 8 changed focus by examining whether prospective metacognitive interventions, involving experimenter-prompted self-explanation, could be used to elicit effective planning, a process that may not automatically be adopted given the degree of ill-structure and the unfamiliarity of constraint satisfaction tasks (Delaney *et al.*, 2004; Ormerod, 2005). Evidence of revised (but also explicit) planning was generated and it was concluded that the change in cognitions underlay increased performance on a transfer task. Experiment 8 further sought to disentangle the effects of simple think-aloud techniques from a metacognitive intervention. Here, simply verbalising thoughts aloud did not appear to alter cognitive processes, as indicated by non-significant differences between a verbalisation and a practice condition. Suggestions as to how metacognitive interventions exert beneficial effects vary. Wetzstein and Hacker (2004) suggest that engaging in metacognition induces a more analytic style. Constraint satisfaction tasks, such as that used presently, involve the application of logical rules and a good deal of decision making. A more analytic style should indeed be of benefit when undertaking such activities.

Both experiments in this chapter have found no advantage of simply having engaged in practice in an office design task. The previous chapter indicated that practice may aid performance on near rather than far transfer tasks. This difference may be due to the level of feedback afforded by practice without further instruction. Here, only intrinsic feedback is available. Within complex, constraint satisfaction contexts, it appears that environmental or contextual feedback is of little utility. It may be that this feedback is not helpful, or that it is unclear how to interpret such feedback in a useful way. In contrast, metacognitive self-explanation interventions provide useful extrinsic feedback. Here, a much more enriched form of feedback is available. The relevance of this metacognitive feedback is readily apparent as it informs the designer as to their current performance progress and also highlights any task misunderstandings. As a result, design strategy is improved when compared to those receiving only intrinsic feedback.

Much of the transfer literature attests to the difficulty of achieving far transfer (i.e., Barnett & Ceci, 2002). Definitions of what constitute a near and far transfer task differ. Presently, a near transfer task is defined as a task with the same structure, carried out in the same context and within the same discipline, here constraint satisfaction. Far transfer is a task within the same discipline and context, but possessing a differing, more complex structure. Therefore, the transfer task used in the present chapter would fall under the definition of a far transfer task. This may explain the lack of any beneficial effects of simple practice interventions. Nevertheless, the utility of metacognitive interventions in attaining a degree of far transfer (relative to training tasks) is demonstrated.

A methodological consideration of the present studies is the lack of the inclusion of a near transfer task. Both Experiments 7 and 8 suggest, due to evidence of the transfer of metacognitive skills to a far transfer task, that it is both probable and likely that beneficial performance effects should be expected for near transfer tasks. However, these speculative effects were not tested. There were a couple of methodological reasons for this omission. Firstly, as indicated in Experiment 4, there were a limited number of general constraints developed for use in the office task paradigm. As each office design task typically incorporates three, and repetition of participant exposure to these general constraints was avoided, a new set of general constraints would need to be developed. This would have required time-consuming validation effects in order to prove that any one set of general constraints were not idiosyncratic. Alternatively, a near transfer task could have been produced by constructing a reduced version of the far transfer task. Here the number of specific constraints and the number of design elements may have been kept constant relative to the practice task. This would reduce the differences between these transfer tasks to the various types of general constraints applied, and the need to complete the task rather than perform a subset of task moves. However, this would also act to reduce the level of dissimilarity between the near and far transfer tasks (due to re-use of the same general constraints), resulting in a less stringent test of far transfer. Therefore, for practical purposes, and in order to preserve and implement the most conservative far transfer test, no near transfer test was used.

A further methodological consideration is the lack of design time limits. Both Experiments 7 and 8 indicated no significant effects of training schedule on design time due to large variation in time taken to complete designs (although both studies indicated stronger, but not necessarily significant associations between time and score in metacognitive training conditions). Therefore, the usefulness of using design time as a dependent measure is called into question. An alternative approach, that may have produced a more sensitive test of training schedule on more important measures of design efficiency, most notably the number of constraints satisfied, may have been to impose a time limit on each design task.

As throughout this thesis, conclusions are limited by the nature of the samples. Here designers, although not necessarily naive, will have little formal design experience. Due to the suggestion that domain experts tend to spontaneously produce verbalisations with metacognitive content (Eteläpelto, 1993; Veenman *et al.*, 2006), present results may not generalise to expert designers. Nevertheless, results indicate that metacognitive training may be of particular benefit to those beginning their training for careers in design. Here educational programs may benefit from incorporating metacognitive teaching interventions.

Despite these methodological considerations, the present experiments have methodological plus points that should also be acknowledged. Presently, little research has examined metacognitive interventions in design contexts. One exception is Wetzstein and

Hacker (2004), who had designers self-explain and justify their designs to another person after which they could chose to revise their design. Here, a differing, more fine-grained, move-by-move metacognitive intervention has been examined. The advantage of doing so may be to prevent any need for complete design revision, as design strategies, and subsequent design progress, is honed on a more regular basis.

To conclude, both retrospective and prospective metacognitive interventions have proved beneficial in respects to the number of constraints satisfied on a subsequent far transfer task. In both instances, metacognition was elicited via experimenter-prompter, self-explanation. Both of these metacognitive interventions have been carried out at the design move level of analysis, that is, before or after a design move was implemented. Flavell (1979) highlighted that metacognitive activities can occur before, during or after an activity. As such, the present studies take a close-up examination of prospective and retrospective techniques used during training. A possibility not yet explored is whether retrospective or prospective techniques before or after a training task could induce similar beneficial effects. Whilst such techniques may be prone to memory inaccuracies, Wetzstein and Hacker (2004) have indicated that a reflective intervention of this nature aided designers in an ill-defined design context. Future research could explore whether these effects may generalise to constraint satisfaction contexts.

CHAPTER EIGHT

General discussion

8.1 Summary of experimental findings

The current series of experiments has addressed two qualitatively different aspects of performance in constraint satisfaction paradigms. The first series of experiments, Experiments 1-5, investigated whether constraint satisfaction performance was affected by variation in the number and nature of constraints, with some further distinctions made between the numbers of constraint, the number of differing general constraint types and the interplay between constraints and the external representation. A summary of these experiments is now given (see Table I in Appendices for an overview of experimental aims, design and results).

Experiments 1 and 2 examined whether variation in the number of constraints produced performance differentiation using both the timetabling and office design paradigm. Both experiments reported that a higher number of constraints led to deterioration on a number of important measures of design efficiency. Experiments 3 and 4 investigated whether variability among design elements, and the number of differing types of general constraint applicable to them, produced differences in design efficiency within the office design paradigm. Here, detrimental effects of having a greater number of constraint types applicable to each design element were found on the number of constraints satisfied. There was also some evidence that greater variability had a detrimental effect on design completion times. Experiment 5 investigated whether distinctions between the nature of certain constraints, and their interplay with the external representation, effects design performance. Here, performance with differing levels of fixed and non-fixed constraints was contrasted. Results indicate that participants undertaking an office design task with a greater proportion of fixed spatial constraints satisfied more of those constraints. A linear effect, with little deviation was apparent. No such effect was found for design time.

A second theme to the experimental work within this thesis was constraint satisfaction training (see Table II in Appendices for an overview of experimental aims, design and results). Given the paucity of empirical evidence concerning human cognition in constraint satisfaction contexts, but the extensive documentation of the utility of metacognitive interventions in various contexts (i.e., Veenman *et al.*, 2006), including some preliminary evidence in design contexts (i.e., Wetzstein & Hacker, 2004), metacognitive interventions aimed at inducing more analytic and systematic approaches to constraint satisfaction design were explored. However, before this was attempted, the potential for individuals to spontaneously improve was investigated. Experiment 6 explored whether designers show improvements in constraint satisfaction performance via practice without further instruction using the timetabling paradigm. Here, only intrinsic feedback is available to inform cognitive strategy. Results indicate that one practice trial was enough to produce a beneficial result on a near transfer task but not far transfer. Indeed, no positive far transfer was achieved.

Experiments 7 and 8 then documented attempts to improve office design performance through the introduction of metacognitive training interventions in comparison with control and practice interventions. Here, experimenter prompted self-explanation techniques were implemented. Experiment 7 contrasted a control group and a practice group with a group performing reflective metacognitive training. Experiment 8 aimed to investigate whether prospective metacognitive training is also beneficial in constraint satisfaction design. Here, an additional verbalisation group was included to control for potential effects of simply voicing thoughts aloud. Both experiments found benefits of metacognitive interventions on the number of constraints satisfied in a following far transfer task. In both instances, control and practice (and in Experiment 8 also verbalisation) groups did not differ. There was no effect of training on design time. However, correlations indicated a greater association between time and constraints satisfied for the metacognitive interventions, indicative that designers undertaking metacognitive training found the transfer task that followed to be less complex than their counterparts in other conditions.

8.2 General conclusions and implications

The research carried out presently was a preliminary attempt to address a gap in the design literature concerning human performance in constraint satisfaction contexts. Whilst not a new area of human endeavour, much of the published research document attempts to automate the process, and focus mainly upon the utility and development of computer algorithms (i.e., Burke & Petrovic, 2002; Burke *et al.*, 2006). The experimental work documented here highlights a number of useful and informative cognitive effects within the domain of human constraint satisfaction design. Several aspects of task constraints have been explored allowing a number of conclusions to be drawn. As such, some headway has been made into uncovering the difficulties entailed in constraint satisfaction.

Collectively, Experiments 1 to 4 indicate that increasing the number of constraints, or the number of qualitatively differing types of constraints, produces performance deterioration. Here, the size of the rule space increases (Simon & Lea, 1974; Zhang & Norman, 1994). This rule space then interacts with the state space to determine the overall size of the problem space (Newell & Simon, 1972) and the complexity entailed in navigating through it. Experiments 3 and 4 also highlight that the consideration of multiple rule types can be problematic. Theories of interactivity (i.e., Sweller *et al.*, 1990) and relational complexity (i.e., Halford *et al.*, 1998) suggest that it may not be the number of items that need to be processed that impact upon task difficulty, but the number of differing items that relate to a decision and need to be considered in parallel. Therefore, the interrelations between qualitative dimensions of the rule space are also a source of performance variation.

Experiment 5 added a little to what can be deduced about the rule space. The distinction between fixed and non-fixed constraints interacted with the level of affordance offered by the external representation (i.e., Gero & Kannengiesser, 2012). The result was differing levels of cognitive fit (i.e., Vessey, 1991). Experiments 1-5 collectively indicate that constraints do matter. Constraint information can affect design performance in various ways, from taxing mental resources, introducing additional complexity, or by being differentially supported by external representations.

All of these findings may have implications for the teaching of constraint satisfaction design. Teaching strategies involving the introduction of subsets of constraints may be expected to alleviate the cognitive load that arises from higher number of constraints. It is unclear whether such a strategy may alleviate the cognitive complexity that arises due to the number of differing constraint types. Here, introducing subsets of constraints may only alleviate task complexity should the differing constraint types be introduced incrementally. This is unlikely to alleviate the overall task relational complexity once all general constraints are revealed. Nevertheless, an incremental increase in relational complexity and also cognitive load may be preferable to an overload of both from the outset. In addition, knowledge and awareness of the difficulties that may be encountered when multiple constraint types apply to a design element may help novice designers to more systematically identify where difficulties are likely to arise and may provoke them to proceed with more caution. In respects to the implications of differential cognitive fit (Experiment 5), design students may benefit from tuition concerning the differences in constraint processing demands and the affordances of differing forms of external representation. They may further benefit from instruction and training in how to transform an external representation into a format that best supports design cognition (Jones & Schkade, 1995).

Experiments 6 to 8 examined whether differing training schemes could bring about improvements in constraint satisfaction performance. Results between differing training groups were interpreted in light of the distinction between intrinsic and extrinsic feedback (Patrick, 1992), and also via the properties of near and far transfer (cf. Barnett & Ceci, 2002). Presently, practice with only intrinsic feedback and no further instruction, proffered performance benefits on near transfer tasks only. As such, it appears brief practice interventions produced a relatively specific form of task learning. No evidence of the acquisition of general, flexible learning was shown. As such, practice alone may not be a suitable training intervention should flexible and adaptive design behaviour be desired. Experiments 7 and 8 focused on metacognitive self-explanation techniques that provide additional extrinsic feedback aimed at strategically changing cognitive behaviour. The growing field of research documenting the utility of metacognitive interventions appeared a promising field. These interventions are relatively easy to develop and implement (as an experimenter is not required to manually and effortfully compute task progress), and appear to prompt individuals to approach and carry out their tasks more strategically and more effectively (i.e., Wetzstein & Hacker, 2004). Both reflective self-explanation and self-explanation of prospective planning strategies produced beneficial constraint satisfaction when contrasted with control, practice or verbalisation conditions. The proposed mechanisms by which metacognition exerts beneficial effects vary within the literature. Whilst the exact micro processes involved may be specific to the intervention and the task undertaken, a generic outcome of engaging in metacognitive activity is that a more analytic and systematic approach is induced whereby feedback concerning task understanding is utilised productively (Wetzstein & Hacker, 2004). Present metacognitive interventions forced participants to scrutinise their strategies on a move by move basis. In doing so, attention was focused on computing task progress as the design was compiled. Here any faulty reasoning assumptions would have been highlighted and adjusted early on.

The training section of this thesis indicates that designers can be trained to perform constraint satisfaction to better effect, and highlights metacognitive activity as an efficient method of achieving this. This conclusion may have implications for the training of novice designers. Designers should be familiarised with metacognitive techniques such as explicit self-explanation, and also taught the potential benefits of its application. Indeed, Ainsworth and Burcham (2007) indicated that informing problem solvers of the benefits of metacognition should encourage the uptake of metacognitive strategising. In addition, metacognitive strategies should focus attention on the most complex aspects of task performance. Given the semi-structured nature of constraint satisfaction, planning and performance monitoring are the most complex and effortful processes. As such, the metacognitive interventions examined presently have entailed explicit monitoring and planning. However, in design contexts where structure is lacking (see Table 1.2), problem structuring (i.e., Restrepo & Christiaans, 2004) may need to take precedence over subsequent planning and progress monitoring. In such contexts, prospective metacognitive interventions may need adjusting accordingly.

8.3 Limitations and methodological considerations

The present stream of research has some limitations that need to be acknowledged. One, already discussed previously in depth, that may not require repeat description here, is the utilisation of undergraduate students as experimental participants. Other limitations are now addressed.

The use of design completion time as a dependent measure did not prove as informative as initially expected. Indeed, results concerning design times were somewhat disappointing. It was expected that greater task difficulty/complexity would result in longer design times, as participants would need to invest more time to reach a satisfactory solution. At the very least, clearer speed-accuracy trade-offs were expected. This was not often the case. The lack of clear stopping criteria, a typical feature of design tasks (Goel & Pirrolli, 1992) may have resulted in the absence of findings in completion times. In uncertain problem solving environments, when faced with a problem of considerable complexity, there are several reasons why an individual may terminate problem solving efforts whilst having not reached a satisfactory solution (Payne & Duggan, 2011), or without demonstrating a preference for speed over accuracy. One such reason may be an inability to monitor whether forward progress is being made, or the evaluation that further efforts will not succeed in improving problem solutions. This could lead to frustration effects and individuals may be unmotivated to expend more energy in working towards task goals. It could be speculated that the inherent complexity of human constraint satisfaction activities may lead to such effects. It is believed that allowing individuals to select their own stop criteria, based on affective reasons or cognitive considerations may have led to the lack of interpretable results concerning time. Future research may consider the utility of using fixed time limits during which to examine design performance. Doing so, whilst losing design time as a dependent measure, may have the added benefit of reducing the level of noise caused in other performance measures by the large variability in design completion times.

Another methodological consideration that should be acknowledged is the difficulty of isolating effects concerning variation between the number of qualitatively different constraint types and their applicability to the design elements. The experiments documented in Chapter 4, concerning relational complexity of the type and number of interactions between design elements and general constraints, had some methodological issues. It appears that introducing variation in one form of relational complexity, necessarily introduces another, unintentional form of interactivity. In Experiment 3, variation in the number of general constraint types applicable to each design element also introduced variation in the number of design elements subject to each general constraint type. Likewise, in Experiment 4, variability between design elements and the number of differing constraint types also introduced variation between the numbers of specific constraints falling under each general constraint. It appears that within the office constraint paradigm, as utilised presently, relational complexity may vary on multiple dimensions. It is hoped that the strength of conclusions made are permissible given the converging evidence of Experiments 3 and 4.

A related limitation concerns the utility of the constraint satisfaction paradigms used presently. Both the office task and the timetabling task were developed specifically for the purpose of providing an experimental platform that allowed human constraint satisfaction to be examined. However, unforeseen limitations of these paradigms limited the scope of some experiments undertaken. As well as issues concerning intractable forms of relational complexity, a limited number of general constraints were available within each paradigm. Experiments 7 and 8 used two office designs tasks, utilising the majority of general constraints presently available. One version was utilised for practice and training purposes, the other as a type of far transfer task. A far transfer task was selected over a near transfer for reasons discussed in Section 7.4 (reasons centred on the idea that the acquisition of flexible, generic skills are of greater desirability). Whilst far transfer is believed to be harder to achieve (Barnett & Ceci, 2002; Gray & Oransanu, 1987), the achievement of any degree of far transfer does not necessarily allow the assumption that beneficial effects in near transfer will also be seen. Had more general constraints been developed it may have been possible to have both a near and a far transfer task, and this may have allowed more comprehensive conclusions to be drawn. However, the development of new general constraints, if possible, would have required an extensive amount of calibration. With the exception of Experiment 5, all office tasks contained two fixed and one non-fixed general constraint, as modelled on

Carroll *et al.*’s (1980) original use of the paradigm. The development of any new general constraints would require a rule with contextual face validity. Fixed constraints would also require the general rule to focus upon one of the few fixed representational anchors. It was felt that such an endeavour would not yield results.

Despite the limitations and methodological nuances noted above, the present experimental paradigms have provided a good basis for a preliminary investigation of constraint satisfaction cognition. Advantages of the current experimental platforms included their wide-applicability to everyday design situations. This allowed the examination of constraint satisfaction cognition without the need for great technical knowledge or extensive task familiarisation. Further practical benefits included the ability to examine the design process in a relatively compact design episode (with participation times varying from 30 to 75 minutes). Nevertheless, other methodologies may prove useful in furthering what may be examined in everyday constraint satisfaction contexts. Logistics, the organisation of supply or delivery networks, is also a constraint satisfaction activity. Such activities combine route-planning involving spatial and temporal constraints, with resource restrictions. Such a context allows a variety of constraint types and distinctions to be introduced and may alleviate some of the restrictions found within the current paradigm.

A final limitation is noted in regards to the conclusions made concerning the utility of undergoing multiple practice trials. In Experiment 6, blocked practice was used. Whilst this is not in itself problematic, the repeated use of the same practice task might be. It may be that the intrinsic feedback available through five practice trials is no greater than that offered by one. Holladay and Quiñones (2003) suggest that variability in practice is more beneficial in contributing to positive transfer. Variable practice was not examined presently. On the basis of results concerning non-variable blocked practice (that five trials has no benefit beyond that gained from one trial), subsequent training experiments incorporated practice conditions entailing only one practice task exposure. This may have precluded the possibility that variable blocked practice may produce improved performance than that seen after one practice trial.

8.4 Recommended future directions

The research detailed presently covers only a small aspect of human constraint satisfaction performance. As such there are many other avenues with the potential to offer further insight. Some avenues of discovery have already been alluded to in the previous section of this chapter. Firstly, one could examine whether performance on complex tasks, containing many constraints, may be improved by breaking the task constraints into subsets. These sets of constraints could then be introduced gradually. This should act to alleviate the cognitive load experienced in comparison to the load experienced when all constraints are presented simultaneously. Secondly, more research could be aimed at discovering whether alternative external representational formats are more advantageous when tackling office design tasks with a greater proportion of non-fixed spatial constraints. Graphical representations such as network maps may prove useful here. Alternatively, research could focus on the utility of using intermediate representations as a cognitive aide memoire. Indeed, in the field of structured problem solving, Jones and Schkade (1995) found that some individuals undergo the costly efforts of translating a problem representation to better capture problem information. A further recommendation for future research may aim to overcome one of the present limitations concerning the contents of practice trials. Comparing the benefits afforded by variable blocked practice with that of identical blocked practice may offer further insight into whether practice may induce some degree of far transfer. In addition, should greater advantage be offered by variable blocked practice, it would be prudent to examine whether training interventions such as the metacognitive ones used presently still produce more advantageous performance then variable practice. Another recommendation involves the greater use of qualitative data. For instance, think aloud protocols could be of greater use than distinguishing between passive commentary and active examination of cognitive strategies (see Experiment 8). Think aloud could be further utilised to identify heuristics used by those undertaking constraint satisfaction design. Analysis of such content may be useful in determining effective and ineffective constraint satisfaction strategies.

Finally, a number of other avenues may allow further insight into factors affecting transfer of training strategies. Whilst the present metacognitive training interventions have demonstrated positive effects, whether these effects are transient in nature is unclear. It may be that the training is a crutch for performance, an effect that may be compounded by the lack of any change in experimental context between training and transfer tests. Should this be the case, performance benefits may be expected to dwindle as the time between training and transfer tests increases. Further research would benefit from testing the longevity of performance benefits. A final, related line of enquiry may offer a way of enhancing the longevity of training interventions. Interest is growing in the role social media may play in enhancing the transfer of training (i.e., Volet, 2013). Given the complex nature of design tasks, even those with a greater degree of task structure such as constraint satisfaction design, training top-ups administered via social media may prove swift and cheap to develop. Here trainers could be available to recap on brief metacognitive interventions. Alternatively, group forums could be set up to allow trainees to act as both designer and experimenter. Whilst such initiatives would bring about issues concerning the motivation to use such a system, its uptake rate may act to moderate the association between training and retention.

REFERENCES

Aguinis, H. & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology, 60*, 451-474.

Ahlum-Heath, M. E., & DiVesta, F. J. (1986). The effect of conscious controlled verbalization of a cognitive strategy on transfer in problem solving. *Memory and Cognition, 14(3),* 281-285.

Ainsworth, S. & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17(3),* 286-303.

Akin, O. (2002). Variants in design cognition. In C. Eastman, M. McCracken, & W. Newstetter (Eds.), *Knowing and learning cognition in design education.* Amsterdam: Elsevier.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89(4),* 369-406.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94(2),* 192-210.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist, 48(1),* 35-44.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51(4),* 355-365.

Annett, J. (1991). Skill acquisition. In J. E Morrison. (Ed.), *Training for performance: Principles of applied human learning*. Chichester: Wiley.

Arthur Jr, W., Bennett Jr, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: a meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88(2)*, 234-245.

Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. *Design Studies, 20(2),* 131-152.

Baddeley, A. (1992). Working memory. *Science*, *255(5044)*, 556-559.

Ball, L. J., Evans, J. B. T., Dennis, I., & Ormerod, T. C. (1997). Problem-solving strategies and expertise in engineering design. *Thinking and Reasoning, 3(4),* 247-270.

Ball, L. J., Lambell, N. J., Reed, S. E., & Reid, F. J. M. (2001). The exploration of solution options in design: A ‘naturalistic decision making’ perspective. In P. Lloyd and H. Christiaans (Eds.), *Designing in context: Proceedings of the Fifth Design Thinking Research Symposium – DTRS.* Delft University Press, Delft, The Netherlands.

Ball, L. J., & Ormerod, T. C. (1995). Structured and opportunistic processing in design: a critical discussion. *International Journal of Human-Computer Studies, 43(1),* 131-151.

Ball, L. J., Ormerod, T. C., & Morley, N. J. (2004). Spontaneous analogising in engineering design: A comparative analysis of experts and novices. *Design Studies, 25(5),* 495-508.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128(4),* 612-637.

Bayazit, N. (2004). Investigating design: A review of forty years of design research. *Design Issues, 20(1),* 16-29.

Beradi-Coletta, B., Buyer, L.S., Dominowski, R. L., & Rellinger, E. R. (1995). Metacognition and problem solving: A process oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(1), 205-223.*

Bernardini, S. (2001). Think-aloud protocols in translation research. *Target, 13(2),* 241-263.

Besnard, D., & Lawrie, A. T. (2002). Lessons from industrial design for software engineering through constraints identification, solution space optimisation and reuse. In *Proceedings of the 2002 ACM symposium on Applied computing* (pp. 732-738). ACM.

Bilda, Z., Gero, J. S., & Purcell, T. (2006). To sketch or not to sketch? That is the question. *Design Studies, 27(5),* 587-613.

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(2),* 109-134.

Burke, E. K., & Petrovic, S. (2002). Recent research directions in automated timetabling. *European Journal of Operational Research, 140(2),* 266-280.

Burke, E. K., Petrovic, S., & Qu, R. (2006). Case-based heuristic selection for timetabling problems. *Journal of Scheduling, 9(2),* 115-132.

Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55(1),* 241-261.

Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(3),* 517-526.

Carroll, J. M., Thomas, J. C., & Malhotra, A. (1980). Presentation and representation in design problem-solving. *British Journal of Psychology, 71(1),* 143-153.

Chan, C. S. (1990). Cognitive processes in architectural design problem solving. *Design Studies, 11(2),* 60-80.

Chandrasekaran, B. (1990). Design problem solving: A task analysis. *AI Magazine, 11,* 59-71.

Chrysikou, E. G., & Weisberg, R. W. (2005). Following the wrong footsteps: Fixation effects of pictorial examples in a design problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5),* 1134-1148.

Cooper, M. M., Sandi-Urena, S., & Stevens, R. (2007). Reliable multi method assessment of metacognition use in chemistry problem solving. *Chemistry Education Research and Practice, 9(1),* 18-24.

Cross, N. (2001). Design cognition: Results from protocol and other empirical studies of design activity. In C. Eastman, M. McCracken & W. Newstetter (Eds.), *Design knowing and learning: Cognition in design education*. Amsterdam: Elsevier.

Davies, S. P. (2003). Initial and concurrent planning in solutions to well-structured problems. *The Quarterly Journal of Experimental Psychology, 56A(7),* 1147-1164.

Delaney, P. F., Ericsson, K. A., & Knowles, M. E. (2004). Immediate and sustained effects of planning in a problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(6),* 1219-1234.

Dorst, K. (2006).Design problems and design paradoxes. *Design Issues, 22(3),* 4-17.

Dorst, K., & Cross, N. (2001). Creativity in the design process: co-evolution of problem–solution. *Design Studies*, *22(5)*, 425-437.

Dym, C. L. (1992). Representation and problem-solving: The foundations of engineering design. *Environment and Planning B: Planning and Design*, *19(1),* 97-105.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education, 94(1),* 103-120.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (Revised Edition).* Cambridge, MA: Bradford Books/MIT Press.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity, 5(3),* 178-186.

Eteläpelto, A. (1993). Metacognition and the expertise of computer program comprehension. *Scandinavian Journal of Educational Research, 37(3),* 243-254.

Fernandes, R., Simon, H. A. (1999). A study of how individuals solve complex and ill-structured problems. *Policy Sciences, 32(3),* 225-245.

Field, A. (2009). *Discovering statistics using SPSS*. Sage Publications Limited, London.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34(10),* 906-911.

Gama, C. (2004). Metacognition in interactive learning environments: The reflection assistant model. *Lecture Notes in Computer Science, 3220,* 668-677.

Gero, J. S., & Kannengiesser, U. (2012). Representational affordances in design, with examples from analogy making and optimization. *Research in Engineering Design, 23,* 235-249.

Gibson, J.J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.). *Perceiving, acting, and knowing: Toward an ecological psychology*. Hillsdale, NJ: Lawrence Erlbaum.

Goel, V. (1994). A comparison of design and nondesign problem spaces. *Artificial Intelligence in Engineering, 9(1),* 53-72.

Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science, 16(3),* 395-429

Goldschmidt, G. (1997). Capturing indeterminism: Representation in the design problem space. *Design Studies, 18(4),* 441-455.

Goldschmidt, G., & Smolkov, M. (2006). Variances in the impact of visual stimuli on design problem solving performance. *Design Studies, 27(5),* 549-569.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organisations: Needs assessment, development and evaluation, (4th Ed.).* Wadsworth Group, Belmont, CA.

Gray, W. D., & Orasanu, J. M. (1987). Transfer of cognitive skills. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning: Contemporary research and applications*, 183-215. San Diego: Academic Press Inc.

Gray, W. D., Sims, C. R., Fu, W-T, & Schoelles, J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behaviour. *Psychological Review, 113(3),* 461-482.

Guindon, R. (1990). Designing the design process: Exploiting opportunistic thoughts. *Human-Computer-Interaction, 5(2),* 305-344.

Hacker, W. (1997). Improving engineering design: Contributions of cognitive ergonomics. *Ergonomics, 40(10),* 1088-1096.

Hacker, D. J., & Dunlosky, J. (2003). Not all metacognition is created equal. *New Directions for Teaching and Learning, 95,* 73-79.

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., Cowan, N. &, Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Science, 11(6),* 236-242.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral and Brain Sciences, 21(6),* 803-865.

Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne Jr., L. E. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(3),* 534-546.

Helsdingen, A. S., van den Bosch, K., van Gog, T., & van Merriënboer, J. J. G. (2010). The effects of critical thinking instruction on training complex decision making. *Human Factors, 52(4),* 537-545.

Holding, D. H. (1965). *Principles of training.* Pergammon Press, New York.

Holladay, C. L., & Quiñones, M. A. (2003). Practice variability and transfer of training: The role of self-efficacy generality. *Journal of Applied Psychology, 88(6),* 1094-1103.

Jones, D. R., & Schkade, D. A. (1995). Choosing and translating between problem representations. *Organizational Behavior and Human Decision Processes, 61(2),* 214-223.

Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology, 93(1),* 59-69.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17(2),* 248-294.

Kruger, C., & Cross, N. (2006). Solution driven versus problem driven design: Strategies and outcomes. *Design Studies, 27(5),* 527-548.

Kwon, S., J., Lee, K., C., & Mustapha, E, E, (2011). The role of cognitive map on influencing decision makers' semantic and syntactic comprehension, and inferential problem solving performance. *Intelligent Information and Database Systems: Lecture Notes in Computer Science, 6952,* 532-543.

Lee, G., Eastman, C. M., & Zimring, C. (2003). Avoiding design errors: A case study of redesigning an architectural studio. *Design Studies, 24(5),* 411-435.

Liikkanen, L. A., & Perttula, M. (2008). Exploring problem decomposition in conceptual design among novice designers. *Design Studies, 30(1),* 38-59.

Lussier, J. W., & Shadrick, S. B. (2006). Components of effective training. In *Virtual media for military applications*. Meeting Proceedings RTO-MP-HFM-136, Paper 3. Neuilly-sur-Seine, France, RTO. http://www.rto.nato.int/abstracts.asp.

Manz, H., Brunner, S., & Wullschleger, L. (2006). Triple vacuum glazing: Heat transfer and basic mechanical design constraints. *Solar Energy, 80(12),* 1632-1642.

Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation, 12(3),* 209-237.

Miller, G. A. (1956). The magical number, seven plus or minus two: Some limitations on our capacity for processing information. *Psychological Review, 63,* 81-97.

Moreno, R., Ozogul, G., & Reisslein, M. (2011). Teaching with concrete and abstract visual representations: Effects on students' problem solving, problem representations, and learning perceptions. *Journal of Educational Psychology, 103(1),* 32-47.

Morgan, P. L., Patrick, J., Waldron, S. M., King, S. L., & Patrick, T. (2009). Improving memory after interruption: Exploiting soft constraints and manipulating information access cost. *Journal of Experimental Psychology, 15(4),* 291-306.

Mullineux, G. (2001). Constraint resolution using optimisation techniques. *Computers and Graphics, 25(3),* 483-492.

Neuman, Y., & Schwarz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem-solving. *British Journal of Educational Psychology, 68(1),* 15-24.

Newell, A., & Rosembloom, A. (1981). Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (Ed.), *Cognitive skills and their acquisition.* Lawrence Erlbaum, Hillsdale, NJ.

Newell, A. & Simon, H. A. (1972). *Human problem solving.* Prentice-Hall Inc, NJ.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7(1),* 44-64.

Novick, J. R. (2001). Spatial diagrams: Key instruments in the toolbox for thought. *The Psychology of Learning and Motivation, 40,* 279-325.

Ormerod, T. C. (2005). Planning and ill-defined problems. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning (Vol. 1*). London: Psychology Press.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84(4),* 429-434.

Paas, F., Renkl, A. & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32(1),* 1-8.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. H. (2007). *Engineering design: A systematic approach (3rd Ed.).* Springer-Verlag, London.

Pashler, H., Capeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate the learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(1),* 3-8.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52,* 629-651.

Patrick, J. (1992). *Training: Research and practice.* Academic Press Limited, London, UK.

Payne, S. J., & Duggan, G. B. (2011). Giving up problem solving. *Memory Cognition, 39(5),* 902-913.

Pennequin, V., Sorel, O., Nanty, I., & Fontaine, R. (2010). Metacognition and low achievement in mathematics: The effect of training in the use of metacognitive skills to solve mathematical word problems. *Thinking and Reasoning, 16(3),* 198-220.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition.* Buros Institute of Mental Measurements, Lincoln, NE.

Purao, S., Rossi, M., & Bush, A. (2002). Towards an understanding of the use of problem and design spaces during object-oriented system development. *Information and Organization, 12(4),* 249-281.

Purcell, A. T., & Gero, J. S. (1996). Design and other types of fixation. *Design Studies, 17(4),* 363-383.

Purcell, A. T., Williams, P., Gero, J. S., & Colbron, B. (1993). Fixation effects: Do they exist in design problem solving. *Environment and Planning B: Planning and Design, 20,* 333-345.

Restrepo, J., & Christiaans, H. (2004). Problem structuring and information access in design. *Journal of Design Research, 4*(2), 1551-1569.

Römer, A., Leinert, S., & Sachse, P. (2000). External support of problem analysis in design problem solving. *Research in Engineering Design, 12(3),* 144-151.

Römer, A., Pache, M., Weiβhahn, G., Lindermann, U., & Hacker, W. (2001). Effort-saving product representations in design- results of a questionnaire survey. *Design Studies, 22(6),* 473-491.

Sabin, D., & Freuder, E. C. (1996). Configuration as composite constraint satisfaction. In *Proceedings of the Artificial Intelligence and Manufacturing Research Planning Workshop* (pp. 153-161).

Sachse, P., Leinert, S., & Hacker, W. (2001). Designing with computer and sketches. *Swiss Journal of Psychology, 60(2),* 65-72.

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52,* 471-499.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45(2),* 185-213.

Schaerf, A. (1999). A survey of automated timetabling. *Artificial Intelligence Review, 13(2),* 87-127.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3(4),* 207-217.

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 53–119). San Diego, CA: Academic Press.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychological Review, 19(4),* 469-508.

Schön, D.A. (1987). *Educating the reflective practitioner.* Basic Books, New York.

Shaft, T. M., & Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *Mis Quarterly, 30(1),* 29-55.

Shin, N., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astrology simulation. *Journal of Research in Science Teaching, 40(1),* 6-23.

Shute, V. J., Gawlick, L. A., & Gluck, K. A. (1998). Effects of practice and learner control on short-and long-term gain and efficiency. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40*(2), 296-310.

Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence, 4*(3), 181-201.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology, 30(1),* 363-396.

Simon, H.A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L.W. Gregg (Ed.), *Knowledge and cognition*. Hillsdale, NJ, Erlbaum.

Smagorinsky, P. (1998). Thinking and speech and protocol analysis. *Mind, Culture and Activity, 5(3),* 157-177.

Smith, G. F., & Browne, G. F. (1993). Conceptual foundations of design problem solving. *IEEE Transactions on Systems, Man, and Cybernetics, 23(5),* 1209-1219.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. C*ognitive Science, 12(2),* 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4(4),* 295-312.

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(2),* 176-192.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10(3),* 251-296.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th Ed.),* Pearson Education, Boston.

Tajika, H., Nakatsu, N., Nozaki, H., Neumann, E., & Murano, S. (2007). Effects of self-explanation as a metacognitive strategy for solving mathematical word problems. *Japanese Psychological Research, 49(3),* 222-233.

Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. *Annual Review of Psychology, 43,* 399-441.

Ullman, D. G., & D'Ambrosio, B. (1995). Taxonomy for classifying engineering decision problems and support systems. *Artificial Intelligence for Engineering, Design, Analysis and Manufacturing, 9(5),* 427-438.

van Merriërboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17(2),* 147-177.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M. & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition Learning, 1(1),* 3-14.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of graphs versus tables literature. *Decision Sciences, 22(2),* 219-240.

Vessey, I., & Gellata, D. (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research, 2(1),* 63-84.

Visser, W. (2004). *Dynamic aspects of design cognition: Elements for a cognitive model of design*. Research report 5144, INRIA, March 2004.

Visser, W. (2006). Designing as a construction of representations: A dynamic viewpoint in cognitive research. *Human-Computer-Interaction, 21(1),* 103-152.

Volet, S. (2013). Extending, broadening and rethinking existing research on transfer of training. *Educational Research review, 8,* 90-95.

Waldron, S., Patrick, J. & Duggan, G. B. (2010). The influence of goal-state access cost on planning during problem solving. *The Quarterly Journal of Experimental Psychology, 64(3),* 485-503.

Waldron, S. M., Patrick, J., Morgan, P. L., & King, S. (2007). Influencing cognitive strategy by manipulating information access. *The Computer Journal*, *50(6),* 694-702.

Wetzstein, A., & Hacker, W. (2004). Reflective verbalisation improves solutions – The effects of question-based reflection in design problem solving. *Applied Cognitive Psychology, 18(2),* 145-156.

Wickens, C. D., & Hollands, J. G. (1999). *Engineering psychology and human performance (3rd Ed.).* Prentice Hall, Upper Saddle River, NJ.

Yamnill, S., & McLean, G. N. (2001). Theories supporting transfer of training. *Human Resource Development Quarterly, 12(2),* 195-208.

Yeo, G. B., & Neal, A. (2004). A multilevel analysis of effort, practice, and performance: Effects of ability, conscientiousness, and goal orientations. *Journal of Applied Psychology, 89(2),* 231-247.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21(2),* 179-217.

Zhang, J. (2000). External representations in complex information processing tasks. In A. Kent (Ed.), *Encyclopedia of library and information science* (Vol. 68). Marcel Dekker, Inc, New York.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18(1),* 87-122.

APPENDIX

Table I. Performance variation in constraint satisfaction due to systematic manipulation of constraints

Experimental aims	Design	Results & conclusions
Exp 1 Timetabling task To establish whether the number of task constraints affects design performance. Hypothesis: Higher numbers of constraints lead to reduced design efficiency.	Within-subjects design. Participants completed three timetabling tasks that varied in the number of constraints: Low: 3 general & 48 specific constraints Medium: 5 general & 67 specific constraints High: 7 general & 97 specific constraints Product measures obtained: classes successfully placed, design time, end constraint violations. Process measures obtained: design moves, error violations throughout.	High constraint tasks resulted in significantly fewer classes placed successfully and in more errors (constraint violations during and at the end of the design process) from other experimental tasks. Performance between the low and medium constraint tasks did not differ significantly on these measures. In addition, significantly faster design completion times and a lower number of design moves were evident in the low constraint task. Here performance on the medium and high constraint task did not differ significantly. These results are consistent with notion that having a larger problem space containing more constraint information results in increased task complexity and a greater degree of cognitive load. Within each performance measure, there was no strict, linear decline in performance deterioration along with increasing numbers of constraints. Instead, threshold effects were apparent. Caveat: number of general and specific constraints co-varied.
Exp 2 Office design task To investigate whether the number of specific task constraints affect design performance. Hypothesis: Higher numbers of specific constraints lead to reduced design efficiency.	Between-subjects design. Participants completed one of three office design tasks that varied only in the number of specific constraints (all had 3 general constraints): Low: 12 specific constraints Medium: 24 specific constraints High: 36 specific constraints Performance measures obtained: proportion of constraints satisfied, design time, the proportion of time taken to satisfy each constraint and design moves.	Participants in the high constraint condition satisfied a significantly lower proportion of their task constraints and took longer to complete their designs than participants in the medium or low constraint conditions. High constraints also utilised more design moves than the low constraint condition. However this effect only approached significance. Performance between the low and medium constraint conditions did not differ significantly, indicating that there may be a threshold beyond which further constraints become problematic. Contrary to expectations, participants in the low constraint condition expended the most time satisfying each constraint. This result cannot be explained by problem space theory but may be indicative of the lack of stringent stop criteria. These results provide further support to the argument that having more information to contend with when searching the problem space leads to deterioration in design efficiency.

Table I. *continued…*

Exp 3 Office design task To investigate whether variation in the number of differing constraint types, but not specific constraints, affects design performance. Hypothesis: Greater levels of variability between design elements and the number of applicable general constraint types will result in deterioration in design performance.	Mixed within- & between-subjects design. Participants completed two office design tasks. Tasks contained 3 general constraint types and 18 specific constraints but differed in the interactions between design elements and general constraints. Low variability: each design element subject to 1 or 2 general constraint types High variability: each design element subject to all 3 general constraint types. Ordering of tasks was fully counterbalanced. Performance measures obtained include the number of constraints satisfied, the design completion time and the number of design moves made.	Performance in the high variability task was consistently worse. The number of constraints satisfied was significantly lower in this high variability task. In addition, the design time and the number of design moves utilised was significantly higher in the high variability task. Presentation had no effect on performance measures with the exception that there was an interaction between task variability and presentation order for the number of design moves utilised. Those who had completed a low variability task first required more design moves to complete a high variability task. Whilst no firm conclusions can be made, it could be suggested that strategies effective in the low variability task may not effectively transfer to the high variability task, but not vice versa. These results indicate that despite having an equivalent amount of specific and general constraint types per task, the interactions between a design element and the differing general constraints is a source of performance variability. Caveat: Co-variation of the number of differing constraint types per design element and the number of design elements per general constraint type.
Exp 4 Office design task To further investigate whether variation in the number of differing constraint types and design elements affect performance. Hypothesis: Greater variation will result in deterioration in design performance.	Between-subjects design. Participants completed either a low or a high variability task, each containing 24 specific constraints. Low variability: 3 differing general constraints High variability: 6 differing general constraints Performance measures obtained include the number of constraints satisfied and the design completion time.	Participants performing the high variability task satisfied significantly fewer constraints than those performing the low variability task. This supports the notion that the greater the interactivity, or relational complexity, between the number of differing constraint types and design elements, the more complex the design task becomes. No significant differences were found in the time taken to complete an office design. This result may have been indicative of a lack of clear stopping criteria but no firm conclusions can be made. Caveat: As in Experiment 3, there are alternative forms of variation (such as the number of general constraint types per task, or the number of specific constraints per general constraint) that may have contributed to the patterns of results.

Table I. *continued…*

Exp 5 Office design task Investigating whether the proportion of constraints referencing a fixed representation point, and the interplay with the external design representation, affects design performance. Hypothesis: Increasing the proportion of fixed constraints leads to improved cognitive fit resulting in greater design efficiency.	Between-subjects design. Participants completed one of four office design tasks. All contained 27 specific and 3 general constraints. The following proportions of fixed and non-fixed constraints were used: All fixed: 3 fixed general constraints 2/3 fixed: 2 fixed, 1 non-fixed general constraints 1/3 fixed: 1 fixed, 2 non-fixed No fixed: 3 non-fixed general constraints Performance measures included the number of constraints satisfied and the design completion time.	Results indicated that those undertaking an office design task with greater proportions of fixed constraints satisfied more of their task constraints. A linear effect with little deviation was apparent, indicating that task performance increased along with the proportion of fixed constraints. Results support the proposition that fixed points within the external problem representation may afford some computational offloading when implementing and monitoring sets of fixed specific constraints. As such, greater cognitive fit results when performing office design tasks containing mainly fixed constraints. The representational format provides no similar support when implementing and monitoring non-fixed constraints that stipulate the positioning of two employees based on neighbouring or surrounding offices. As in Experiment 4, no significant differences were found in the time taken to complete an office design. Again this may have been indicative of a lack of clear stopping criteria but no firm conclusions can be made.

Table II. Training and performance within constraint satisfaction tasks

Experimental aims	**Design**	**Results & conclusions**
Exp 6 Timetabling paradigm Investigating whether simple practice, and the repetition of it, has an effect on performance on near and far transfer tasks. Hypothesis: Practice should lead to improved constraint satisfaction performance, with greater benefits shown by those undergoing multiple practice trials	Between-subjects design. Experimental tasks undertaken differed according to experimental condition: A 5 trial group performed a practice task 5 times. A 1 trial group performed a practice task once. A control condition performed an irrelevant maths task. All participants then completed a near (similar to the practice task but with superficial differences) and far transfer task (different structurally). All tasks were limited to a maximum of 12 minutes. Performance measures included the number of successful class placements and the number of omissions (classes not scheduled).	For the near transfer test, the control group placed significantly fewer classes successfully, and omitted more classes from the end design than practice groups (who did not differ from each other). For far transfer, there were no significant group differences. In respect to the 5 trial condition, analysis of performance throughout practice trials indicated that significantly more classes were placed successfully and fewer omissions were made in the 5th trial in comparison to the 1st trial. Practice without further instruction offers only intrinsic feedback. It appears that such feedback and any learning or skills acquired during the practice phase may be specific to the structure of the task. As the structure of the near transfer task was identical, performance on this task improved as a function of practice. Multiple practice trials did not offer any advantage above that gained from one practice task. The lack of any beneficial effects of practice in the far transfer task indicates that no general timetabling skills were acquired.

Table II. *continued…*

Exp 7 Office design task Investigating whether a reflective metacognitive self-explanation method can produce beneficial transfer effects in comparison to a practice only intervention and a control condition. Hypothesis: A metacognitive self-explanation group will show higher levels of performance on a transfer task.	Between-subjects design. Experimental tasks undertaken differed: A reflective self-explanation group performed a simplified practice task with self-explanation prompts following design moves. A practice group performed the practice task without further instruction. A control group performed an irrelevant maths task. All participants performed a far transfer task. Performance measures included the number of constraints satisfied and the time taken to complete the transfer task. Additional measures involved the correlation between these measures.	Results indicated that the reflective self-explanation group satisfied more constraints than those in the practice or control group (who did not differ). No differences were found in design time. Correlational analysis of number of constraints satisfied and design times indicated a significant positive association for the reflective self-explanation group only. Here, stronger correlations are attributed to a clear speed-accuracy trade-off function, enabled by the lower task complexity experienced by the self-explanation group, relative to other experimental groups. Practice alone, did not produce any performance benefits compared to the control group. Collectively, results indicate that self-explanation techniques aimed at prompting metacognitive activity has beneficial effects on constraint satisfaction; it appears to improve performance and reduce perceived complexity relative to conditions where participants have undergone either practice without further instruction, or performed an irrelevant task. In order to achieve this, the present metacognitive intervention produced additional extrinsic feedback helpful in improving cognitive strategy and becoming more systematic. The lack of differences in design time may be indicative of the lack of clear stopping criteria. Caveat: Whilst the act of verbalisation is not expected to have induced the beneficial effects in the reflective self-explanation group, this cannot be ruled out.
Exp 8 Office design task Investigating whether a prospective metacognitive planning method can produce beneficial transfer effects in comparison to a practice only intervention and a control condition Hypothesis: A metacognitive planning group will show the highest levels of performance on a transfer task.	Between-subjects design. A metacognitive planning group performed a simplified practice task with prompts to evaluate and potentially alter proposed moves before they are implemented. A verbalisation group performed the practice task whilst thinking aloud. A practice group performed the practice task without further instruction. A control group performed an irrelevant maths task. All participants performed a far transfer task Performance measures included the number of constraints satisfied, the time taken to complete the transfer task and correlations between these measures.	The metacognitive planning group satisfied more constraints than other experimental groups (who did not differ). Correlational analysis found no associations between number of constraints satisfied and design completion times. However, the association within the metacognitive planning group approached significance and was the largest in magnitude. This may indicate that those undertaking the metacognitive training found the transfer task simpler in comparison to participants in other conditions. These results indicate that prospective metacognitive strategies are beneficial in eliciting planning strategies that are not automatically engaged in complex or unfamiliar tasks. As in Experiment 7, the mechanism by which metacognitive planning may produce benefits involves the production of useful extrinsic feedback acting to make participants focus their attention more systematically on crucial task features. As in Experiment 7, there was no effect of training schedule on time taken to complete a design, which may be indicative of the lack of clear stopping criteria. Once again, practice without further instruction was not sufficient to produce performance benefits on a far transfer task. In addition, thinking aloud during the practice task did not offer any benefits in terms of transfer performance. It appears that the present act of verbalising did not alter designers’ thought processes.