

**Mitigating the Effects of Implicit Constraints in Verbal
Insight Problem Solving through Training**

Afia Ahmed

BA (Hons), MSc, C Occ Psychol

A Thesis submitted for the
Degree of Doctor of Philosophy

School of Psychology, Cardiff University
70 Park Place
Cardiff, CF10 3AT

UMI Number: U584516

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U584516

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

I would like to extend my warmest gratitude to my supervisor, Professor John Patrick, for his time, patience and support. He has been an excellent teacher not just during this PhD but also throughout my professional career at Cardiff University. I am particularly grateful to him for providing me with the many opportunities that allowed me to develop my research and teaching skills.

I sincerely thank all the people who have provided assistance in some way during my research. In particular, I would like to thank the following undergraduate students: C. Brennan, B. Cotterell, M. Curry, S. Wilkinson, and R. Walker. Also, thank you to R. Chadwick, O. De Condappa, Dr. S. King, Dr. P. Morgan, T. Patrick, and V. Smy for their help in coding the verbal protocols in this thesis.

Thanks to all my 'old' university friends, particularly Eileen, Hazel, Laura, and Lorraine for their support. Thanks also go to Eman, Lisa, Michal and Phil. A big thank you to my childhood friend, Dr. Sheeja Joy, for her emails of encouragement and for frequently reminding me that 'it will be over soon'.

Thank you to my parents and siblings for their love and support throughout my academic life. Thank you especially to my mum and sister for being there for my son, Ibrahim, to enable me to have the time to write this thesis.

I would like to thank my husband, without whose encouragement, I never would have embarked on this journey. I have learnt so much from you since we started our lives together. Most of all, thank you for showing me that anything is possible when you have the belief and determination.

Last, but certainly not least, a very big thank you to Ibrahim, my beautiful boy. This PhD became a bigger challenge once you came along! You are a constant reminder that I have so much more to learn. Thank you for helping me to see the brighter side of life and for ending my days with a smile. I dedicate this thesis to you.

TABLE OF CONTENTS

Acknowledgements	ii
Table of Contents	iii
List of Tables	vii
List of Figures	viii
SUMMARY	ix
CHAPTER 1	
Introduction to insight problem solving	1
1.1 What is insight problem solving and how is it different?	1
1.2 Why is insight problem solving practically important?	14
1.3 The importance of training	17
1.3.1 The benefits of training in different domains	18
1.3.2 Theoretical approaches to transfer	20
1.3.3 Factors that inhibit positive transfer of training	21
1.4 Overview of thesis layout	23
CHAPTER 2	
Review of problem solving theories relevant to achieving insight	26
2.1 Problem Space Theory (Newell & Simon, 1972)	26
2.2 Gestalt Theory (e.g., Kohler, 1924; Wertheimer, 1945)	29
2.3 Representational Change Theory (Ohlsson, 1992)	36
2.4 Restructuring Change Theory (Knoblich et al., 1999)	42
2.5 Progress Monitoring Theory (MacGregor et al., 2001)	45
2.6 Summary	51
CHAPTER 3	
Review of methods that facilitate insight problem solving	54
3.1 Incubation	54

3.2	Hints approach	57
3.3	Training	60
	3.3.1 Specific training and positive transfer	61
	3.3.2 Generic training and positive transfer	65
3.4	Summary	72

CHAPTER 4

	Heuristic-based training for verbal insight problems	74
4.1	Heuristic-based training and transfer to verbal insight problem solving	74
4.2	Experiment 1	78
	4.2.1 Method	78
	4.2.2 Results and Discussion	85
4.3	Experiment 2	89
	4.3.1 Method	91
	4.3.2 Results and Discussion	93
4.4	General Discussion	97

CHAPTER 5

	Specific training for functional fixedness verbal insight problems	102
5.1	Divergent thinking and representational change in solving functional fixedness problems	102
5.2	Experiment 3	107
	5.2.1 Method	108
	5.2.2 Results and Discussion	113
5.3	General Discussion	117

CHAPTER 6

	Inconsistency checking as form of generic training for verbal insight problems	120
6.1	Inconsistency checking as a mechanism for representational change	120
6.2	Experiment 4	125
	6.2.1 Method	125

6.2.2	Results and Discussion	130
6.3	Conclusions	137

CHAPTER 7

Enhanced training in inconsistency checking for verbal insight problems		140
7.1	An iterative process of inconsistency checking	140
7.2	Explaining and justifying solution attempts during problem solving	141
7.3	Experiment 4	144
	7.3.1 Method	147
	7.3.2 Results and Discussion	150
7.4	General Discussion	156

CHAPTER 8

A measure of constraint activation in verbal insight problem solving		160
8.1	Representational change during insight problem solving	160
8.2	The Lexical Decision Task (Mayer & Schvaneveldt, 1971)	162
8.3	Pilot study	164
	8.3.1 Method	164
	8.3.2 Results	165
	8.3.3 Discussion	166
8.4	Experiment 6	167
	8.4.1 Method	168
	8.4.2 Results and Discussion	171
8.5	General Discussion	172

CHAPTER 9

General Discussion		176
9.1	Summary of experimental findings	176
9.2	Some implications	181
9.3	Limitations	183
9.4	Recommendations for future research	184

9.5	Conclusions	186
	REFERENCES	188
	APPENDICES	203
	Appendix A	203
	Appendix B	205
	Appendix C	206
	Appendix D	207
	Appendix E	209
	Appendix F	211
	Appendix G	212
	Appendix H	213
	Appendix I	215
	Appendix J	216
	Appendix K	217
	Appendix L	218

LIST OF TABLES

CHAPTER 4

Table 4.1	Effect of training on solution rate of problem categories (three problems in each category).....	85
Table 4.2	Incorrect hypotheses generated by condition (average rank)	87
Table 4.3	Solution rates for test problems (as a proportion)	96

CHAPTER 5

Table 5.1	Effect of training on solution rate of problem categories (three problems in each category)	114
-----------	---	-----

CHAPTER 6

Table 6.1	Types of training related behaviours in during problem solving	133
Table 6.2	Examples of questions verbalised by trained participants that were related to the problem statement	134

CHAPTER 7

Table 7.1	Table of design for Experiment 5	146
Table 7.2	Solution rates across conditions (out of seven problems)	151
Table 7.3	Types of hypotheses generated between conditions (across seven problems)	153

CHAPTER 8

Table 8.1	Mean and SD for target and non-target words for each problem	166
Table 8.2	Table of design for Experiment 6	170

LIST OF FIGURES

CHAPTER 1

Figure 1.1	Solution to the nine-dot problem	6
------------	--	---

CHAPTER 4

Figure 4.1	Percentage solution rate by conditions and problem category	86
Figure 4.2	Percentage solution rate by conditions and problem category	97

CHAPTER 5

Figure 5.1	Percentage solution rate by conditions and problem category	114
------------	---	-----

CHAPTER 6

Figure 6.1	Percentage solution rate by conditions and problem category	131
------------	---	-----

CHAPTER 7

Figure 7.1	Percentage solution rate by conditions and problem category	152
------------	---	-----

SUMMARY

The main focus of this thesis was to design training to mitigate the effects of constraints underlying verbal insight problem solving. Concurrent verbal protocols were collected. Experiments 1 and 2 tested training that was specific to solving problems that were exemplars of trained categories. In Experiment 1, heuristic training was provided for two categories of constraint: those with ambiguous words and those with human names that should be associated with animals. Transfer was positive for novel problems within the two trained categories but not for problems outside. Experiment 2 improved performance on problems with ambiguous words once shortcomings in Experiment 1 were addressed. Experiment 3 tested training that was specific to solving functional fixedness verbal insight problems. An iterative process of considering many functions of individual items was successful in facilitating performance but not for problems outside. Experiments 4 and 5 investigated the effectiveness of different types of generic training in facilitating solution of novel verbal insight problems. In Experiment 4, participants were trained to identify assumptions, which they made during problem solving and were inconsistent with the problem statement. In Experiment 5, participants were trained to iteratively consider different parts of the problem specification and to identify inconsistencies as in Experiment 4. Solution rate was improved in both experiments although instruction to explain and justify oneself during problem solving was also sufficient in facilitating performance in Experiment 5. Finally, Experiment 6 tested a novel method for identifying when a participant was constrained by an incorrect representation during verbal insight problem solving. The results supported that there

is variability in the nature of stereotypical constraints involved and demonstrated how training can be designed to induce restructuring or a shift in representation.

CHAPTER 1

Introduction to insight problem solving

The focus of this thesis is on a special form of problem solving entitled 'insight problem solving' and the overarching objective is to design training to facilitate solution of insight problems. This chapter begins with an introduction to the nature of insight problem solving and its importance. Next, training is discussed as a possible methodology for improving insight problem solving, which is followed by an overview of the thesis.

1.1. What is insight problem solving and how is it different?

Any form of problem solving entails generating novel solutions to problems that a person may never have encountered before, or has very limited or no knowledge of what is required to reach the correct solution (Dominowski & Dallob, 1995). Some well-known general definitions of problem solving include:

'Problem solving is defined as any goal-directed sequence of cognitive perspectives.' (Anderson, 1980, p. 257)

'...cognitive processing directed at transforming a given situation into a goal situation when no obvious method of solution is available to the problem solver.' (Mayer, 1990, p. 284).

Both definitions emphasise that problem solving involves cognitive processing to reach a particular goal, i.e., the solution. Further, Mayer's definition notes that the solver may not know how to progress with problem solving to reach the solution.

However, insight problem solving is a special type of problem solving in that the solver defaults to a high probability habitual conceptualisation that acts as a major impediment to solution leading to an impasse, a state of mind of not knowing how to progress to attain the solution (Ohlsson, 1992). In other words, knowledge that is normally activated by the problem specification or context is not relevant and has to be rejected, and information in the problem specification has to be interpreted in an unusual way in order to reach the solution (Duncker, 1945; Gick & Lockhart, 1995; Isaak & Just, 1995; Mayer, 1995; Wertheimer, 1959). Insight problems are thus ill-defined, in that both the goal and steps for its completion are unclear and the solution possibilities are endless (Chrysikou, 2006).

The definition of insight is problematic (MacGregor & Cunningham, 2008). Some propose that insight is characterised by the sudden illumination of the solution (Metcalfe & Wiebe, 1987) whereas others have focused on the psychological processes involved in achieving insight, namely restructuring or representational change following an impasse (Knoblich, Ohlsson, Haider, & Rhenius, 1999; MacGregor & Cunningham, 2008; Ohlsson, 1992). Consequently, several definitions of insight have emerged. According to Mayer (1995, p. 3)

‘the term insight has been used to name the process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve it.’

This definition notes that insight problem solving involves a sudden transition from not knowing to knowing the solution. However, it does not state that having insight will provide the solution to the problem; rather it places the solver on the correct

solution path. A criticism of this definition is that Mayer did not specify the process or processes that underlie what he refers to as insight.

Dominowski and Dallob (1995, p. 33) contend that

‘Insight will be characterised as a form of understanding (of a problem and its solution) that can result from restructuring, a change in a person’s perception of a problem situation.’

Ansburg (2000, p. 143) proposed that

‘Insight occurs when a solver restructures a previously intractable problem such that a new understanding of what needs to be done appears in the consciousness.’

The definitions of Dominowski and Dallob (1995) and Ansburg (2000) are similar in that both suggest insight occurs as result of restructuring, which may indirectly be suggesting the suddenness of solution as Mayer (1995). Dominowski and Dallob (1995) suggest restructuring results in a new ‘form of understanding’ but this does not necessarily mean that restructuring will result in insight. This problem is, however, avoided in Ansburg’s (2000) definition as restructuring must result ‘in a new understanding of what needs to be done’ to reach the solution.

Ohlsson (1992, p. 10) defined insight problems as those

‘... which have a high probability of triggering an initial representation which has a low probability of activating the knowledge needed to solve the problem.’

Ohlsson's (1992) definition is different to the aforementioned definitions as it focuses on why insight problems are difficult. It suggests that insight problems contain certain information within their specification that triggers a misleading representation that inhibits the information needed to solve the problem. Although not mentioned in his definition, Ohlsson (1992) also suggests that restructuring of the initial representation is required to facilitate solution.

Apart from Mayer's (1995) and Ohlsson's (1992) definitions, all the above definitions place importance on restructuring the faulty initial problem representation, without which solution is unattainable. This is what makes insight problem solving unique. The insight experience itself is characterised as follows: the solver is misled in his or her interpretation of a problem; the solver continues working on the problem until they reach an impasse in which problem solving ceases; the impasse is eventually overcome by a representational change and the solution of a problem occurs consequently (Ansburg & Dominowski, 2000; Chronicle, MacGregor, & Ormerod, 2004; Fleck, 2008; Jones, 2003; Kaplan & Simon, 1990; Ohlsson, 1984b, 1992; Ormerod, MacGregor, & Chronicle, 2002; Segal, 2004; Sternberg & Davidson, 1995).

One type of insight problem that has not been explicitly referred to as an insight problem is called a lateral thinking problem by DeBono (1967). This type of problem will be used extensively in this thesis and involves the requirement for representational change or restructuring to attain solution. An example is:

Anthony and Cleopatra are lying dead on the floor in an Egyptian villa.

Nearby is a broken bowl. There are no marks on their bodies and they were

not poisoned. Not a person was in the villa when they died. How did they die?

(Sloane, 1992, p. 13)

The above problem is difficult because the names Anthony and Cleopatra are commonly associated with humans. This high frequency response is unhelpful to finding the solution to this problem. Typically, the problem solver brings unwarranted assumptions to the problem that is based on their past experiences, which create unconscious biases that are unhelpful for problem solving (Ohlsson, 1984b). Thus, solvers draw unhelpful stereotypical inferences from the theme of the problem or from certain words/phrases in the problem specification, which obstructs the production of the correct, insightful solution (Ohlsson, 1984b, 1992; Weisberg & Alba, 1981). This is an example of the effect of a negative habit (William James, 1890) or what Reason (1990) termed 'strong but wrong rules' where a person tends to default to high frequency responses that are useful in most situations but can be inappropriate when applied to an unusual situation, which consequently results in an error. The solution that Anthony and Cleopatra are fish is only achieved when the unwarranted assumption that the names refer only to humans is removed (Ohlsson, 1992).

The nine-dot problem (Scheerer, 1963) is a well researched visuo-spatial insight problem that has also been used to illustrate the role of assumptions leading to an incorrect problem representation. In this problem, nine dots are arranged in a three-by-three two-dimensional square and the goal is to link all nine dots using four straight lines without lifting the pen off the paper or without retracing any lines. The solution (see Figure 1.1) is only achieved once the solver abandons the assumption

that they cannot go beyond the imagined square boundaries to link all dots in four straight lines (Weisberg & Alba, 1981).

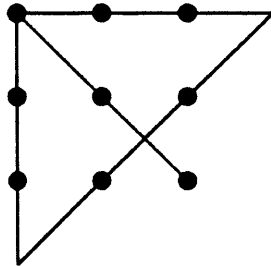


Figure 1.1. Solution to the nine-dot problem

Laboratory experiments investigating ‘fixation’ in insight problem solving have also suggested that habitual directions can prevent the solver from changing their representation. The role of ‘fixation’ has been investigated using ‘functional fixedness’ problems, which require an everyday object to be used in a novel way to attain solution (Duncker, 1945). For example, Duncker (1945) found that when participants were asked to fix a candle to a wall using a box of drawing pins, matches and a candle, participants fixated on the typical function of the box that blocked the generation of a novel function for the box to be used as a candle holder. That is, solvers tended to default to, and thus fixate on, the common function for an object, which impaired problem solving.

On the whole, insight problems are difficult to solve because inappropriate reliance on past knowledge and experiences make it particularly difficult to realise that the initial representation of the problem is incorrect. In other words, our strong thinking habits can impede solution to insight problems making the ever necessary representational change psychologically difficult.

Despite our attempt to characterise the unique nature of insight problem solving, there is still much debate concerning whether insight problem solving represents a distinct type of problem solving involving unique processes, or whether it is achieved through the same processes as those involved in solving non-insight or analytic problems (i.e., problems that are solved by working incrementally towards the solution). There are four areas in which research has attempted to identify differences between insight and non-insight problems; metacognitions (Metcalf & Wiebe, 1987), individual differences in cognitive abilities (DeYoung, Flanders, & Peterson, 2008; Gilhooly & Murphy, 2005); the role of working memory (Ash & Wiley, 2006; Fleck, 2008; Lavric, Forstmeier, & Rippon, 2000); and neural activity (Luo & Nicki, 2003; Jung-Beeman et al., 2004). Each of these will be discussed briefly below.

Metcalf and Wiebe (1987) tested whether insight and non-insight problems generate different metacognitive experiences. Participants in their study were required to give feelings of warmth ratings at 10 second intervals to indicate how near they felt they were to the solution when solving five insight and five non-insight problems. They found that participants showed awareness when approaching the correct solution to non-insight problems but were unaware of their progress towards the solution to insight problems until just before they reached the solution. These results were interpreted as suggesting that the solution to an insight problem occurs suddenly, as Gestalt psychologists argued, hence why solvers were poor at predicting how close they were to the solution. In contrast, non-insight problems require a more methodical approach to reach solution; therefore the solver is likely to have some knowledge concerning their progress. However, Weisberg (1992) questioned these results as raw data were used to report frequencies of insight problems whereas proportions were

used for the non-insight problems. Nevertheless, these results suggest, although do not prove, that different processes may be involved in insight and non-insight problem solving.

Many attempts have been made to identify whether success on insight versus non-insight problems is determined by individual differences in cognitive abilities (e.g. DeYoung et al., 2008; Gilhooly & Murphy, 2005). Gilhooly and Murphy (2005) investigated whether success on 24 insight (e.g., Nine-dot; Inverted pyramid, Ohlsson, 1992; Marriage, Dominowski, 1994) and 10 non-insight problems (e.g., anagrams, Gilhooly & Johnson, 1978; Hobbits and Orcs, Thomas, 1974; Tower of Hanoi, Egan & Greeno, 1974) was linked to measures of working memory, fluid intelligence, and fluidity and flexibility of thought. Regression analyses indicated that fluid intelligence predicted success on non-insight problems, whereas fluidity and flexibility of thought predicted success on insight problems. When insight problems were categorised into verbal and spatial problems, it was found that success on spatial insight problems was predicted by figural fluency (the ability to generate spatial interpretations, $r = .41$). However, success on verbal insight problems was related to flexibility of thought ($r = .32$) and vocabulary scores ($r = .41$). Gilhooly and Murphy (2005) suggest that the latter finding is likely due to many insight problems containing verbal ambiguities that need to be identified and resolved to attain solution. An additional finding was that success on insight problems was better predicted by measures of strategic switching and inhibition (e.g., figural fluency). Therefore, it was inferred that executive processes of switching and inhibition play a role in solving insight problems. That is, when the initial representation and approach lead to an impasse, a switch to an alternative approach takes place after inhibiting the unhelpful representation. However, a later study by Gilhooly and Fioratou (2009) found that

measures of inhibition and switching did not predict success on insight problems, thus suggesting that restructuring involves spontaneous processes.

A recent study by DeYoung et al. (2008) examined whether individual differences in three cognitive abilities were associated with performance on nine verbal insight problems (e.g., Inverted pyramid, Ohlsson, 1992; Marriage, Dominowski, 1994). The three abilities were convergent thinking (i.e., moving logically and linearly toward a solution), divergent thinking (i.e., non-linearly searching through memory and recognising a pattern), and breaking frame (i.e., abandoning a faulty initial representation and searching for a new one). Convergent thinking was assessed using the Wonderlic Personnel Test of intelligence, whereas divergent thinking was assessed using three of the Torrance Tests of Creativity (fluency, originality in responses, and flexibility i.e., the ability to switch between perspectives). Breaking frame was assessed using the anomalous card task in which participants were presented a series of normal playing cards on a screen and asked to describe what they saw. Within the trial an anomalous playing card (a black four of hearts) was presented repeatedly and the ability to break frame was indicated by the number of trials prior to the identification of the anomalous card. In other words, the higher the number of trials before identification of the anomalous card, the lower the ability to break frame. Regression analyses revealed that all measures correlated significantly with insight, except for the originality index of divergent thinking. It was suggested that the ability to break frame is important to avoid persevering with an incorrect representation, while divergent thinking is necessary to generate elements of a new representation. Further, convergent thinking is helpful in identifying shortcomings in the existing problem representation or to validate novel representations. The association found between flexibility and insight provides added

support for the suggestion made by Gilhooly and Murphy (2005) regarding the process of switching to overcome an impasse in insight problem solving. However, as only verbal insight problems were studied, generalisability of the results is limited.

Working memory is another individual difference variable that has been examined to differentiate insight and analytic forms of problem solving. Working memory is critical to cognition as it is used to store and manipulate information required for cognitive tasks such as language, comprehension, learning and reasoning (Baddeley, 1992), and is broken down into verbal and spatial storage units.

Alternatively, short-term memory is used to hold a small amount of information that is readily available for a brief period of time. Whilst some researchers have reported working memory to be associated with analytic problems (Fleck, 2008; Gilhooly & Murphy, 2005; Lavric et al., 2000), others have reported an association with insight problems (Ash & Wiley, 2006). However, recently Gilhooly and Fioratou (2009) suggested that working memory contributes to both insight and non-insight problem solving, although insight problems that require restructuring are likely to make more demands on working memory.

Lavric et al. (2000), Gilhooly and Murphy (2005), and Fleck (2008) reported similar findings in that working memory was related to analytical problems and not insight problems. Lavric et al. (2000) found that when participants performed Duncker's (1945) candle problem (insight) and a version of the Watson's Selection Task (analytical) either alone or whilst completing a working memory task that required counting the number of tones presented, the memory task disrupted performance on the analytical problem but not the insight problem. The authors concluded that because insight problems are ill-defined with an ambiguous problem situation, little planning takes place during solution attainment. In contrast, analytic

problems have a clear defined problem space, allowing more opportunities for planning, which is dependent on working memory. However, these results need to be interpreted with caution as only one insight and one analytical problem was used. Hence, a study by Gilhooly and Murphy (2005), discussed earlier, used a larger set of insight versus non-insight problems and found that performance on Raven's Matrices, a measure of working memory, predicted success on non-insight problems ($r = .41$). More recently, Fleck (2008) found similar findings to Lavric et al. and Gilhooly and Murphy using four insight and four analytical problems. In addition, Fleck found that verbal short-term memory was associated with insight problems, but this finding was less conclusive once individual problems were analysed. As no two insight problems are the same; it seems plausible that the demand on memory is likely to vary.

In contrast to the above, Ash and Wiley (2006) reported an association between working memory and (visuo-spatial) insight problem solving, although this finding was dependent on whether an insight problem had either a large or small search space to reach the solution. A configuration of six insight move problems (e.g. eight-coin, matchstick arithmetic) were manipulated to create many move available (MMA) and few move available (FMA) problems. MMA problems had a large problem space, which the solver needed to explore before reaching an impasse that triggered restructuring. In FMA problems, the initial representation allowed few or no moves, therefore the search through the problem space was limited, which in turn enabled the solver to reach an impasse and restructuring sooner. Regression analyses revealed that measures of working memory span (that demonstrate the ability to actively control and allocate attention) were correlated with success on insight problems with a large search space i.e., insight problems that involve a search phase

and a restructuring phase, whereas the ability to control attention did not predict success on problems that only involved a restructuring phase i.e., FMA problems.

An alternative perspective was provided by Gilhooly and Fioratou (2009) who argued that working memory should be involved in solving both insight and non-insight problems because both require temporary storage of possible solutions. They also noted that previous studies in this area have used a limited number of test problems (Fleck, 2008) and, or, only one type of insight problem (e.g., Ash & Wiley, 2006; Lavric et al., 2000), which limits generalisability of the findings. Gilhooly and Fioratou addressed these criticisms by using a larger set of insight (18) and non-insight (10) problems to examine the role of verbal working and visuo-spatial working memory capacities (as measured by working memory measures) in solving insight and non-insight verbal and spatial problems. It was found that verbal working memory capacity predicted success on both verbal insight and verbal non-insight problems but not to either type of spatial problems. However, spatial working memory capacity predicted success on both spatial and verbal types of insight and non-insight problems. A possible explanation provided for the latter finding was that verbal problems may also induce spatial short-term storage processes during problem solving, for example by triggering strategies that are based on imagery. It was concluded that the internal representation of an insight or non-insight problem tends to be complex; therefore it needs to be held in working memory whilst the solver considers alternative interpretations or actions. Further, the key elements of the problem that require reinterpretation are likely to be held in working memory. This is likely to be the case for insight problems as they require restructuring to overcome an impasse.

Finally, researchers have attempted to identify whether different neural activities are associated with success on insight and non-insight problems. Research suggests two important parts of the brain that are associated with insight problem solving: the hippocampus (Luo & Nicki, 2003) and the right hemisphere (Jung-Beeman et al., 2004). Luo and Nicki (2003) used functional magnetic resonance imaging (fMRI) to show that when solution to insightful Japanese riddles (that had a surprising solution) was presented to participants, stronger activation occurred in the right hippocampus. The converse was true for non-insight riddles (that did not have a surprising solution). The authors concluded that the hippocampus is involved in restructuring in insight problem solving. In particular, they proposed three possible roles for the hippocampus. First, stronger activation for insightful solutions may be due to novel representations formed among already existing concepts. Second, the hippocampus may be involved in overcoming impasse. Third, stronger activation may be due to the formation of a new representation.

A further study by Jung-Beeman et al. (2004) demonstrated that the hippocampus is not the only brain region that is involved in restructuring. Participants in their experiments completed remote association tasks in which three words were presented and the task was to find the target that could be combined with the three words to form a new word or phrase (e.g., the word 'cheese' is associated with the following three words: 'cottage', 'blue', and 'mouse'). After finding the solution, participants reported whether or not it was accompanied by an 'aha' experience i.e., whether the solution came to them suddenly. They found that solution was preceded by greater activity in the right anterior superior temporal gyrus, a region of the brain in the right hemisphere which is associated with making connections between information during comprehension. EEG activity was also greater in that region just

before the insight solution was reported. It was argued that this brain region is responsible for linking mental concepts in a novel way and may promote restructuring. However, these results were based only on verbal problems which require contrasting with non-insight problems to obtain conclusive evidence that the right anterior superior temporal gyrus is in fact associated with insight problem solving.

In summary, research that has reported differences in metacognitive experiences (Metcalf & Wiebe, 1987) during insight and non-insight problem solving is questionable due to criticisms of methodology or test problems used. However, studies of individual differences in cognitive abilities highlight that flexibility in thinking (Gilhooly & Murphy, 2005) or divergent thinking (DeYoung et al., 2008) is predictive of success on insight problems. This is consistent with the notion that solution to insight problems are more likely to be attained once the solver considers alternative interpretations that go beyond the initial faulty representation. Research also suggests that working memory is linked to insight problem solving, particularly in facilitating restructuring (Ash & Wiley, 2006; Gilhooly & Fioratou, 2009), which is a prerequisite to solving insight problems. Finally, neuroscientific studies also suggest that different brain regions are associated with solving insight and non-insight problems. In conclusion, the research discussed hitherto suggests that distinct processes are indeed involved in solving insight and non-insight problems.

1.2. Why is insight problem solving practically important?

There are many practical examples where a shift in representation is imperative but difficult to initiate because solvers repeatedly default to a habitual response, which in most occasions is the correct response. However, difficulties

experienced in shifting the default response have resulted in disastrous consequences such as the David-Besse (1985) and Three Mile Island (1979) incidents. The author has research experience in the nuclear industry; therefore examples from this area are discussed below to demonstrate the detrimental effect of habitual responses on problem solving. In particular, examples concerning incorrect assumptions regarding instrument readings and plant faults are provided. It is proposed that the study of insight problem solving has important implications for understanding everyday problem solving and even catastrophes in which a habitual mode of representation results in assumptions that inhibit the production of the correct interpretation and solution of the problem.

For example, operators are very good at diagnosing routine problems within the nuclear plant that often involve a single fault yet operators encounter difficulties in diagnosing unusual faults, particularly those that are unusual and involve multiple faults. One suggestion for this finding is that habitual assumptions (e.g., symptoms are caused by a single-fault or that instrument readings are always valid) interfere with diagnostic reasoning . This was tested in a field study by Patrick, Gregov, Halliday, Handley, and O'Reilly (1999). Operations personnel were presented with three fault-finding scenarios in a simulator of a plant control room. The scenarios consisted of one single fault (a small leak) and two multiple-faults: a) plant fault masked by instrument failure; and b) two leaks in the secondary circuit. In the first multiple-fault scenario, participants correctly identified a fault with the plant but when they checked the instrument panel, the instrument incorrectly indicated no such fault. However, despite this contradictory information, participants chose to reject their correct hypothesis and to believe the instrument reading. It can be inferred that the habitual experience that instruments readings are nearly always correct had a strong inhibitory

effect on hypothesis generation. This is because participants generated the correct hypothesis several times due to the substantial evidence indicating a plant fault but they chose to reject it every time they viewed the instrument reading that did not corroborate their hypothesis.

The David-Besse nuclear power plant incident of 1985 in Germany is another example in which operators incorrectly assumed that instrument readings were valid in a multiple-fault situation. A capacitor failure caused a loss of main feedwater but the instrument incorrectly indicated that the pressure relief valve was shut when it was still open. Consequently, an operator attempted to start the emergency feedwater pumps but this error resulted in both pumps tripping. In this case it can be inferred that the operators' action was based on the assumption that the valve reading was valid. Further, it appeared that this assumption was so strong that the operator did not attempt to validate the instrument reading by checking the pressuriser or the acoustic monitor.

Similarly, the accident at the Three Mile Island nuclear power plant in USA in 1979 was due to a combination of habitual interpretations, along with instrument failures, which resulted in a near catastrophe. Cooling water leaked as a relief valve was stuck open, which caused the core of the number two reactor to overheat. However, instrumentation incorrectly indicated that the pressure was high. Operators were trained to reduce the flow of coolant through the core when the pressure was high. This habitual response was triggered by the incorrect instrument reading, which under normal circumstances would have been the right action, but in this case was fatal. Another example is that operators incorrectly interpreted the indicator light was showing the relief valve to be closed when in actuality the light was indicating that power had been removed from the solenoid. This suggests that the operators were

unable to overcome this wrong interpretation/representation although it conflicted with the instrument reading.

In summary, the above examples illustrate that solvers' past experiences lead them to default to a high probability habitual response, which under normal circumstances is correct. However, in these examples, the habitual interpretation or representation was wrong, which solvers failed to recognise, even when faced with information that was contradictory. In such situations, the solver requires insight to be able to make the switch to an alternative representation. It is argued that these examples do not provide direct evidence that an incorrect habitual response is responsible for errors in problem solving. However, the solvers' resistance to change even when faced with contradictory information would suggest that habits have a strong inhibitory effect that make the much needed shift in representation difficult to execute. Therefore, research on insight problem is of importance because it serves to inform our understanding of how the problem solver represents problems and how the generation of novel solutions may be facilitated. Next, the importance of training is discussed in different contexts including problem solving.

1.3. Training

There has been extensive growth in training research in the last 40 years and the benefits of training has been reported in occupational, educational, cognitive, design, military, and engineering domains to name a few (Aguinis & Kraiger, 2009; Salas & Cannon-Bowers, 2001). 'Training' is defined as

'the systematic acquisition of skills, rules, concepts, or attitudes that result in improved performance in another environment.' (Goldstein & Ford, 2002, p.1)

Therefore, training is designed to promote transfer of learning such that a trainee is able to apply what was learnt during training to a different situation. To illustrate some of the benefits of training, this section begins with a brief overview of how training has benefited two different domains: organisations and aviation human factors. This is followed by a discussion of different theoretical formulations of transfer and how they dictate the design of training. It is suggested that training is only useful when trainees are aware of the relevance of their newly learnt skills in solving problems presented in a different context. As training has benefited a variety of domains, it is suggested insight problem solving may also be improved through training.

1.3.1. The benefits of training in different contexts

Aguinis and Kraiger (2009) recently reviewed the training and development literature since the year 2000 and reported many benefits of training, which corroborates the review of the training research literature conducted by Salas and Cannon-Bowers in 2001. That is, training works as it results in many positive benefits. This section briefly summarises some of the benefits training research has produced within organisations and aviation human factors.

The benefits of training reported within the organisational context include increased job performance, improvement in technical skills, increased declarative knowledge (i.e., knowledge about facts, meanings of words) and procedural knowledge (i.e., how to perform behavioural skills), as well as increased strategic knowledge (knowledge when to apply a particular skill) (cf. Aguinis & Kraiger, 2009).

Training has produced additional benefits in areas such as leadership, which are not necessarily unrelated to job performance. For example, Dvir, Eden, Avolio

and Shamir (2002) conducted a longitudinal study field study, using a sample of cadets in the Israel Defense Forces, in which group leaders received transformational leadership training. A transformational leader is one that exhibits charismatic behaviours and is able to motivate and stimulate followers as well giving followers individual consideration. Results showed that followers of leaders who received transformational leadership training were motivated, reported increased levels of internalised moral values for the organisation, and were high in empowerment.

Another important area that has benefited from training research is aviation human factors. Air crashes such as the crash of the Boeing 737 at Kegworth (Ladkin, 1996) have consistently highlighted the need for training to reduce human error. Safety and team-based training programs in this area have resulted in a reduction in time lost in dealing with injuries as well as improving team performance, respectively (Edkins, 2002). As teamwork is critical to achieving common goals in areas such as aviation, team-based training has received much attention. In particular, Crew Resource Management (CRM) training has been conducted using flight simulators to improve teamwork and communication. Studies that have addressed CRM training reported positive reactions to training by aircrews, increased knowledge of teamwork principles, aircrew communication and performance, as well as increased situation awareness (O'Connor, Flin, & Fletcher, 2002; Salas, Burke, Bowers, & Wilson, 2001).

In summary, research suggests that training is useful at an individual and team level across domains and thus further confirms the importance of training research. Research on insight problem solving have tended to focus on identifying the processes that lead to the attainment of insight with less focus on how insight problem solving

can be improved through training. The subsequent sections summarise the transfer literature and the implications for the design of training.

1.3.2. Theoretical approaches to transfer

Transfer can either be positive (when previous habits or skills facilitate learning) or negative (when previous habits or skills interfere with the learning of new skills). Further, positive or negative transfer can take place during both the learning of new skills as well as during the performance of old skills (Patrick, 1992). Transfer and training invariably go hand in hand, thus the design of training is dictated by theoretical approaches of transfer, which are briefly summarised below.

Theoretical approaches to transfer consistently highlight the importance of ensuring similarities between training and test situations for positive transfer to occur. Thorndike and Woodworth (1901) proposed that transfer was determined by the extent to which two tasks shared 'identical elements' that were of a cognitive nature. Thus, if two tasks shared common methods, principles, and procedures, then positive transfer was likely to occur.

Support for the theory of identical elements came from Anderson (1987) and his theory of skill acquisition which was concerned with how procedural knowledge (knowledge about how to do something) was developed from declarative knowledge (knowledge about facts), and how it was turned into skilled performance using IF...THEN production rules. Anderson argued that the extent of positive transfer between two tasks was determined by the extent to which the production rules in the context of one task (i.e., training task) could be used in the performance of another task (i.e., transfer task).

According to Judd's theory of deep structure, transfer depends on the extent to which the learner notices underlying shared causal principles between two problems (Judd, 1908). Gick and Holyoak (1987) suggested the transfer was determined by the perceived similarity between the old and new situation. In other words, positive transfer will occur when two situations are perceived similar or identical and the responses required are similar. In short, although the above transfer approaches ignore the psychological nature of the skills required in completing tasks (Patrick, 1992), they all highlight that training and test problems must share similarities to promote positive transfer of training.

1.3.3. Factors that inhibit positive transfer of training

Although the above theoretical approaches suggest that positive transfer will occur between two cognitive tasks that share 'identical elements' or components, this is not always the case. That is, successful performance on training tasks cannot be taken to assume that a trainee will be successful on test tasks. There are primarily two reasons for the lack of transfer between cognitive tasks, which have implications for the design of training in general as well as for training on insight problem solving. Firstly, trainees may lack awareness of the range of tasks or situations to which the newly learnt skill or procedure can be applied. As it was mentioned earlier, perceived similarity between two tasks is considered important in facilitating positive transfer. However, trainees may not perceive two tasks to be similar even when they actually are (Gick & Holyoak, 1987). Secondly, knowledge and skills are highly dependent on the context in which they were learnt, which make it difficult to facilitate positive transfer to another, new context (Patrick, 1992). The issue of context and transfer are discussed below.

There are many examples which have demonstrated that the context can have inhibitory effects on the transfer of training (e.g., Bassok & Holyoak, 1989; Gick & Holyoak, 1980; Schoenfeld, 1988). Examples from research on analogical problem solving (Gick & Holyoak, 1980) and mathematical problem solving (Bassok & Holyoak, 1989) are provided below to illustrate this point.

Analogical problem solving requires problem solvers to use similarities between the current problem (called the 'target') and a problem solved in the past (called the 'source') in attaining the solution. Gick and Holyoak (1980) employed Duncker's (1945) radiation problem, in which a patient with a malignant tumour could only be saved by a special kind of ray. However, a ray of high strength destroys both the tumour and healthy tissues whereas a weak ray does not destroy healthy tissues but it does not destroy the tumour either. The solution was to destroy the tumour by attacking with many low intensity rays from different directions simultaneously. Only 10% of participants who were given the problem on its own reached the solution. Other participants were given three stories to memorise prior to completing the problem. One of these stories was structurally similar to the radiation problem in which a general captured a fortress by dispatching small platoons along many roads to attack simultaneously. When participants were given a hint i.e., that the story was relevant to solving the radiation problem, solution rate increased to 80% in contrast to 40% when no hint was provided. Thus, participants in the no hint condition tended not to make use of the analogue. This suggests that although relevant rules or strategies may be stored in long-term memory, there is no guarantee that they will be perceived as relevant and applied when the context is changed.

Bassok and Holyoak (1989) examined transfer of procedures learnt within an algebraic context to a physics context. They found that students who had learnt

algebra could apply their knowledge to solve an isomorphic (structurally similar) physics problem. Thus, students had recognised the similarity between the problems that enabled them to apply their algebraic knowledge to solve a new problem. However, students who had learnt physics found it difficult apply their knowledge to solve an isomorphic algebra problem. Thus, learning to apply mathematics within the physics context did not prepare students to solve a more abstract mathematical problem in which the physics context was removed. This suggests that even if a problem solver has the right skills to solve a given a problem, a change in the problem context is likely to inhibit performance due to the reliance on past problem solving experiences.

In summary, theoretical formulations of transfer suggest that training and test tasks that are of a cognitive nature must share similarities in terms of structure, procedures, principals or method to facilitate positive transfer. However, trainees must be aware of the relevance of their new skills to a problem presented in a new context for positive transfer to take place. In other words, trainees must be able perceive the relevance of the training to solve test problems. Given that training has resulted in many benefits across a variety of domains, training designed to improve insight problem solving is needed, which is what this thesis aims to achieve whilst addressing the above issues.

1.4. Overview of thesis layout

The primary aim of this thesis is to make a contribution to the field of insight problem solving. This is done by examining theories of insight problem solving advocating the importance of representational change that is relevant to how training may be designed to improve success on verbal insight or lateral thinking problems.

Also there is a lack of qualitative analysis of insight problem solving, which this thesis attempts to correct. Hence, the experiments reported here utilised concurrent verbalisations to gain an in-depth understanding of the process of insight problem solving, as well as to validate the effects of training.

The thesis is organised into nine chapters. The present chapter provided an introduction to insight problem solving, how it is different and why research in this area is of importance. It also discussed how training may serve to improve insight problem solving, which is the focus of this thesis. *Chapter 2* provides a review of the different theoretical accounts of insight problem solving whereas *Chapter 3* reviews the different methods for facilitating insight.

Chapters 4 to 8 form the empirical chapters of this thesis. Specifically, *Chapters 4 to 7* report experiments that attempted to overcome habitual conceptualisations by introducing different training approaches to increase the probability of attaining the insight solution. The approach to training can either be specific or generic to test problems. That is, specific training concerns facilitating positive transfer to a particular problem for which training was designed (e.g., Burnham & Davis, 1969; Chronicle et al., 2001; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981), whereas generic training concerns facilitating transfer to not only the particular problem for which the training was designed but also to novel insight problems (e.g., Ansburg & Dominowski, 2000; Chrysikou, 2006; Dow & Mayer, 2004; Wicker, Weinstein, Yelich, & Brooks, 1978). Therefore, *Chapters 4 and 5* focus on specific training to promote transfer to verbal insight problems as much of the research has been restricted to the consideration of visuo-spatial problems (Burnham & Davis, 1969; Chronicle et al., 2001; Kershaw & Ohlsson, 2004;

Weisberg & Alba, 1981). *Chapters 6* and *7* focus on generic training to promote transfer to a wider range of verbal insight problems.

Chapter 8 forms the last empirical chapter in this thesis. It reports an experiment that tested a novel method to determine whether problem solving is impaired as result of an incorrect representation, other than by asking someone to solve a problem. The aim of the experiment was to determine how powerful an incorrect representation is, and consequently, to determine the degree of difficulty on a given verbal insight problem.

Chapter 9, which is the final chapter, summarises the findings, implications and limitations of the thesis and makes suggestions for future research.

CHAPTER 2

Review of problem solving theories relevant to achieving insight

This chapter provides a review of problem solving theories that may serve to increase our understanding of insight. It begins with a review of Newell and Simon's (1972) Problem Space Theory and how its applicability is limited to well-defined problems. This is followed by a critical evaluation of insight problem solving theories including Gestalt Theory (Köhler, 1924; Wertheimer, 1945), Representational Change Theory (Ohlsson, 1992), Restructuring Change Theory (Knoblich et al., 1999, 2001), and Progress Monitoring Theory (MacGregor et al., 2001). It is concluded that representational change is critical to solving insight problems.

2.1. Problem Space Theory (Newell & Simon, 1972)

Newell and Simon's (1972) Problem Space theory has been instrumental in understanding general problem solving. Although it is not a theory of insight problem solving per se, it is of interest to evaluate this theory in the light of its relevance to insight problem solving. An outline of this theory is provided below followed by a discussion of the type of problem in scope to this theory.

According to Problem Space Theory, problem solving centers around the problem space, which consists of knowledge states i.e., the initial state of the problem, the goal state, all possible mental operators or moves that can be applied to a state to change it to a different state, and all of the intermediate states of the problem. The solver creates an internal representation of a problem, referred to as the initial state, and problem solving entails applying operators to reduce the distance between the initial state and goal state, and a number of intermediate states are created along the

solution path. Newell and Simon also proposed that problem solving is constrained by working memory limits, and the speed at which information is retrieved from long term memory. Hence, heuristics or rules of thumbs are used to reduce the number of alternative solutions possible during problem solving although the use of a heuristic may not always yield results.

Problem Space Theory has been used to successfully explain problem solving in well-defined problems such as the Tower of Hanoi (Anzai & Simon, 1979) and the missionaries-cannibals (Thomas, 1974) problem. In the former problem, subjects are presented with three pegs where the first peg consists of three discs of decreasing size i.e., the initial state, and the solution or goal state is to pile the discs in the same order in the last peg. The rules are that only one disc can be moved at a time and a larger disc cannot be placed on top of a smaller one. In the missionaries-cannibals problem, the task is to transfer three missionaries and three cannibals across a river on a boat. The boat is small, therefore up to two people can be taken across at any one time and the boat must be accompanied by someone when returning to the other side. However, at no point can there be more cannibals than missionaries on one side of the river or the cannibals would eat the missionaries. The rules in both problems therefore restrict the possible number of operators that can be applied at each point of problem solving.

There are three main differences between these problems and insight problems. Firstly, the Problem Space Theory conceptualises the Tower of Hanoi in terms of clear initial and goal states, which are used as a basis for planning a series of problem solving moves. Information about the states of a problem can be used by the theory to predict the size of the problem space. In contrast, for example, the verbal insight Anthony and Cleopatra problem cannot be represented in terms of well defined initial and goal states. Therefore, the size of the problem space for this insight

problem is dependent on the problem solver's representation, which is influenced by their past habits and knowledge. Secondly, the Tower of Hanoi has clear rules, which minimise the potential operators that can be applied, and thus guide problem solving. In contrast, the Anthony and Cleopatra problem is not bounded by any such rules. Finally, the Tower of Hanoi has clearly defined operators or moves that can be applied to reduce the distance between problem states as the solver progresses through the problem space until the goal state is reached. General heuristics can be used to reduce the possible number of alternative moves, thus directing the solver towards the goal state. However, the more ambiguous nature of the Anthony and Cleopatra problem makes it difficult for the solver to define a problem space that can be broken down into states that can be acted upon. There are no clear operators that can be used, particularly as success on an insight problem is dependent on the solver recognising and correcting an inappropriate problem representation.

In conclusion, the differences between insight problems and the problems in scope to Problem Space Theory suggest that this theory is more applicable to understanding well-defined move problems that can be explicitly represented in terms of knowledge states. More importantly, the initial representation of a well-defined problem is defined by the initial state of a problem at the outset. On the other hand, insight problems are ill-defined with no clear states, therefore what the problem solver brings in terms of their own experiences and knowledge interacts with the presentation of the problem, and together these two factors determine the initial representation which in turn influences subsequent problem solving. Despite the lack of applicability to insight problems, Problem Space Theory has paved the way for insight theories such as MacGregor et al.'s (2001) Progress Monitoring Theory, which is discussed in *Section 2.5* in this chapter.

2.2. Gestalt Theory (e.g., Köhler, 1924; Wertheimer, 1945)

Gestalt Theory was originally applied to understanding perception, with particular emphasis on searching for patterns and interpretations within data or elements. Gestalt psychologists (Duncker, 1945; Köhler, 1969; Wertheimer, 1945) later applied the theory to understanding human problem solving. This section will outline key distinctions made by Gestalt Theory, which was directly concerned with insight and how it might be facilitated. It will be argued that despite shortcomings, Gestalt Theory has been influential in stimulating later theories of insight problem solving that focus on the importance of restructuring or representational change.

Gestalt Theory is based on the idea of how people organise or 'group' elements of a visual stimulus in order to view it as whole (Wertheimer, 1922). The basic principle followed by Gestalt psychologists is that the whole is greater than that sum of its parts. That is, viewing something as a whole has greater meaning than viewing it by individual parts. The idea of 'grouping' was used by Gestalt Theory to make two important distinctions in understanding problem solving. Firstly, the role of reorganisation or restructuring in problem solving was emphasised (e.g., Maier, 1931). A solver must reorganise how they perceive a problem or parts of a problem in order to increase the probability of solution attainment. Secondly, a distinction was made between two types of thinking: reproductive thinking and productive thinking (Wertheimer, 1923). Reproductive thinking entails using previous experiences or solutions to solve a new problem. However, productive thinking is considered more complex as it requires the solver to produce a new organisation (representation) of the problem elements to attain solution. That is, a novel restructuring of the problem is required that enables the solver to perceive the problem in a way that was not considered initially in order to overcome an impasse during insight problem solving.

The Gestalt view contends that reproductive thinking may impede problem solving as the initial interpretation/representation of a problem may result in the solver fixating on aspects of a problem (Duncker, 1945; Luchins, 1942), thus preventing novel interpretations that might lead to a solution. As a consequence, productive thinking is required to view the problem in a novel way before progress can be made.

What also distinguishes Gestalt Theory, from other theories of insight problem solving, is the argument that insight results from the sudden reorganisation of one's understanding of the problem, without the awareness of the solver, and that it is accompanied by the so-called 'aha' experience. This 'aha' experience is what Wallas (1926) labelled as illumination (when a creative idea suddenly emerges in one's head), which was proposed to be the third phase in a series of four phases of creative problem solving.

In short, Gestalt Theory characterises insight problem solving by three criteria: firstly, a problem is solved suddenly via a 'aha' experience; secondly, it is solved after a period of impasse or no progress; and thirdly, a problem must be approached in a new way that enables restructuring of the problem (Weisberg, 2006). Evidence for the 'aha' experience largely comes from the observation of the apparent suddenness of the emergence of the solution after a period of no progress (e.g., Jung-Beeman et al., 2004; Köhler, 1925; Metcalfe & Wiebe, 1987). For example, Köhler (1925) observed that apes encountered an impasse during problem solving (e.g., attempting to reach an unreachable banana from a cage), which was followed by an apparent sudden restructuring of the problem, and hence achieving the solution (e.g., reaching the banana by joining two sticks together). Further support in human problem solving was reported by, for example, Maier (1931) using the two-string problem. In this problem participants were taken into a room containing several objects (e.g., poles, pliers) and

two strings that were hanging from the ceiling. The task was to tie the two strings together yet the strings were far apart such that the participant was unable to reach one string whilst holding onto another. The insightful solution was to attach pliers to one of the strings and to swing it like a pendulum, thus allowing one string to be held whilst catching the other on its upswing. Participants found this problem difficult until the experimenter appeared to accidentally brush against the string to produce a swinging motion. Maier concluded that the participants initially reached an impasse after which the hint triggered reorganisation or restructuring of the problem to produce the correct solution. Furthermore, most participants who reached the solution after observing the hint reported that they had not noticed the experimenter brush against the rope.

Support for the 'aha' experience also comes from the previously mentioned study by Metcalfe and Wiebe (1987) who found that participants' warmth ratings (i.e., ratings of how close they thought they were to the solution) increased as they approached the solution to search (algebra) problems. In contrast, warmth ratings for anagrams and insight problems were stable prior to solution, at which point they increased sharply. It was concluded that this provided support for the 'aha' experience in insight problem solving as solution to the insight problem came suddenly to the solver, which is argued to be the result of restructuring the problem. However, although the participants did not perceive their solution to be gradual for insight problems, it does not necessarily imply that it was not. Moreover, a general criticism of this study and indeed others is that there is no direct evidence that restructuring had in fact occurred. Such evidence is problematic and it will be argued later in this thesis that some indirect evidence can be gleaned from the collection and analysis of concurrent verbal protocols.

A later study by Weisberg and Alba (1981) used insight problems such as the nine-dot problem (Scheerer, 1963) to directly challenge the Gestalt view that solution to insight problems arises suddenly as a result of restructuring. Gestalt psychologists contend that the nine-dot problem is difficult because people 'fixate' on the square formed by the dots and therefore incorrectly assume that they cannot go beyond this square. It follows that telling participants that they can go beyond the boundaries of the square should facilitate restructuring, and therefore the insight solution. Weisberg and Alba tested this hypothesis but found no training effect unless this instruction was accompanied by a hint that gave away part of the solution e.g., giving the first line of the solution (Experiment 1). It was argued that the instruction to remove the alleged fixation, in the form of an unwarranted assumption, was not sufficient in facilitating insight. Further, it was observed that the solutions were not sudden and direct but were preceded by a number of failed attempts, thus contradicting the suddenness of solution proposed by Gestalt Theory. A similar argument was made by Novic and Sherman (2003) who suggested that success on the nine-dot problem was likely to occur after several (failed) attempts at drawing lines that connected the dots. Weisberg and Alba concluded that the Gestalt terms of 'fixation' and 'insight' were unhelpful in explaining the processes involved in solving insight problems. Instead, emphasis was placed on the role of problem-specific knowledge (e.g., solution related hints) in triggering restructuring to facilitate solution success.

Durso, Bea and Dayton (1994) used the following verbal insight problem to investigate the suddenness of solution following restructuring or representational change:

A man walks into a bar and ask for a glass of water. The bartender points a shotgun at the man. The man say 'Thank you' and walks out. What was going on? (Durso et al., 1994, p. 95)

The solution to this problem is that the man had the hiccups, therefore the bartender startled him with the gun which cured the hiccup. This solution is argued to occur suddenly once solvers make sense of the problem (Dayton, Durso, & Shepard, 1990). In Experiment 1, after completing the problem, participants rated the relatedness of all possible pair-wise combinations of 14 terms that were associated with the problem. Some terms were mentioned explicitly in the problem (e.g., bartender - man), some were objects found in a bar (e.g., TV - pretzel), and some were related to the solution (e.g., surprise – remedy). It was found that non-solvers rated words mentioned in the problem highly whereas solvers rated words that were related to the solution more highly. These results were interpreted to indicate that solvers had a different mental representation to non-solvers. However, as ratings were taken retrospectively, these results do not provide conclusive evidence for representational change. In particular, it was not known whether differences in the ratings were due to differences in the solution process or due to differences in subjects' initial representation (Ash, Cushen, & Wiley, 2009)

Subsequently, participants in Durso et al.'s Experiment 2 were asked to judge critical pairs of words, as indicated by Experiment 1, on five occasions: before and after they heard the problem, every 10 minutes during problem solving until the problem was solved, and immediately after the solution was reached. Participants judged two insight pairs (e.g., surprise – remedy, relieved – thank you), two related pairs (e.g., bartender – bar, shotgun – loaded), and two unrelated pairs (e.g., pretzel –

shotgun, TV – remedy). It was not clear why ‘remedy’ was used in one of the insight and unrelated pairs. Six filler pairs were used to disguise the aim of the experiment. It was found that relatedness ratings for ‘insight pairs’ changed as the solver progressed to the problem solution, which was interpreted to suggest that the solver’s representation of the problem changed during problem solving. Thus, ratings on the first two occasions were dissimilar, then moderately similar on the next two occasions, and then highly similar after solution. No change in ratings was observed for the related and unrelated pairs. In other words, participants rated related pairs as related and unrelated pairs as unrelated to the problem on all five occasions. These results were interpreted to suggest that even successful solvers begin with an incorrect representation, but more importantly, restructuring did not occur suddenly, as proposed by Gestalt psychologists.

There is a notable similarity between the studies of Durso et al. (1994) and Metcalf and Wiebe (1987) as both collected judgment ratings during problem solving in an attempt to demonstrate representational change. However, the results were contradictory in that the former reported evidence against the Gestalt view of the suddenness of solution whereas the latter was in support of this view. Ash et al. (2009) noted that the research by Durso et al. was confounded by design issues which question the conclusiveness of their results. Firstly, they did not compare the results with a problem that did not require restructuring. Secondly, participants in their study were allowed to ask the experimenter questions that required either a yes or no response. Thus, it was possible that the participants had gained additional information that was not available in the problem statement, which may reflect the gradual changes in similarity ratings rather than a change in representation per se. Finally, it was not known whether the pairs of critical concepts served as an indirect hint to

those who solved the problem. Although Durso et al. used verbal ratings that could be subject to biases or indeed have no relationship to on-going representations, their study however made a promising attempt to demonstrate representational change during insight problem solving.

In summary, Gestalt psychologists argued that problem solving was more complex than simply reproducing learned responses. Particular emphasis was placed on the importance of understanding the structure (i.e., the relationship among parts) of the problem (Wertheimer, 1922) in a novel way to overcome an impasse. Unlike Gestalt Theory, the Problem Space Theory (Newell & Simon, 1972) focuses on the step by step process of searching for a solution path that connects the initial state to the goal state in well-defined problems. The most distinguishing feature of Gestalt Theory is the conceptualisation of insight as the sudden appearance into consciousness which is accompanied by a 'aha' experience following the reorganisation or restructuring of the problem. Indeed, studies do appear to show support for the suddenness of solution which is consistent with achieving insight (e.g., Jung-Beeman et al., 2004; Köhler, 1925; Metcalfe & Wiebe, 1987), although this is extremely difficult to validate. However, studies that challenge this view have demonstrated that the insightful solution is often preceded by several unsuccessful solution attempts during problem solving (Durso et al., 1994; Weisberg & Alba, 1981). In other words, it is argued that the problem solver accumulates partial information at each solution attempt which in turn facilitates insight. Indeed, Ohlsson (1992) noted that even Gestalt psychologists acknowledged that the full (correct) solution to an insight problem may not occur at once. Further, research in support of the 'aha' experience has been criticised for lack of rigour and representativeness of problem solving as a limited number of problems were used to test the theory

(Weisberg, 1986). The Gestalt Theory has been described as a general theory (Ohlsson, 1984a) and has been criticised for its vagueness, particularly for not answering what insight is and what processes underlie restructuring (Davidson, 2003). The subsequent sections will review representational change theories of insight problem solving and the processes proposed by such theories to overcome an impasse (Knoblich et al., 1999; Ohlsson, 1984b, 1992).

2.3. Representational Change Theory (Ohlsson, 1992)

Ohlsson's (1992) Representational Change Theory attempted to incorporate the constructs of 'insight' and 'restructuring' proposed by Gestalt Theory into an information processing theory of problem solving. Insight was conceptualised in terms of breaking out of an impasse. Hence, if there is no impasse in a problem, then that problem does not require insight. As in Gestalt Theory, Ohlsson also argued that the interpretation of a problem situation triggers a faulty initial representation that leads to an impasse. Further, an impasse is overcome or broken when the representation is changed i.e., it is reinterpreted, re-represented or is restructured (Ohlsson, 1984b). In this section Ohlsson's explanation of the nature of insight is discussed first. Comparisons are made between Ohlsson's approach and Gestalt Theory to explain why a solver experiences an impasse, how impasses are broken, and what happens after an impasse is broken. In particular, Ohlsson's three mechanisms (elaboration, re-encoding, and constraint relaxation) by which an impasse may be broken to bring about representational change are discussed. It will be concluded that Ohlsson's theory contradicts the suddenness of solution, as proposed by Gestalt Theory.

According to Ohlsson, representational change can either result in partial or full insight. Partial insight entails breaking out of an impasse to enable (stepwise) problem solving to resume. For example, in the Anthony and Cleopatra problem solvers may correctly consider the names to not be associated to humans yet fail to realise that they refer to fish. In this case, partial insight is achieved as a result of restructuring although stepwise problem solving would not be feasible in such an ill-defined problem. Ohlsson contends that a series of partial insights may be needed before reaching the insightful solution. Contrary to the Gestalt view, Ohlsson suggests that restructuring can result in further impasses and that restructuring may not always lead to the solution. However, full insight entails breaking out of an impasse and 'seeing' the complete solution. That is, once the solver sees the solution to a problem, that solution can be attained swiftly without further errors. Based on the distinction between partial and full-insight, Ohlsson claimed that what Gestalt psychologists referred to as the suddenness of solution that is accompanied by a 'aha' experience is essentially the achievement of full insight by a solver that may be the result of a series of partial insights.

Ohlsson (1992) argued that a problem solver encounters an impasse because their initial encoding of a problem (i.e., the initial representation) is based on their past knowledge and experiences, which do not activate the relevant operators to solve the problem. A similar explanation was provided by the Gestalt approach. However, Ohlsson claimed that a solver has the knowledge operators to solve a problem, which are retrieved from memory via an unconscious automatic process called 'spreading activation'. To activate these operators, the solver must change their representation of the problem such that a new representation can serve as a new source of activation (Ohlsson, 1984b). Whereas Gestalt Theory did not specify what processes may

underlie representational change, Ohlsson suggested three mechanisms. These were: 1) elaboration or addition of new information; 2) re-encoding, which entails focusing on correcting the faulty representation of a problem that is mistaken rather than incomplete by re-categorising or deleting some information; and 3) constraint relaxation, in which incorrect assumptions or constraints on the goal are altered. Each mechanism is examined below.

Elaboration entails adding new information, without deleting or revising existing information, by closely examining and noticing features of a problem that were previously unnoticed. Alternatively, Ohlsson (1992) notes that new information can be added to a representation by recalling information from memory e.g., remembering that a metal lid on a jar will expand if placed in hot water. Ohlsson gave the example of Kaplan and Simon's (1990) mutilated checkerboard problem to illustrate how the addition of information to a representation occurs via elaboration. In this problem, the board is initially covered by 32 dominoes which occupied two squares each. Two squares were removed from two opposite corners of the checkerboard and participants were asked to determine whether the remaining 62 squares could be filled with 31 dominoes. Ohlsson argued that it is unlikely that solvers initially encode the colours of the squares as they do not recognise its relevance to the solution that 31 dominoes cannot cover the board. To attain the solution, the solver must recognise two features of the problem: that all squares on the diagonal line are the same colour and that each domino must cover one square of one colour and another square of another colour. Ohlsson notes that when these two features were noticed by the solvers, the solution was attained quickly. This observation was interpreted as evidence for elaboration of the problem that enabled

the solver to acquire a new representation of the checkerboard in terms of it losing two white or black squares.

Another mechanism for representational change is re-encoding. Ohlsson (1992) contended that elaboration focuses on an initial representation of a problem that is incomplete whereas re-encoding focuses on an incorrect or mistaken representation. That is, the solver must abandon or reject some component of their current representation via re-encoding. Ohlsson gave the examples of Duncker's (1945) candle problem and Maier's (1931) two string problem to illustrate re-encoding. In the former problem, a box must be re-encoded as a platform rather than a container to attain the solution. In the latter problem, the pliers must be used as a pendulum; therefore the solver must re-encode its function as a weight rather than a tool.

The final mechanism, constraint relaxation, focuses on the implicit constraints imposed by the solver that must be relaxed in order to overcome an impasse. Thus, constraint relaxation is argued to be different from elaboration and re-encoding because it changes the mental representation of the problem goal (solution) rather than the problem situation. For example, Ohlsson suggested that the nine-dot problem is less difficult once the solver relaxes the constraint that they can draw a line beyond the imaginary square formed by the nine dots. However, research by Weisberg and Alba (1981), discussed earlier, found that increasing a solver's awareness of this constraint did not facilitate insight unless a solution-specific hint was also given. More recently, MacGregor et al. (2001) also found that instruction to remove this self-imposed constraint did not always produce solution on the nine-dot problem. In fact, only 31% solved the problem. Indeed, Ohlsson acknowledged that the validity of the constraint relaxation analysis for the nine-dot problem is open to debate.

Although Ohlsson (1992) provided examples of insight problems to illustrate how each mechanism may facilitate representational change, it is however unclear to what extent these mechanisms are distinct. Further, if several impasses are encountered when attempting to solve a problem, then it follows that each impasse may be overcome by a different mechanism. For example, instead of elaborating on the colour of the squares on the mutilated checkerboard problem (Kaplan & Simon, 1990), it is possible that the board is re-encoded such that pairs of squares are the same size as the dominoes being used to cover it. Alternatively, problem solvers may be constrained by assuming that the problem is not solvable and thus show fewer solution attempts. In the case of the nine-dot problem, elaboration may also help focus the problem solver to think beyond the boundaries of the nine dots. Fleck and Weisberg (2004) suggest that the candle problem may not be solved using re-encoding, as proposed by Ohlsson. Instead, solution may be reached by elaborating the problem information and realising that a platform is required. This solution actually arose from several unsuccessful solution attempts in Fleck and Weisberg's study.

Research has largely focussed on investigating Ohlsson's (1992) constraint relaxation mechanism and support has been reported using a variety of insight problems including the six-coin (Chronicle, MacGregor, & Ormerod, 2004), matchstick arithmetic (Knoblich et al., 1991, 2001), the unlimited move car park game (Jones, 2003), and verbal insight problems (Ansburg & Dominowski, 2000). Indeed, Ohlsson noted that most insight problems incur impasses because the solver imposes constraints that make a problem unsolvable. However, elaboration and re-encoding have received little attention. A recent study by Ansburg and Dominowski (2000) suggests that these two mechanisms may not be distinct. They found that

training geared towards elaboration problems i.e., problems requiring a change in representation through elaboration of information facilitated performance on both elaboration and constraint relaxation test problems. These results suggest that the same underlying processes may be involved in solving both elaboration and constraint relaxation type problems.

Representational Change Theory and Gestalt Theory also differ in terms of their explanation concerning what happens after an impasse is broken or after the 'aha' experience. Firstly, in contrast to Gestalt Theory, Ohlsson argued that representational change does not always lead to a better representation. Further, it does not guarantee solution attainment. Ohlsson argued that representational change merely facilitates an impasse to be broken to enable the solver to resume problem solving. Secondly, Gestalt psychologists contend that the insight solution appears suddenly into consciousness as a result of reorganising or restructuring the problem representation. Ohlsson hypothesised that this suddenness of solution is in fact an illusion due to the solver's lack of introspective access to the cognitive processes that lead to the solution. Therefore, a solution to a problem is constructed at the moment of insight but due the capacity limits of human cognition, this construction is simple and of a short duration, which is why it is unnoticed by the solver. This hypothesis appears to be consistent with studies that have demonstrated solution attainment preceding several failed solution attempts (Durso et al., 1994; Weisberg & Alba, 1981).

In summary, Representational Change Theory provides a more detailed account of the Gestalt terms 'insight' and 'restructuring'. Insight is explained in relation to impasses experienced during problem solving. Further, the theory specified three possible cognitive mechanisms (elaboration, re-encoding, and constraint relaxation) that may facilitate representational change although it was acknowledged

that other mechanisms may exist. However, there is a lack of research that has attempted to investigate the elaboration and re-encoding mechanisms, although constraint relaxation has received greater support, thus suggesting that it plays an important role in problem solving. A criticism of the theory is that it is not always possible to predict when and how the representation of a problem will change during problem solving. The key difference between the Gestalt approach and Ohlsson's theory is that the former contends that insight is attained following a sudden reorganisation of one's understanding of a problem whereas Ohlsson rejected this proposition. Ohlsson suggested that solvers have a lack of awareness of the cognitive processes that lead up to the solution, hence why the solution may appear to be sudden. Further, it was suggested that the insight solution may be reached after a series of partial insights that involve much information processing, which in turn may lead the solver closer to the solution. In other words, partial information is accumulated during problem solving which contributes to solution attainment.

2.4. Restructuring Change Theory (Knoblich et al., 1999)

Knoblich et al.'s (1999) Restructuring Change Theory is similar to Ohlsson's (1992) Representational Change Theory in two ways. Firstly, both theories agree that insight problems trigger a faulty representation that leads the solver into an impasse. Secondly, both theories emphasise that impasses are resolved by altering the faulty representation via particular cognitive processes. Ohlsson (1992) proposed three processes underlying representational change (elaboration, re-encoding, and constraint relaxation), whereas Restructuring Change Theory advocates two processes, namely, constraint relaxation and chunk decomposition, where the former process is the same as that proposed by Ohlsson. Research by Knoblich and colleagues that tested the

Restructuring Change Theory are discussed below. It is concluded that chunk decomposition is in scope to a particular type of insight problem whereas support for constraint relaxation extends to a variety of insight problems.

Knoblich et al. (1999) argued that chunk decomposition corresponds to an element in a problem representation that can be broken down further. For example, in matchstick arithmetic, the Roman Numerals VII can be decomposed into the chunks V and II whereas tight chunks such as X are less likely to be decomposed as two slanted sticks because they have no meaning within the problem context and thus are not chunks themselves. The second process, constraint relaxation, is not proposed to be either deliberate or voluntary. Knoblich et al. suggest that by relaxing the constraints in a problem the set of options available to a problem solver is expanded. Similarly, Ohlsson (1992) in his Representational Change Theory suggested that the relaxation of self-imposed constraints allows the activation of dormant operators. However, Knoblich et al. argued that the ease at which constraint relaxation occurs is dependent upon the scope of the constraint, which is determined by how much a problem representation is affected by relaxing that constraint. It was proposed that the wider the scope of a constraint, the lower the likelihood that the constraint will be relaxed.

Knoblich et al. (1999) provided support for their theory using matchstick arithmetic problems that were presented using Roman numerals. The goal was to move a single stick so that an initial false equation is transformed into a true equation. For example, moving from the false equation $VI = VII + I$ to the true one $VII = VI + I$. Some problems (Type A) required changing the values (numbers) in the equation whereas others problems (Type B) required a change in representation of the equation by altering operators (i.e., +, -, and = signs). Knoblich et al. argued that our

experience of arithmetic is more towards changing values in an equation rather than changing operators, therefore this knowledge can inhibit problem solving. It was found that participants encountered difficulties in solving Type B problems, which was interpreted to suggest that solvers were unable to relax the constraints of arithmetic.

Knoblich et al. (2001) argued that the above experiment did not provide direct evidence for what was causing difficulties with Type B problems. Hence, they monitored eye movements whilst participants solved the matchstick problems used by Knoblich et al. (1999). They found that participants fixated on values rather than operators when attempting to solve both types of problems, which was interpreted to suggest that participants' initial representation of the problems was based on the incorrect assumption that the values required change. Participants showed increased fixation for problems that required representational change, which was interpreted to suggest that an impasse was encountered. Those who were successful at solving Type B problems showed longer fixation on operators, presumably because they realised that a change in an operator was needed to reach the solution. Further, successful participants showed an ability to shift their attention to a part of a problem that they had previously neglected, which was argued to be the result of representational change. In conclusion, these results also suggest that the representation of a problem has a strong influence on problem solving.

In general, matchstick arithmetic problems are particularly suitable in illustrating support for the two mechanisms of representational change proposed by Knoblich et al. i.e., constraint relaxation and chunk decomposition. Constraint relaxation is indeed one process by which representational change takes place in a variety of insight problems including verbal insight problems (Ansburg &

Dominowski, 2000; Wicker et al., 1978), yet the applicability of chunk decomposition does not extend as well to verbal problems as it is the semantic meaning of words in a problem that influences the mental representation. Therefore, decomposing the words to individual letters in such problems will not be useful.

In summary, Restructuring Change Theory is consistent with Gestalt Theory and Representational Change Theory in that all three approaches contend that insight is achieved once restructuring is facilitated by overcoming an impasse. However, Gestalt Theory did not specify any processes that may facilitate restructuring, whereas Representational Change Theory specified three mechanisms and Restructuring Change Theory specified two processes. In particular, Restructuring Change Theory drew our attention to an additional cognitive process i.e., chunk decomposition, which is more suited to explaining how insight problems such as matchstick arithmetic problems may be broken down into individual chunks to aid representational change.

2.5. Progress Monitoring Theory (MacGregor et al., 2001)

Progress Monitoring Theory provides an alternative view to the representational change accounts of insight problem solving by suggesting that an impasse results from the inappropriate application of heuristics rather than an inappropriate mental representation of the problem (MacGregor et al., 2001). This section begins with an overview of Progress Monitoring Theory which is followed by an evaluation of research that tested this theory (MacGregor et al., 2001; Ormerod et al., 2002). Next, research that attempted to test the predictions of both Progress Monitoring Theory and Representational Change theories are evaluated (Jones, 2003; Öllinger et al., 2006). It is argued that recent research suggests that an impasse results from an inappropriate problem representation rather than the inappropriate application

of heuristics. Moreover, representational change determines what heuristics are applied during insight problem solving. It will be concluded that the generality of this theory is limited to visuo-spatial insight problems.

Progress Monitoring Theory has some similarity to Newell and Simon's (1972) Problem Space Theory, as both emphasise the role of heuristics in problem solving, although the latter is in scope to well-defined problems. This theory is based around the hill-climbing idea that problem solving proceeds with the solver seeking to minimise the distance between the current state and the goal or subgoal state. An impasse occurs when hill-climbing does not result in a solution, therefore alternative searches are explored to overcome this impasse. Two heuristics are central to Progress Monitoring Theory: maximisation and progress monitoring. The maximisation heuristic is similar to means-end analysis proposed by Problem Space Theory in that the solver works towards attaining the goal with each move. Progress monitoring is used by the solver to assess their rate of progress towards the goal. When the rate of progress is considered too slow to solve a problem within the maximum number of permissible moves, criterion failure occurs, which leads the solver to consider alternative searches that may lead to insight. However, solution is likely to be achieved earlier the sooner criterion failure is experienced. In contrast to Representational Change Theory, Progress Monitoring Theory contends that constraint relaxation is often necessary but not sufficient for achieving insight (MacGregor et al., 2001). That is, constraint relaxation will facilitate problem solving for those who experience criterion failure. The probability of experiencing an impasse in a problem is dependent on a person's lookahead value i.e., the capacity of potential moves an individual can lookahead and remember. It was proposed that those with

high lookahead will achieve insight quicker as they will realise sooner that the problem cannot be solved with traditional methods (MacGregor et al., 2001).

MacGregor et al. (2001) demonstrated support for their theory using the nine-dot problem, which is a visuo-spatial insight problem. Participants were given this problem but with a line already drawn. It was found that a line that extended horizontally across the top three dots and to a nondot point was less helpful in comparison to a diagonal line that went from the top left dot to the bottom right dot. According to Representational Change Theory, the former should have been beneficial as it removes the constraint that the lines are bound inside the square. However, those given the diagonal line demonstrated better performance, which MacGregor et al. argued forced participants to consider alternative solutions more quickly because drawing two additional lines whilst keeping inside the square would have resulted in three more dots being crossed out (Jones, 2003). However, this result is not surprising as earlier research has demonstrated that solution-specific hints facilitate performance on the nine-dot problem (e.g., Weisberg & Alba, 1981).

A later study by Ormerod et al. (2002) found similar results as MacGregor et al. (2001) using the eight-coin problem, which is also a visuo-spatial insight problem. Participants were presented with a configuration of eight coins and instructed to remove two coins, within a specified number of moves, so that each of the eight coins touched only three coins. The solution required the solver to represent the problem in three-dimensional terms such that the coins can be stacked on top of each other. Participants were given different versions of the problem that manipulated the number of moves available in which a coin can touch three other coins. It was found that in versions where there were no moves available, the problem was solved quicker than in versions where several moves were available. Ormerod et al. argued that, as in the

case of the nine-dot experiment (MacGregor et al., 2001), the number of moves available determined how quickly alternative solutions were sought. These results support Ash and Wiley (2006) who suggested that insight problems involving many moves require the solver to search a larger problem space, whereas problems with a few or no moves require a shorter search, therefore the solver reaches an impasse and restructuring sooner.

Jones (2003) suggested that the problems in scope to Progress Monitoring Theory and Representational Change Theory differ in that the former cannot be applied to problems that can be solved in one move because it relies on the constant monitoring of progress during problem solving. Jones tested both theories using the car-park game, which is an unlimited move (visuo-spatial) insight problem. The game entails manoeuvring a taxi out of a car park but the way to the exit is obstructed by other cars. The goal is to determine how the cars can be moved out of the way to enable the taxi to leave the car park. This problem is considered difficult because solvers fail to realise that the taxi must be moved before an exit route can be created. In other words, the lack of awareness of this critical move places a constraint on the problem solution. Participants were given practice problems although none required them to move the taxi to create a path to the exit. The knowledge gained from the practice was manipulated such that it could be applied to the car-park problem. Representational Change Theory predicts that practice problems should lead participants to assume that the taxi does not need to be moved, which should hinder performance on the car-park problem. In contrast, Progress Monitoring theory predicts that the practice problems should have no effect as performance is based on reducing the difference between the initial state and goal state. The results provided greater support for Representational Change Theory as performance was higher when

easy practice problems or a rotated version of the problem was given. Despite the different emphasis and scope of application of both theories, both advocate that insight problem solving involves some kind of restructuring of the initial faulty representation. A key difference between Progress Monitory Theory and Representational Change Theory is that the former accounts for the solution process until an impasse is reached and representational change is required, whereas the latter accounts for how an impasse is overcome to achieve insight (Jones, 2003).

Öllinger et al. (2006) investigated the relationship between heuristics and representational change in insight problem solving. In particular, they tested whether the main source of difficulty in insight problem solving arises from the use of inappropriate heuristics (MacGregor et al., 2001), or from self-imposed constraints on a given problem (Knoblich et al., 1999, 2001; Ohlsson, 1992), or whether both sources of difficulty influence insight problem solving (Jones, 2003). Matchstick arithmetic problems used by Knoblich et al. (1999) were adapted to create two different types of problem that required solvers to make two moves to transform an incorrect statement into a correct one. One type of problem required a value (numbers) move and an operator (i.e., +, -, or = sign) move in order to reduce the distance to the goal. Another type of problem required a value and an operator move, which initially increased the distance to the goal, and the solution had a tautological structure. The latter type of problem is considered more difficult because the constraints that operators are constant and that an equation may have two operators have to be relaxed. The distance to a goal is defined as the numerical difference between the left and right side of the equation. For example, the following equation $IV = IV + VI$ has a distance of four, which requires a value and an operator move. Hence, I from the VI must be moved before the V to make $IV = IV + IV$, thereby

increasing the distance to six. Then the vertical stick from the + operator must be placed horizontally to make the + into a =. The solution ($IV = IV = IV$) has a tautological structure with a distance of zero.

Öllinger et al. (2006) used these problems to test the assumptions of both Progress Monitoring Theory and Representational Change Theory. According to Progress Monitoring Theory, the maximization heuristic should be used to reduce the distance to make the left and right side of the equation more similar. Further, the progress monitoring heuristic is used to assess progress after every move to determine whether there is a consecutive move that would balance both sides of the equation. The theory predicts that problems requiring moves to reduce the distance to the goal should be easier than problems that require moves to increase the distance to the goal, which has been previously demonstrated in well-defined problems such as the Hobbits and Orc problem (Thomas, 1974). Representational Change Theory predicts that problem difficulty in matchstick problems is determined by the degree to which constraint relaxation is required. Thus, problems requiring a value and an operator move to produce a tautological structure will be more difficult because of the greater nature of constraints involved.

The results by Öllinger et al. (2006) indicated that the type of move required and whether a move increased or reduced the difference to the goal did not influence problem difficulty. Instead, problem difficulty was influenced by whether or not representational change was required. Hence, performance was poorer for equations requiring a solution with a tautological structure, which was likely to be due to the greater degree of representational change required by these problems. These findings were interpreted to suggest that solvers do not prefer an approach that reduces the distance to the goal and thus similar conclusions to Jones (2003) were drawn. That is,

Progress Monitoring Theory explains problem solving before an impasse, in which inappropriate heuristics are applied to an inappropriate representation, and Representational Change Theory explains the impasse phase in which the representation is changed by applying more appropriate heuristics.

In summary, support for Progress Monitoring theory comes from visuo-spatial insight problems. To some degree, these problems can be broken down into states that are determined by a move made by the solver. In contrast, verbal insight problems cannot be conceptualised in terms of states, therefore the heuristics emphasised by Progress Monitoring Theory are not applicable to solving these problems. However, recent research suggests that representational change determines which heuristics are applied to overcome impasse in the types of insight problems in scope to Progress Monitoring Theory (Öllinger et al., 2006). Therefore, the importance of representational change in overcoming difficulties in insight problem solving was emphasised.

2.6. Summary

Insight problems are considered difficult because a problem solver's past knowledge activates unhelpful information which consequently inhibits the knowledge that is required to solve a problem. A faulty problem representation is formed and solvers inevitably experience an impasse. The theoretical approaches to understanding insight problem solving discussed above differ mainly in terms of their focus on how insight, or solution, to a problem is achieved. Although Problem Space Theory is not in scope to insight problems, it is nonetheless useful in explaining how problem solving occurs within the problem space of a problem. The premise of Gestalt Theory is that the problem representation, i.e., how people perceive, interpret,

or organise information influence solution generation. Insight problem solving is characterised by an initial period of no progress (i.e., impasse), followed by a sudden restructuring of the problem representation and the sudden appearance of the insight solution, although some research suggests that a solution is preceded by several failed solution attempts (Durso et al., 1994; Weisberg & Alba, 1981). However, this view did not explain how a problem is re-represented.

The representational change accounts also suggest that impasses occur due to an inappropriate problem representation that constrains the problem space (Knoblich et al., 2001; Ohlsson, 1992). However, unlike Gestalt Theory, Representational Change Theory and Restructuring Change Theory specify possible mechanisms underlying representational change, with constraint relaxation receiving much support. Further, Representational Change Theory contradicts the Gestalt view of the suddenness of solution and instead suggests that information gathered during failed solution attempts i.e., partial insights lead the solver to the insight solution. The solution appears to be sudden to the solver because they have a lack of awareness of the processes that lead up to the solution. The alternative viewpoint proposed by Progress Monitoring Theory is that impasses occur as a result of inappropriate application of heuristics within the problem solving process (MacGregor et al., 2001) rather than due to an inappropriate problem representation. However, research suggests that Progress Monitoring Theory is in scope to visuo-spatial insight problems only. Further, recent research suggests that representational change is the key factor that determines the appropriate application of heuristics on the representation (Öllinger et al., 2006). In conclusion, it would seem that a faulty problem representation plays a critical role in explaining the difficulties experienced in solving insight problems. The next chapter, *Chapter 3*, provides a review of the different

methodological approaches that have aimed to promote representational change to facilitate insight.

CHAPTER 3

Review of methods that facilitate insight problem solving

In this chapter, methods that facilitate insight problem solving are reviewed including the role of incubation (Sio & Ormerod, 2009; Wallas, 1926), hints (Kaplan & Simon, 1990; Maier, 1931; Weisberg & Alba, 1981); and training (Ansburg & Dominowski, 2000; Chrysikou, 2006; Dow & Mayer, 2004; Kershaw & Ohlsson, 2004; Wicker et al., 1978). It is argued that the positive effects of incubation are limited to a narrow range of problems whereas the hints approach is only effective when it is tailored to the solution of specific test problems. Therefore, training may offer the best approach for investigating the facilitation of a range of novel insight problems. Furthermore, the focus of training, as suggested by insight theories, should be on promoting representational change or restructuring to achieve solution.

3.1. Incubation

It has been documented that a flash of insight occurs when a problem is put aside and attention is temporarily shifted away from the problem (Wallas, 1926). Wallas (1926) labelled this temporary abstention from conscious problem solving as 'incubation', where incubation is defined 'as an increase in the probability of finding the solution to a problem after a pause' (Schooler et al., 1993, p. 171). Thus, providing a break during insight problem solving allows for incubation to take place. This section will discuss the evidence in support of incubation and provide possible explanations for why incubation works.

A recent meta-analytic review of 117 empirical studies by Sio and Ormerod (2009) provides support for the existence of incubation effects in problem solving and

suggests some potential moderators including whether the incubation period is filled or not, length of the preparation period, and the type of problem solved. Thus, greater performance was found when the incubation period was filled with intervening activities in contrast to working continuously on a problem or having a rest whilst solving the problem. Performance was also positively correlated with longer preparation periods. Finally, positive incubation effects were found more for creative writing problems (i.e., problems that have multiple solutions) than linguistic problems (i.e., those that require the solver to consider linguistic information in the problem e.g., riddles) and visual problems (i.e., those that require the solver to consider visuo-spatial information in the problem e.g., Duncker's 1945 candle problem). These findings imply that incubation is less likely to be helpful in solving verbal insight problems, which are used by experiments reported in this thesis.

Three of the most prominent hypotheses proposed to explain the effects of incubation are spreading activation, selective forgetting, and problem restructuring hypotheses. The spreading activation hypothesis (Smith, 1995; Yaniv & Meyer, 1987) proposed that activation spreads to previously ignored items that are relevant to the solution. The selective forgetting hypothesis (Smith, 1995) proposed that time not spent on the problem weakens the activation of inappropriate information that distracts the solution path in the initial attempts at problem solving. Finally, the problem restructuring hypothesis proposed that a solver's mental representation of a problem will be reorganised after several unsuccessful attempts. Hence restructuring may occur through adopting a new problem strategy (MacGregor et al., 2001) or by relaxing self-imposed constraints that inhibit problem solution (Knoblich et al., 1999; Ohlsson, 1992; Segal, 2004).

Sio and Ormerod (2009) reported that support for the above hypotheses is specific to particular problem types. Hence, support for spreading activation comes from creative problems that have multiple solutions requiring a wide search of the solver's knowledge, which is facilitated by an incubation period. If the solution to the problem lies within already activated knowledge that is represented incorrectly, then widening the search for new knowledge is unlikely to be helpful. Support for the restructuring hypotheses comes from linguistic (or verbal) and visual (or visuo-spatial) problems, which have one possible solution. Verbal problems illustrate restructuring resulting from the relaxation of inappropriate constraints (Ansburg & Dominowski, 2000), whereas visuo-spatial problems illustrate restructuring resulting from a strategic shift in the search process following an impasse (MacGregor et al., 2001). Finally, support for the forgetting hypothesis is mixed.

To summarise, insight problems are difficult because inappropriate knowledge leads a solver to form an incorrect representation of a problem. Therefore, incubation provides time for the solver to forget such inappropriate knowledge and to view a problem from a fresh perspective upon returning to the problem. Positive effects of incubation have been linked to linguistic and visual problems although the phenomenon of incubation does not explain how the correct problem representation is identified. Moreover, research has demonstrated that insight still occurs even when no incubation period is provided (Dow & Mayer, 2004; Knoblich et al., 1999; Macgregor et al., 2001; Wicker et al., 1978). Thus, incubation may partly explain the process of insight on some occasions. In conclusion, the evidence in support of incubation effects is mixed (Dorfman, Shames, & Kihlstrom, 1996; Sio & Ormerod, 2009), as it is never clear-cut whether the problem solver has actually stopped trying to solve a problem during the break. As suggested by Sio and Ormerod (2009), incubation is

perhaps best understood through close examination of the problem to which it is applied and the conditions under which it is observed.

3.2. Hints approach

Another approach to facilitating insight problem solving is to provide the problem solver with a problem-specific hint. Research that has investigated the hints approach, which has mainly focussed on visuo-spatial insight problems, is reviewed below. It will be concluded that hints are only helpful when they are tailored specifically to the solution of the test problem.

Studies investigating the effects of hints can be traced back as early as Maier's (1931) experiment of the two-string problem, in which the experimenter appeared to accidentally brush against a string hanging from a ceiling which prompted the solver to attain the solution that items can be used as weights to swing one rope closer to the other. The underlying mechanisms for how hints operate are explained by insight problem solving theories. According to Progress Monitoring theory (MacGregor et al., 2001), hints expose the solver to problem-relevant information. Consequently, the solver is able to discover alternative moves when they experience criterion failure as a result of an impasse. In terms of Representational Change theory (Ohlsson, 1992), hints guide the solver to relax self-imposed constraints that otherwise block them from gaining insight into a problem.

The hints approach has predominantly been applied to facilitate performance on visuo-spatial problems (e.g., Burnham & Davis, 1969; Kaplan & Simon, 1990; Lung & Dominowski, 1985; Weisberg & Alba, 1981). For example, Kaplan and Simon (1990) tested the effectiveness of different types of hints to aid performance on the mutilated checkerboard problem. Verbal protocols were collected to examine

changes in the problem space more thoroughly. They found that when participants were given a 'parity' hint regarding what they might do to solve the problem, performance was greater than when given an 'impossible' hint i.e., what actions would be impossible in facilitating solution. These results were interpreted to suggest that 'parity' hints guided the solver to focus on finding an appropriate problem representation within the specific domain of the problem space, whereas the 'impossible' hint directed the solver to generate other unhelpful representations of the problem. However, Kaplan and Simon's observations do not indicate how participants identified the need for a new problem representation or what processes were involved in generating a new representation (Bohannon, 1994). Furthermore, it is not known whether such hints are generalisable to other insight problems (Cheung, Kong, Li, Wong, & McBride-Chang, 2005).

Several studies have developed hints to facilitate performance on the nine-dot problem, which is another visuo-spatial problem. Burnham and Davis (1969) gave participants an instruction to extend the line beyond the boundary of the nine-dot square. A training effect of 23% was found in the hint condition, whereas a later study by Lung and Dominowski (1985) found a 34% training effect. However, Weisberg and Alba (1981, Experiment 1) found 100% improvement when this hint was combined with a drawn line that was part of the solution to the problem. Weisberg and Alba's result is not surprising as giving hints that are directly related to the problem solution narrow the problem space and provide a critical starting point for the solution process.

Few studies have investigated the effects of hints, sometimes referred to as strategic instructions, in solving verbal insight problems (e.g., Ansburg & Dominowski, 2000; Perfetto, Bransford, & Franks, 1983; Wicker et al., 1978). The

hints utilised by these studies have differed in terms of their specificity. That is, some studies have provided general hints such as re-interpreting a problem several times to avoid making incorrect assumptions (Ansburg & Dominowski, 2000; Wicker et al., 1978), with training effects ranging between 48% (Ansburg & Dominowski, 2000) and 63% (Wicker et al., 1978, Experiment 2). On the other hand, others have investigated whether giving hints that were directly related to problem solution improved performance on several verbal insight problems such as the following (Perfetto et al., 1983, p. 25):

‘A man living in a small town in the U.S. married twenty different women in the same town. All are still living and he has never divorced one of them. Yet, he has broken no law. Can you explain?’

Perfetto et al. (1983) presented participants with hints, referred to as ‘clues’ in their study, prior to problem solving. For example, the clue related to the above problem was ‘A minister marries several people each week’. The solution was that the man was a minister/priest. It was found that only the participants who were explicitly told that the clue would help them to solve the problems performed well (54%). It is surprising that the training effect was not much higher as the clue was directly linked to the problem solution. However, participants who were not prompted to use the clues performed as poorly (29%) as those who were not given any hints at all (19%). An important implication of these findings is that to be able to use new knowledge or problem-relevant hints, it is critical that participants are made aware of the relevance of this knowledge.

In summary, research suggests that the hints approach facilitates positive transfer although the transfer effect is higher when the hint is directly linked to or gives a way a part of the solution to a problem. However, a general criticism of the hints approach is that, in providing a hint, the novelty of the problem is diminished whilst prompting the solution to the problem. Further, in order to be able to devise a useful hint, the solution to a problem must be known in advance. From an applied perspective, the generalisability of solution-specific hints is likely to be limited, particularly given the variability among insight problems. In any case, everyday problems are not accompanied by explicit hints and research indicates that problem solvers habitually default to past knowledge and experiences to aid them in their problem solving (Knoblich et al., 1992; Ohlsson, 1992; Sternberg & Davidson, 1995).

3.3. Training

A number of experimental studies have demonstrated that insight can be facilitated through training (Ansburg & Dominowski, 2000; Chrysikou, 2006; Cunningham & MacGregor, 2008; Lung & Dominowski, 1985; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981; Wicker et al., 1978). The central aim of training is to facilitate transfer, in which previous learning is applied to solve a current task or problem. Positive transfer occurs when past experiences enable a person to solve a problem. In contrast, negative transfer occurs when past experiences inhibit problem solving (e.g., Duncker, 1945; Luchins, 1942). There are two types of training: specific training and generic training. The former is aimed at facilitating transfer to test problems that involve the application of the same rules or procedures as those inculcated during training whereas the latter aims to facilitate transfer to a wider range of problems. Much of the research on training in problem solving is specific in that

positive transfer is facilitated to a similar problem soluble with the same process or heuristics that were learnt during training (e.g., Burnham & Davis, 1969; Chronicle et al., 2001; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981).

This section will first review research studies that have investigated specific training and transfer on insight problem solving to demonstrate that transfer effects are dependent on the degree of similarity between training and test problems. This is followed by a review of key generic training studies that have facilitated transfer to a range of different insight problems. It will be concluded that whilst specific training can be successful to some degree in facilitating positive transfer, the most important and difficult challenge is to develop more generic training that does not cue the specific answer to a problem and results in positive transfer to a variety of different problems.

3.3.1. Specific training and positive transfer

In the case of insight problem solving, the training literature highlights that transfer tends to be specific (e.g., Chronicle et al., 2001, Kershaw & Ohlsson, 2004a; Weisberg & Alba, 1981) such that what is learnt during training can be directly applied to solve test problems that are similar to training problems. However, research has generally focussed on improving performance on visuo-spatial insight problems such as the four-dot problem (e.g., Weisberg & Alba, 1981, Experiment 4) and the nine-dot problem (e.g., Burnham & Davis, 1969; Chronicle et al., 2001, Experiment 3; Kershaw & Ohlsson, 2004a; Weisberg & Alba, 1981, Experiment 2), with less attention paid to verbal insight problems. This section provides a review of specific training studies that have investigated visuo-spatial insight problems to conclude that the generality of the training is limited to a few specific problems.

An example of specific training that aimed to facilitate transfer on the nine-dot problem is provided by Kershaw and Ohlsson (2004). The authors assert that insight problems are difficult to solve because the solver is hindered by multiple sources of difficulty posed by a problem. They speculated three sources of difficulty that need to be alleviated to improve performance on the nine-dot problem. These sources include difficulties in processing the problem information, the influence of prior knowledge, and the way the problem is perceived by the solver. To overcome these difficulties, Kershaw and Ohlsson (2004, Experiment 3) trained participants in making non-dot turns i.e., making a turn where there is no dot, which is considered crucial to solving the nine-dot problem. The solution rate was raised from 7% to 40%, thus suggesting that other sources of difficulties excluding the ones identified by Kershaw and Ohlsson may operate in the nine-dot problem.

An earlier study by Weisberg and Alba (1981, Experiment 2) found training to go outside the pattern of dots for a four and ten dot problem raised the solution rate from 0% to 43% for the nine-dot problem. However, it is unlikely that training was maximised in the study by Weisberg and Alba as participants were not provided with feedback concerning their performance on the training problems, which may have limited improvement. Indeed, more recent research using verbal problems has found that solution feedback on practice problems can facilitate the solution of test problems (Ansburg & Dominowski, 2000, Experiment 4).

'Shape training' is another approach to specific training that has been utilised to facilitate transfer to visuo-spatial problems such as the four-dot problem (e.g., Weisberg & Alba, 1981, Experiment 4). In Weisberg and Alba's (1981) study, the shape that was practised during training could be directly transposed to the solution of the four-dot test problem. It was found that shape training was successful when this

shape was still relevant after one dot had been eliminated from the test problem. This highlights that the degree of specificity of the training with respect to the solution of a problem is very important in facilitating transfer. However, in addition to shape training in the Weisberg and Alba study, the practice problems allowed participants to join the dots together, thus facilitating outside dot turns on the four-dot problem. This may have contributed to the 91% solution rate that was found for the four-dot problem. It is not surprising that such problem-specific training is successful given that training is so closely oriented to the problem solution. However, this type of training is not 100% successful, which suggests that the context generated by the problem specification can still have strong inhibitory effects despite the fact that training is specifically tailored to the solution of that problem. Indeed, as discussed above, the transfer literature highlights that solvers need to be made aware of the relevance of their newly learnt skills to a problem in a new context in order for positive transfer to take place.

There is a lack of specific training studies that have investigated positive transfer using verbal insight problems. This is unsurprising as, unlike visuo-spatial insight problems, verbal insight problems do not have a clear goal state; therefore it is difficult to devise specific training for such problems. The closest to specific training for verbal insight problem solving is a study by Dow and Mayer (2004), which examined the effects of training in facilitating transfer to different categories of insight problems i.e., verbal, spatial, mathematical, or a combination of verbal and spatial problems. In Experiment 2 of their study, participants were trained on one or more of these types of problem and then tested on problems from both trained and untrained categories. Participants were trained in a three-step process for solving each type of problem. For example, training in verbal problems instructed participants to

notice that “one of the words is a trick or play on words” (p. 394) and then demonstrated, using three training problems, how this could facilitate problem solution through a three step explanation process of defining key words in the problem, analysing the meaning of these words in the context of the problem, and concluding with the solution. Neither verbal nor mathematical training improved transfer to the same category of test problem and an improvement using this criterion was only found with visuo-spatial training. Another experiment included a control group that received no training and, again, verbal training was no better than visuo-spatial training or the control group at solving verbal problems. It is possible that the categories were too broad, which increased the variability in the nature of the problem constraints.

In summary, specific training studies on visuo-spatial problems have demonstrated weak transfer effects in spite of the strong correspondence between training and test problems, thus suggesting that there is still much room for improvement. It can be argued that by giving training on problems that are structurally similar to test problems, participants are given components of the solution (Chronicle et al., 2001), therefore a problem is no longer viewed as a novel problem by participants. Furthermore, perceptual problems, as proposed by Kershaw and Ohlsson (2004), are more likely to be associated with visuo-spatial problems whereas verbal insight problem solving is likely to be influenced by the problem solvers past conceptual knowledge and how the problem information is processed (Knoblich et al., 1999; Ohlsson, 1992). From an applied perspective, such problem-specific training is not useful unless we have knowledge of both the insight problems and their solutions, which defeats the objective of facilitating transfer to novel, unseen problems. Indeed, in the training studies discussed above, both training and test problems were

structurally and perceptually similar e.g., four-dot and nine-dot problems. Verbal insight problems, on the other hand, differ in terms of problem content, which suggests that training is needed that capitalises on any commonalities shared by such problems, which is explored in Experiments 1 to 3 reported in this thesis.

In conclusion, specific training studies encounter the same criticism as the hints approach in that training is likely to facilitate performance if it is linked to particular problem solutions. At a more general level, the success of any training should be measured by its ability to facilitate positive transfer to a range of novel insight problems, which was the aim of generic training studies reviewed in the next section.

3.3.2. Generic training and positive transfer

The alternative to specific training is to devise generic training that facilitates transfer to not only the particular problem for which the training was designed but also to novel insight problems (e.g., Adeyemo, 2003; Ansburg & Dominowski, 2000; Chrysikou, 2006; Wicker et al., 1978). However, a few generic training studies have attempted to investigate transfer to a range of insight problems, which are reviewed below to demonstrate that, although transfer effects tend to be modest, they highlight the importance of focusing solver's attention on the incorrect representation of a problem (e.g., Ansburg & Dominowski, 2000; Chrysikou, 2006; Wicker et al., 1978). It is argued that the difficulty in devising generic training lies in the fact that the nature of the problem constraint or the stereotypical assumption associated with insight problems vary so much between problems.

One example of generic training is what Adeyemo (2003) labelled 'definitions pre-training', which entailed providing definitions of key terms that were required to

solve test problems (Adeyemo, 2003; Mayer, Dyck, & Cook, 1984). Adeyemo (2003) trained children aged 12 and 13 in the use, meaning and definitions of key words such as 'x-ray', 'tumour', and 'malignant' that are used in Duncker's (1945) x-ray problem. Two other problems used in this study were Maier's (1933) candle problem and Gick and Holyoak's (1980) oil well problem. Participants then used these key words in sentences to demonstrate their understanding. Following this, participants either solved a problem that was considered near (oil well problem) or far (candle problem) from the experimental problem before attempting to solve Duncker's (1945) radiation problem. Solution rate was increased from 15% to 37.5%. Adeyemo concluded that definitions pre-training improved problem solving performance on both near and far problems i.e., problems that were either specific or novel to training problems when compared to a control group that received no training. However, a criticism of such pre-training is that it may have either drawn attention to critical aspects of the problem or cued increased understanding of the context and the nature of the problem, thus enabling participants to reach the solution. Further, both the candle and oil well problems are actually structurally similar to the x-ray problem, which suggests that the approach to training was more specific than generic. Finally, it was not clear what aspect of the definition pre-training facilitated transfer.

The earliest attempt at some form of generic training for insight problem solving was reported in two experiments by Wicker et al. (1978). The effectiveness of training in two problem solving strategies was investigated: a problem reformulation strategy that encouraged problem solvers to keep changing their view of the problem to avoid making unnecessary assumptions and a visualisation strategy that instructed the problem solver to form a detailed visual representation of the problem.

Experiment 1 tested three training conditions: 1) a dual training condition that

involved a combination of a reformulation and visualisation strategy; 2) a visualisation only strategy condition; and 3) a no strategy condition receiving only practice. Participants in the training conditions attempted eight training problems, that were defined as insight problems as each involved an assumption that had to be overcome, and solution feedback was provided for each. All conditions including a control condition were tested on 11 different test problems. A five point scoring scheme was devised for each attempted test solution although the criteria for this scheme were not made explicit by the authors. Results showed that performance was significantly better for the dual training condition (58%) than both the control (44%) and visualisation (26%) conditions. A second experiment was conducted in order to determine whether the problem reformulation strategy alone was responsible for improved performance in Experiment 1. Results confirmed that performance was best for the reformulation strategy (63%) in comparison to two control conditions although participants were given only three practice problems in Experiment 2 as opposed to eight in Experiment 1.

Although Wicker et al.'s (1978) study was the first to provide some form of generic training for insight problem solving, it is difficult to assess its value because a list of the training and test problems was not given, which make it difficult to gauge the transfer relationships between them for the different conditions. The experiments also incur other criticisms. Firstly, the reformulation strategy was not specified in detail. Secondly, participants did not receive practice in identifying and rejecting inappropriate assumptions, which is addressed by the training utilised in Experiments 4 and 5 in this thesis. Finally, participants were not required to verbalise how they attempted to solve problems during the training and testing phases, and therefore there was no qualitative evidence concerning whether participants had actually applied the

trained strategies. This criticism is also addressed by the collection and analysis of verbal protocols in this thesis.

Ansburg and Dominowski (2000) conducted a series of five experiments to investigate the effects of various forms of generic training in facilitating performance on novel verbal insight problems. In Experiment 1, training consisted of four components:- (i) reading strategic instructions that emphasised reading a problem carefully, interpreting it in alternative ways, and not getting stuck on the most obvious interpretation; (ii) practising 15 verbal problems and when unsuccessful re-reading the problem twice, paraphrasing the problem in writing and receiving a hint; (iii) explaining, if still unsuccessful, why the provided solution is correct; and (iv) comparing and identifying the procedural similarities between practice problems in an ongoing fashion. Solution rate on 15 test problems for the training and a control condition was 60% and 34%, respectively.

Ansburg and Dominowski (2000) conducted a further four experiments in an attempt to extricate the effect of the different components of training in Experiment 1. Experiment 2 demonstrated that the use of strategic instructions alone (component i above) resulted in a 48% training effect. Experiment 3 found that employing the same training as Experiment 1 with only three practice problems made little difference to the solution rate (57%). Experiment 4 compared the effect of strategic instruction and solution feedback with another condition that also received guided practice (component ii above). Solution rate for both trained conditions was better than a control condition although there was no difference between the two trained conditions, thus suggesting that the guided practice had no extra beneficial effect. Experiment 5 provided training that differed from that in Experiment 1 in two respects. Firstly, no practice problems were given although solution feedback was

provided and, secondly, participants were encouraged to understand the similarities between problem solutions (component iv above) and to explicitly report a general strategy for solving such problems. Solution rate for the training condition was similar to that in Experiment 1, thus suggesting that attempted generation of practice solutions was not a necessary component of the training effect. Ansburg and Dominowski concluded that training should involve advance strategic instructions, practice with feedback on structurally similar problems and problem comparison to be most effective in promoting transfer.

A recent study by Cunningham and MacGregor (2008) attempted to replicate Ansburg and Dominowski's (2000) finding. Training consisted of written instructions warning participants that the first or obvious interpretations can be mistaken and that alternative interpretations need to be sought to find solutions to insight problems. An explanation in how to apply the training was given using the nine-dot problem and practice given on three matchstick arithmetic problems (Knoblich et al., 1999). A wider range of test problems were employed that consisted of three spatial problems, six verbal problems and 21 rebus problems, which combine spatial and verbal elements as words or symbols are presented that represent a well-known phrase or saying (MacGregor & Cunningham, 2008). Cunningham and MacGregor argued that spatial and verbal problems, in their original form, appear to be puzzle-like, and therefore may not be perceived as realistic which in turn may influence performance on such problems. Therefore, for each spatial and verbal problem, an analog version that had a more realistic cover was created. Half of the participants completed the original, puzzle-like versions of problems whereas the other half completed realistic versions of test problems. It was found that for spatial problems, training raised solution rate to 67% from 27% in the no training condition. However, training was not

effective in raising solution rates on verbal problems or realistic spatial problems. A possible explanation for the lack of replication for Ansburg and Dominowski's findings is likely to be due to differences in the training procedure. Cunningham and MacGregor's training focussed on spatial problems whereas Ansburg and Dominowski trained and tested participants on verbal problems. Further, the spatial problems were presented verbally, which suggests that training did improve performance on verbal problems. However, as participants were not aware of the change in context from the training to test problems, it would explain why participants did not show transfer to verbal or rebus problems.

Chrysikou (2006) proposed an alternative approach to training, labelled 'goal-derived categorisation training', by attempting to break stereotypical assumptions concerning typical objects (e.g., shoe, fork) in facilitating performance on insight problems. In Experiment 1, participants were allocated to one of four training conditions. Two conditions involved participants completing either the Alternative Categories Task (ACT) or the Alternative Categories with Critical Items Task (ACT-C). These tasks required participants to list alternative categories to which common items may belong. However, ACT-C also included items that were important for the solution of test problems. There were also two control conditions; one involved the Embedded Figures Test in which participants had to find previously seen simple figures embedded in larger complex figures; the other comprised of a Word Association task. All participants completed seven test insight problems (i.e., two functional fixedness problems, four verbal insight problems, and one visuo-spatial insight problem). Although the training appears to be specific to functional fixedness problems, there was no significant difference in performance between the ACT and ACT-C conditions, thus suggesting that performance was unaffected by whether the

training included the problem-specific object or not. A training effect of 57% was found in the ACT condition. Further, a second experiment replicated this finding although participants were not reminded to apply the training in solving test problems, as in Experiment 1. A possible explanation for this finding is that in breaking stereotypical categorisations of objects, participants were equipped with the general skill that the first interpretation is often incorrect and that alternative interpretations are required to solve an insight problem.

Both Wicker et al. (1978) and Ansburg and Dominowski (2000) demonstrated that generic training, which highlight the importance of interpreting a problem in alternative ways along with practice in doing this can improve performance on verbal insight problems. Further, Chrysikou (2006) demonstrated that training that focuses on breaking stereotypical categorisations of objects can facilitate solution on a range of insight problems including verbal problems. One commonality between these three studies is that solvers were made aware of the importance that the first interpretation is often incorrect and that the consideration of alternative interpretations is likely to lead to the problem solution. This is consistent with Gestalt Theory (e.g., Kohler, 1924; Wertheimer, 1945) and theories of representational change (Knoblich et al., 1999; Ohlsson, 1992) that advocate the importance of restructuring or representational change to overcome an impasse experienced during insight problem solving. However, despite the positive effects of training, solution rates still need to be improved.

In summary, the results of the generic training studies reviewed above are notable as they demonstrate that a short training procedure can facilitate transfer to novel insight problems. Thus, it is possible to train general insightful skills in facilitating performance across a range of insight problems. However, they also

highlight that further research on training is needed as transfer effects were not 100%. The generic studies discussed hitherto suggest that for generic training to be effective, problem solvers need to be given practice in identifying and correcting their faulty representations in order to overcome impasses during insight problem solving. Further, the nature of assumptions associated with insight problems vary from one another, therefore a generic form of training is needed that addresses this issue. Both of the above points are investigated by Experiments 4 and 5 reported in this thesis.

3.4. Summary

Three methods that were discussed that facilitate insight included incubation, hints, and training. The evidence in support of incubation effects are mixed, whereas the hints approach is only effective when the hint is specifically tailored to the solution of the test problem. The final approach, training, has demonstrated that it is possible to promote positive transfer to insight problems that are either specific or novel to training problems. In particular, research on specific training suggests that the degree of specificity of the training is a critical factor in determining the rate of transfer to the insight test problem. However, limited research is available that demonstrates spontaneous transfer to a novel set of insight problems, although generic training is the way forward in facilitating transfer across a range of problems. It seems that there are two important prerequisites for inducing positive effects of training; highlighting the relevance of training to trainees along with practice in representational change (Ansburg & Dominowski, 2000).

In conclusion, all three approaches demonstrate modest transfer effects, suggesting that no approach is completely successful and that there is room for improvement. However, of the three approaches, training requires further exploration

as it has been fruitful in facilitating transfer to either specific problems or a range of insight problems. The next four chapters in this thesis report experiments that investigated transfer to verbal insight problems either through specific training (*Chapters 4 and 5*) or generic training (*Chapters 6 and 7*).

CHAPTER 4

Heuristic-based training for verbal insight problems

As discussed in *Chapter 3*, training is one of the main methods that have been investigated to improve the solution of insight problems (e.g., Ansburg & Dominowski, 2000; Burnham & Davis, 1969; Chronicle et al., 2001, Chrysikou, 2006; Dow & Mayer, 2004; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981; Wicker et al., 1978), along with incubation (e.g., Dorfman et al., 1996; Schooler et al., 1993; Sio & Ormerod, 2009), and providing the problem solver with a hint (e.g., Burnham & Davis, 1969; Gick & Holyoak, 1980; Weisberg & Alba, 1981). However, much of the research on specific training has focussed on facilitating transfer to visuo-spatial insight problems although weak transfer effects have been reported. This chapter begins with a brief summary of this research, followed by a discussion of heuristic-based training as a form of specific training to promote transfer to verbal insight problems. Two experiments are reported that employed heuristic-based training as a possible methodology to promote transfer to specific categories of verbal insight problems.

4.1. Heuristic-based training and transfer to verbal insight problem solving

Much of the training research investigating specific training has focused on the nine-dot problem, which is a visuo-spatial insight problem (e.g., Burnham & Davis, 1969; Chronicle et al., 2001, Experiment 3; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981, Experiment 2), with training effects ranging between 40% (Kershaw & Ohlsson, 2004, Experiment 3) and 43% (Weisberg & Alba, 1981, Experiment 2). When training entails practice in drawing non-dot turns, which is critical to solving

this problem, performance is higher (Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981), which suggests that the degree of specificity of the training to test problems is crucial for positive transfer to occur. However, the training effects are modest, and suggest that other sources of difficulties may operate that affect performance on the nine-dot problem (Kershaw & Ohlsson, 2004). Further, the context generated by this problem has a strong inhibitory effect, therefore despite the fact that training is designed to improve performance on a specific problem, the solution rate is far from 100%. Finally, specific training reduces the novelty of the test problem by giving practice in problems that are structurally similar. From an applied perspective, the generality of these studies is limited to the test problem under study.

There is a lack of specific training studies designed to improve performance on verbal insight problems. This is most likely due to the fact that, unlike visuo-spatial problems, verbal insight problems do not have a clear goal state. Further, there is greater variability in the content and constraints associated with verbal problems. Therefore, the challenge is to devise and implement specific training that can facilitate transfer to verbal insight problems. One suggestion is to train solvers to use heuristics or rules of thumbs to solve insight problems that share certain characteristics. The remainder of this section provides a discussion of how heuristic-based training can be used to solve verbal insight problems.

Heuristic-based training is one type of training that has been used successfully in other problem solving domains, including industrial faultfinding (Shepherd, Marshall, Turner, & Duncan, 1977) and mathematics (Schoenfeld, 1979). General heuristics such as hill-climbing and means-end analysis have been successfully applied to solve well-defined, move problems (Chronicle et al., 2004; MacGregor et al., 2001; Newell & Simon, 1972; Öllinger et al., 2006; Ormerod et al., 2002). For

example, Kaplan and Simon (1990) suggested that, in solving the mutilated checkerboard problem, a visuo-spatial insight problem, solvers applied heuristics to narrow the space of possible moves to achieve solution. However, such heuristics cannot be applied to solve verbal insight problems as their goal state is ill-defined. It is also important to note that most training studies provide training with only one test problem rather than looking at a wider range of problems, which limits the applicability of the training to other insight problems. Thus, general problem solving heuristics that might apply across a wider range of problems is lacking in the insight problem solving domain (Chronicle et al., 2004), and further research is needed in this area, which is explored by the two experiments reported in this chapter. Although heuristic training is also narrow in some sense, it nonetheless is better than some literature that has only looked at one test problem (e.g., Weisberg & Alba, 1981).

One of the difficulties faced in designing heuristic-based training for verbal insight problems is that the nature of the stereotypical assumptions or constraints associated with the problems are so idiosyncratic (cf. Isaak & Just, 1995), such that it is difficult to envisage how knowledge of any one of these could be used to facilitate the solution of other problems. One possibility is to develop an intermediate categorisation that identifies commonalities between particular types of insight problem and to design heuristic-based training based on these categories. This was the rationale for Experiments 1 and 2 reported in this chapter.

One study that attempted to categorise insight problems for the purpose of training is that by Dow and Mayer (2004; reviewed in *Chapter 3, Section 3.3.1*). Dow and Mayer categorised problems by their overall nature i.e., verbal, visuo-spatial, mathematical or a combination of these. They trained participants in solving one or more of these types of problems and performance was tested on the different problem

categories. It was found training in verbal or mathematical problems did not improve performance on test problems of the same category. Only visuo-spatial training improved solution to the same category of test problems. In another experiment, no difference was found between verbal and visuo-spatial training in solving verbal problems when compared to a control group. It was possible that there was too much variability between the nature of the problem constraints and the categories were too broad which affected the results. Although it is useful to categorise problems for the purpose of training, it may be more beneficial to categorise problems in terms of constraints in order to narrow the categories and thus increase the rate of transfer

Consequently, as part of the training, Experiment 1 in this chapter first identified commonalities in constraints between particular types of insight problem and heuristics were developed. Thus, participants were first made aware of problem constraints and then trained on two heuristics that involved identifying ambiguous words and ambiguous names within a problem that lead to solution. For example, 'guide' is an ambiguous word as it can refer to either a human/animal guide or a map. For ambiguous names, the names in the Anthony and Cleopatra refer to animals, not humans. The heuristic to consider names as ambiguous should discourage participants from making this assumption during testing. Experiment 2 aimed to improve solution of problems containing ambiguous words by adapting the training utilised in Experiment 2 and by using a new set of test problems. It was predicted that, in each experiment, positive transfer would be restricted to trained category of test problems as suggested by theoretical formulations of transfer (e.g., Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901), and no transfer would take place on problems that were out of scope to the training. The think aloud methodology was utilised in both experiments.

4.2. Experiment 1

The aim of Experiment 1 was to design and implement specific training to facilitate transfer to verbal insight problems. Common problem constraints were identified and then simple heuristics were developed to examine the effects of heuristic training on verbal insight problems. The heuristics concerned searching for any ambiguous words or names in the problem statements and identifying their alternative interpretations. In addition, a third ‘out of scope’ category of verbal insight problem was used that could not be solved by either of these heuristics in order to shed light on the level of specificity associated with transfer of this training. It was predicted that positive transfer would be restricted to the two trained categories of verbal insight problems. Further, it was hoped that the verbal protocols might reveal whether solution attempts or the hypotheses generated by training and control (no training) participants were influenced by their problem representation. Thus, no training participants were more likely to generate hypotheses that were based on an inappropriate representation which they were unable to overcome spontaneously.

4.2.1. Method

Participants

Forty-six psychology students from Cardiff University participated in this experiment as partial fulfilment of course requirements¹. Ten participants were rejected as they were familiar with at least one of the test problems. The final sample comprised 36 participants and ages ranged between 18-27 years ($M = 20.69$, $SD = 2.16$).

¹ All materials were pilot tested using 10 participants, which revealed that four-minutes per problem were sufficient for participants to read and solve test problems. As no changes were required, datasets for these participants comprised part of the final sample.

Materials

Training Programme

Training covered two categories of problem concerning either ambiguous words or ambiguous names and involved the following three stages:-

1. Solving problems that contained ambiguous words. This entailed:

a) awareness of ambiguous words in a problem and provision of a first heuristic.

b) practice in using this heuristic to solve a problem containing ambiguous words.

2. Solving problems that contained ambiguous names. This entailed:

a) awareness of the ambiguous nature of names in a problem and provision of a second heuristic.

b) practice in using this heuristic to solve a problem containing ambiguous names.

3. Practice solving problems containing either ambiguous words or names.

For awareness training in ambiguous words (Stage 1a), participants read an example of a problem that contained an ambiguous word and were given the heuristic 'If you cannot make sense of the problem then search for and identify any ambiguous word(s) and its alternative meaning(s)':-

A woman shoots her husband. Then she holds him under water for over 5 minutes. Finally, she hangs him. But 5 minutes later they both go out together and enjoy a wonderful dinner together. How can this be?

In this problem people might assume that the husband had been killed. Participants were told that 'shoot' was the ambiguous word and asked to think of alternative meanings for this word and to solve the problem. Participants then read that if 'shoot'

is thought of in the photography context, then it would be the case that the wife had shot a picture of husband, developed it under water, and then hung the photograph to dry. This interpretation would explain why the husband was alive.

Next, participants were required to practise using the heuristic to solve a problem containing an ambiguous word (Stage 1b):-

A man walked into a bar, and before he could say a word, he was knocked unconscious. Why?

In this problem the word 'bar' is ambiguous. The man had walked into a metal bar rather than a drinking bar.

For awareness training in names associated with animals (Stage 2a), participants were told of the ambiguous nature of names, which do not always refer to humans, and then were given the heuristic 'If you cannot make sense of the problem then search for a name and identify what animal could be involved'. Participants read the following problem:

Spike, an adult, brings the paper to Mr. Hopkins every day. Spike is never paid for this. Why does he do this?

The solution was that Spike was a dog taking the paper to his owner, Mr Hopkins.

Next, participants were required to practise using the heuristic to solve the following problem (Stage 2b):-

It was a Sunday morning and music was playing in the background. Charlie was sitting, minding his own business. However, when the music stopped, a shadow fell over Charlie which led to him being crushed to death. Why?

The solution was that Charlie was a small animal such as a bug that was crushed to death because someone sat on him.

In the final part of the training participants were presented with two problems (Stage 3). Participants were required to apply the training to identify any words or names with alternative interpretations that the problems may contain and to think aloud whilst solving the problems. The problems were:

Problem 1

While on safari in the wild jungles of Africa, Professor Quantum woke one morning and felt something in the pocket of his shorts. It had a head and tail, but no legs. When Quantum got up he could feel it move inside his pocket. Quantum, however, showed little concern and went about his morning rituals. Why such a casual attitude toward the thing in his pocket?

Problem 2

Bobby had not taken anything and was feeling fine but he couldn't help repeating everything Mr Jenkins said. Why is that?

Problem 1 contained the ambiguous words 'head' and 'tail', and thus referred to a coin, which is why Quantum was not concerned. Problem 2 contained two names but as Bobby was repeating what Mr Jenkins was saying, it suggests that Bobby was a parrot.

Test Problems

Nine test problems were used: six that were in scope to the training of which three problems contained ambiguous words (Married, Guide, and King & Queen) and three problems contained ambiguous names that were related to animals rather than humans (Anthony & Cleopatra, Mr Jones, and Jason). A further three problems were used where neither of the rules can be applied, and thus were out of scope to the training (Bombs Away, Rope, and Sons). The problems were taken from Sloane

(1992), Ansborg and Dominowski (2000), and Dow and Mayer (2004) (see Appendix A).

Design

The independent variables were condition (training and no training) and problem category (ambiguous words, ambiguous names, and out of scope). Participants were randomly allocated to a condition and completed all problems from each category. Presentation of the test problems was randomised. The dependent variable was whether the problems were solved.

Procedure

Each participant was given an introduction to the experiment and then requested to 'think aloud' during each problem. To facilitate this, practice exercises were given involving different tasks and contexts as recommended by Ericsson and Simon (1980, 1993). These involved solving a multiplication problem, calculating the number of windows in the participant's house, and naming 20 animals.

After completing the think aloud training, participants in the control condition completed the test problems whereas participants in the training condition completed the training programme beforehand. Participants were given a four-minute time limit for each problem and if they were silent for a period of time, the experimenter again used two non-directive prompts ('What are you thinking?' and 'Please keep talking'). Verbalisations were recorded continuously. After completing each test problem, participants were required to rate how familiar they were with that problem on a 5-point scale (1 = very unfamiliar, 5 = very familiar). Participants were not given solution feedback. Finally, participants were debriefed and asked not to reveal

information about the experiment to others. The duration of the experiment was approximately one hour.

Qualitative categorisation and assessment of reliability

Verbal protocols were analysed primarily for two reasons. Firstly, to verify that training influenced the types of hypotheses or solutions generated by participants and secondly, that training was responsible for the improvement in problem solving performance. The following two-stage procedure was developed to identify and categorise hypotheses and to assess inter-coder reliability.

1. Three coders (study researcher and two unrelated researchers) read the protocols and individually identified what they considered a hypothesis/solution for each of the test problems completed by participants. A hypothesis or solution was defined as 'any verbalisation that attempts to answer the question posed in the problem statement'. Practice was given on a verbal protocol that was excluded from the final dataset to ensure coders had understood the task. Once all three coders had completed identifying the hypotheses, any discrepancies were discussed until a consensus was reached and a final list of agreed hypotheses was produced. A total of 787 hypotheses were initially identified between all three coders of which 750 or 95.3% were identified by all three coders, 14 or 1.78% were identified by two coders, whereas 23 or 2.92% were identified by one coder at the first stage of the categorisation process. After discussion, eight hypotheses were rejected making the final total 779.

2. The second stage involved the same three coders independently categorising the agreed 779 hypotheses into one of four categories: incorrect ambiguous word, incorrect ambiguous name, incorrect other, or correct. Incorrect ambiguous word hypotheses were so called because participants correctly identified the ambiguous word but incorrectly hypothesised its alternative meaning. For example, in the Guide problem, some participants correctly identified 'guide' as the ambiguous word but hypothesised that it referred to a 'goat', 'sheep', or 'dog'. Incorrect ambiguous name hypotheses were so called because participants correctly identified the ambiguous name but incorrectly identified which animal the name referred to. For example, in the Anthony and Cleopatra problem, participants incorrectly hypothesised that the names referred to 'insects', 'dogs', 'cats', 'butterflies', or 'birds' but not 'fish'. Incorrect other hypotheses were those that could not be categorised using the other categories. Of the 779 hypotheses, it was found that 774 or 99% were categorised by all three coders using the same categories and the remaining 5 or 1% were categorised by two coders only. The Perreault and Leigh (1989) reliability index was selected to calculate inter-coder reliability between pairs of coders as it accounts for differences in reliabilities due to the number of categories and also focuses the issue of reliability on the whole coding process (Kolbe & Burnett, 1991). The reliability indices across the 779 hypotheses for the three pairs of coders were 0.99 respectively, which is considered acceptably high as it exceeds 0.80 (Gremier, 2004; Krippendorf, 1980).

4.2.2. Results and Discussion

Ten participants rated at least one of the test problems above three on the familiarity scale and therefore were rejected from the final sample, which comprised 36 participants. An exploratory data analysis did not reveal any violations of homogeneity and normality of the data, therefore Analysis of Variance (ANOVA) was conducted. It was predicted that transfer would be specific and therefore positive transfer would be restricted to the two trained categories of verbal insight problem solving. As predicted, a 2 (condition: training and no training) x 3 (problem category: ambiguous words, ambiguous names, and out of scope) ANOVA revealed that more problems were solved in the training than in the no training (control) condition ($F(1, 34) = 13.11, MSE = 11.34, p < .01$) although this effect interacted with problem category ($F(2, 68) = 6.94, MSE = 3.90, p < .01$). Simple main effect analyses indicated that training was better than the control condition in the ambiguous words ($p < .01$) and ambiguous names categories ($p < .01$) but there was no difference in the out of scope category, as predicted. Solution rates are given in Table 4.1.

Table 4.1. Effect of training on solution rates of problem categories (three problems in each category)

Problem Category	No Training (Control)		Training	
	Mean	SD	Mean	SD
Ambiguous Words	0.78	1.11	1.78	0.73
Ambiguous Names	0.11	0.32	1.17	1.15
Out of Scope	0.78	0.65	0.67	0.59

Figure 4.1 reports percentage solution rate between training conditions and problem category. For each category, the solution rate for the no training condition was taken as the baseline measure. It can be seen that within the ambiguous word category, training raised solution from 26% (no training) to 59% (training) whereas

for the ambiguous name category, solution was raised from 4% (no training) to 39% (training). Thus, a facilitation of 33% was observed for the ambiguous word category whereas a facilitation of 35% was observed for the ambiguous name category.

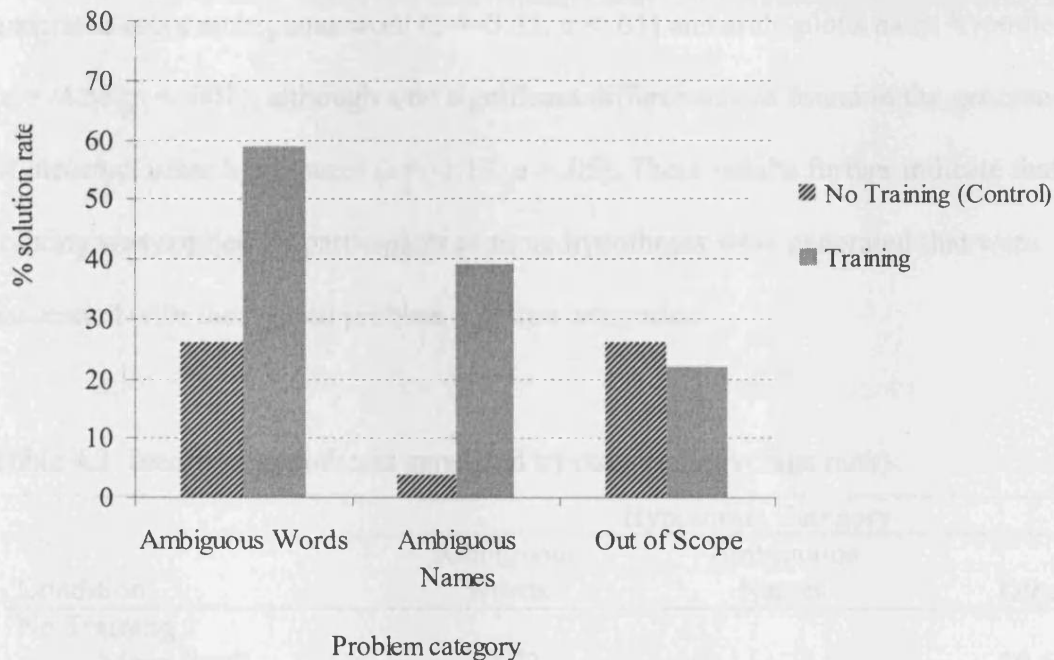


Figure 4.1. Percentage solution rate by conditions and problem category

Analysis of hypotheses generated

To further evaluate the effectiveness of training, verbal protocols were analysed to determine whether participants in the training condition had successfully identified the correct ambiguous word or name but failed to generate the correct alternative interpretation of that word or name. It was expected that training participants would generate a greater number of hypotheses that were associated with ambiguous words or names, but fewer (other) hypotheses that fell outside these two categories.

Exploratory data analysis revealed that the data was positively skewed for ambiguous words and names in the no training condition, which could not be eradicated using log, square root or reciprocal transformations, therefore Mann-Whitney *U* tests were conducted (Table 4.2). As expected, training participants generated more ambiguous word ($z = -3.53, p < .01$) and ambiguous name hypotheses ($z = -4.63, p < .001$), although a no significant difference was found in the generation of incorrect other hypotheses ($z = -1.17, p > .05$). These results further indicate that training was applied by participants as more hypotheses were generated that were associated with the trained problem solution categories.

Table 4.2. Incorrect hypotheses generated by condition (average rank)

Condition	Hypothesis Category		
	Ambiguous Words	Ambiguous Names	Other
No Training			
Mean Rank	12.72	11.11	20.56
Training			
Mean Rank	24.28	28.89	16.44

A tentative observation that was made during inspection of verbal protocol data was that some training participants attempted to solve a problem by initially looking for a name followed by an ambiguous word although the order of training was first on ambiguous words and then ambiguous names. On some occasions participants alternated between both when they were unsure of the problem solution. Thus, a strategy by which to approach problem solving was adopted by some of the solvers which was not observed in the control condition. An example of how training encouraged reasoning to a successful solution in the Mr Jones problem was:

“Mr Jones, it doesn’t state that he is a person. So he could be some kind of animal. It’s kind of ambiguous about what the medical practitioners are...they

might not be specialised in helping him with his leg. Although he broke his leg he could have died because of shock so although that's not an injury, that is a result of the accident...It could be an animal that has been run over...So he died as a result of breaking his leg and he was attended to immediately...He could be some kind of cat or dog. That's if he was run over or is an animal that can't survive without his leg. ...Unless he was really small, an insect... He was attended to by medical practitioners so it's more likely to be a bigger animal...He could be a horse because when horses break their leg they are often killed.”

On the whole, evidence from analysis of the verbal protocols indicated that training resulted in both qualitative and quantitative differences in the nature and number of hypotheses generated that were consistent with the nature of training. Thus, training encouraged participants to generate more solution related hypotheses concerning ambiguous words or names than no training participants.

In conclusion, training to be aware of and identifying ambiguous words and names improved the solution rates for novel exemplars of these problem categories. It is of interest that the transfer of training effect was quite specific and that positive transfer in terms of solution rate was observed only with respect to the trained problem categories, a finding consistent with theoretical predictions of transfer effects (e.g., Gick & Holyoak, 1980; Thorndike & Woodworth, 1901). However, there was some evidence that this transfer effect was slightly better for the ambiguous names category. A reason for this may be that the problem space is greater when searching for ambiguous words rather than ambiguous names that can be linked with animals. Consequently, problem solvers may have found problems belonging to the ambiguous name category slightly easier. Also, the think aloud training required participants to

list 20 animal names which may have helped the participants in the training condition to solve the test problems containing ambiguous names. Another explanation is that it was possible that the results may have been confounded by the choice of three test problems (Guide, Married, and King and Queen) belonging to the ambiguous word category. That is, the words King and Queen within the King and Queen problem refer to a married couple which may have reinforced the problem constraint in the Married problem, and vice versa. The aim of Experiment 3 was to both replicate and increase performance on problems belonging to the ambiguous word category. The training in Experiment 2 was adapted to increase participants' likelihood of correctly identifying and interpreting ambiguous words within a problem. Furthermore, to overcome the criticism of problem choice discussed above, the effects of training were investigated using a different selection of test problems.

4.3. Experiment 2

The training in Experiment 1 directed participants to identify the ambiguous meanings of words in verbal insight problems. A training score of 59% was found thus suggesting room for improvement. Experiment 2 aimed to both replicate this finding and to increase performance on problems with ambiguous words by using a new set of test problems to overcome the confounding issues concerning Experiment 1, as discussed above. Unlike Experiment 1, the approach to training in Experiment 2 focussed on one category of verbal insight problem solving i.e., ambiguous words, that could be solved by the use of a simple heuristic. Consequently, training was devised that covered awareness of this problem constraint coupled with practice at using the relevant heuristic to solve this type of problem. Thus, a problem containing an ambiguous word is one where the word can be interpreted in more than one way

other than the most common interpretation e.g., guide can mean 'map' or a 'person that directs'. The number of test problems in the ambiguous word category was increased from three, as used in Experiment 1, to four in Experiment 2. As in Experiment 1, a second out of scope category of verbal insight problems was also used that could not be solved by the heuristic. It was predicted that positive transfer would occur to test problems of the trained category.

An additional aim of Experiment 2 was to investigate whether problem length influenced performance on verbal insight problems. It is apparent that verbal insight problems tend to vary in terms of length, as illustrated throughout this thesis (cf. Isaak & Just, 1995). Hence, it could be argued that when participants in Experiment 1 were presented with longer verbal insight problems, the search through the problem space to find a solution was increased, which consequently resulted in no solution. Surprisingly, past studies on insight problem solving have overlooked the possible effects of problem length on performance, therefore, in Experiment 2, original versions of verbal insight problems were deliberately lengthened by 12 significant words i.e., words with substantial meaning such as nouns and adjectives and not words like 'the' 'and' or 'it' and performance on original and lengthened versions was compared. It was ensured that sentences and words that were added had no relation to the ambiguous word in the test problem. It was predicted that performance would be poorer on lengthened problems as the search through the problem space to find the solution should take longer in contrast to original version of problems which are inherently shorter in length. Verbal protocols were collected, although not analysed, in order to keep the methodology as similar to the one employed in Experiment 1.

4.3.1. Method

Participants

Forty first year undergraduate psychology students from Cardiff University participated in this experiment as partial fulfilment of course requirements². Ages ranged between 18 to 40 years ($M = 20.13$, $SD = 3.88$).

Materials

The no training (control) condition completed a Word Association task (Cohen, 1975) containing a list of 100 words that served as a filler task. A similar filler task has been used by Chrysikou and Weisberg (2004) and Chrysikou (2006), which had no adverse effect on problem solving performance. Unlike Chrysikou's (2006) study in which participants read the words and wrote down their responses, the words were presented orally by the experimenter and participants were instructed to say out loud the first word that came to their mind. Participant's responses were noted beside the respective word by the experimenter. The aim was to ensure that participants in the no training condition were thinking aloud as the same was required of participants who completed the training. Words that had associations with solutions to test problems were omitted and replaced with words from a version of the list by Winer (2002).

Training Programme

The training programme was similar to that used in Experiment 2 although the focus was only on training in solving verbal insight problems that contained

² All materials were pilot tested using 10 participants, which revealed that four-minutes per problem were sufficient for participants to read and solve original and lengthened version of test problems. As no changes were required, datasets for these participants comprised part of the final sample.

ambiguous words. The difference was that different training problems were used in Experiment 2 to allow for a larger set of test problems to be used. The full training programme can be found in Appendix C.

Test Problems

Seven test problems were used: four problems were in scope to the training (Coin, Guide, Island, and Shoot) of which only the Guide problem was the same as in Experiment 1. A further three problems where the rule cannot be applied, and thus out of scope to the training, were used (Captain, Twins, and Water Tower). Original versions of the seven test problem can be found in Appendix B.

Seven lengthened version of the original problems were created by adding 12 significant words i.e. nouns and adjectives that elaborated the problem (see Appendix C). It was predicted that the addition of 12 words, that had no relation to the ambiguous word, would increase the search through the problem space for the solution.

Design

The independent variables were condition (training and no training), problem length (original and lengthened), and problem category (ambiguous words and out of scope), where the latter variable was repeated. Twenty participants were assigned to each condition of which 10 participants completed original version of the test problems and 10 completed lengthened version of the test problems. Thus four experimental groups were formed (trained participants who completed original problems, trained participants who completed lengthened problems, no training participants who completed original problems, and no training participants who

completed lengthened problems). Presentation order of test problems was randomised. The dependent variable was whether the problems were solved.

Procedure

Participants read the introduction to the experiment and then were given practice in 'thinking aloud', as in Experiment 2. After completing the think aloud training, participants in the control condition completed the Word Association task whereas participants in the training condition completed the training programme. Participants then completed original versions or lengthened versions of test problems depending on which experimental group they were assigned to. Participants were given a four-minute time limit for each problem and if they were silent for a period of time, the experimenter again used two non-directive prompts ('What are you thinking?' and 'Please keep talking'). Verbalisations were recorded continuously. After completing each test problem, participants were required to rate how familiar they were with that problem on a 5-point scale (1 = very unfamiliar, 5 = very familiar). Participants were not given solution feedback. Finally, participants were debriefed and asked not to reveal information about the experiment to others. The duration of the experiment was approximately one hour.

4.3.2. Results and Discussion

No participants rated the test problems above three on the 5-point familiarity rating scale, therefore data for all 40 participants were included in the analyses. An exploratory data analysis did not reveal any violations of homogeneity and normality of the data. It was predicted that positive transfer would be restricted to the trained category of verbal insight problem solving. Furthermore, solution rate for original

problems would be higher than on lengthened problems as the latter increases the search for solution.

A 2 (condition: training and no training) x 2 (problem category: ambiguous words and out of scope) x 2 (problem length: original and lengthened) ANOVA revealed no significant effect of problem length on performance ($F(1, 36) = .89, MSE = .05, p > .05$) or a significant three-way interaction between condition, problem category and problem length, ($F(1, 36) = .71, MSE = .05, p > .05$). The lack of a significant interaction suggests that training participants did not find original versions of test problems any easier to solve than lengthened problems with problem category having no effect on solution rate. One possible reason for these null results could be due to lack of power. The pilot test revealed that both lengthened and original versions of test problems could be solved within the four minute time limit, which provided an early indication that an effect of length may not be found, as confirmed by the above results. This oversight may also account for the no effect of length that was found in this experiment. Another explanation is that the problems actually needed to be lengthened by more than 20 words in order to observe an effect of length.

As no effect of length was found in the preceding analysis, the data was aggregated and a 2 (condition: training and no training) x 2 (problem category: ambiguous word and out of scope) ANOVA was conducted. As predicted, a significant effect of condition was found with more problems solved in the training ($M = 2.85, SD = .99$) than in the no training condition ($M = 1.05, SD = .94$), ($F(1, 38) = 33.25, MSE = 1.82, p < .001$). Further, no significant effect of problem category ($F(1, 38) = 2.89, MSE = .19, p > .05$) or an interaction between condition and problem category ($F(1, 38) = 3.07, MSE = .20, p > .05$) was found. Contrary to predictions,

simple main effect analyses indicated that training improved performance on problems belonging to both the ambiguous word and out of scope categories.

Consequently, a 2 (condition: training and no training) x 7 (problems) was conducted to elucidate the effects of training on individual problems. Although the data are binary, ANOVA is applicable based on the assumptions of Greer and Dunlap (1997) that a repeated ANOVA on such data preserves the type 1 error rate and that power tends to be unaffected. Further, other researchers have successfully analysed binary data using ANOVA (e.g., Knoblich et al., 1999; MacGregor et al., 2001; Öllinger et al., 2006; Ormerod et al., 2002). Solution rates are given in Table 4.3. As predicted, more problems were solved in the training than in the no training (control) condition ($F(1, 38) = 46.77, MSE = 8.45, p < .001$) although this effect interacted with problem ($F(6, 228) = 2.40, MSE = .39, p < .05$). Simple main effect analyses indicated that training was better than the control condition for all exemplars of the ambiguous words category. These were the Coin and Shoots problems (all $ps < .05$), and the Guide and Island problems (all $ps < .001$). Contrary to predictions, the Water Tower problem from the out of scope category also showed an effect of training ($p < .05$). The information in the Water Tower problem misleads the solver to think that the painter is painting the outside of the tower and but this interpretation does not explain why no one saw the painter. As training encouraged participants to consider alternative meanings of words, this may have directed participants to consider the alternative interpretations of the problem that the painter was painting the inside of the tower, which is the solution to the problem.

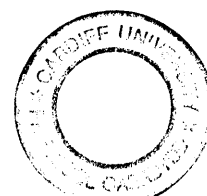


Table 4.3. Solution rates for test problems (as a proportion)

Problem Category	No Training (Control)		Training	
	Mean	SD	Mean	SD
Ambiguous Words				
Coin	0.50	0.51	0.80	0.41
Guide	0.40	0.50	0.95	0.22
Island	0.05	0.22	0.65	0.49
Shoots	0.10	0.31	0.45	0.51
<i>Overall Ambiguous Words Category</i>	<i>0.31</i>	<i>0.29</i>	<i>0.71</i>	<i>0.25</i>
Out of Scope				
Captain Scott	0.75	0.44	0.85	0.37
Twins	0.05	0.22	0.15	0.37
Water Tower	0.15	0.37	0.55	0.51
<i>Overall Out of Scope Category</i>	<i>0.32</i>	<i>0.20</i>	<i>0.52</i>	<i>0.23</i>

Table 4.3 illustrates a ceiling effect for the out of scope ‘Captain Scott’ problem in the no training condition, thus suggesting this problem was too easy. This finding may provide some explanation as to why a null effect of problem category was found earlier.

It was observed that training raised the solution rate from 26% (no training) to 71% (training) within the ambiguous word category (Figure 4.2). Thus, a facilitation of 45% was observed for the trained category. However, a t-test was conducted on proportion solution data after checking assumptions of homogeneity and normality, which revealed that training in Experiment 2 did not significantly raise performance on ambiguous word problems ($M = .71, SD = .25$) when compared to performance in Experiment 1 ($M = .59, SD = .25$), ($t(36) = 1.50, p = .07, 1$ -tailed).

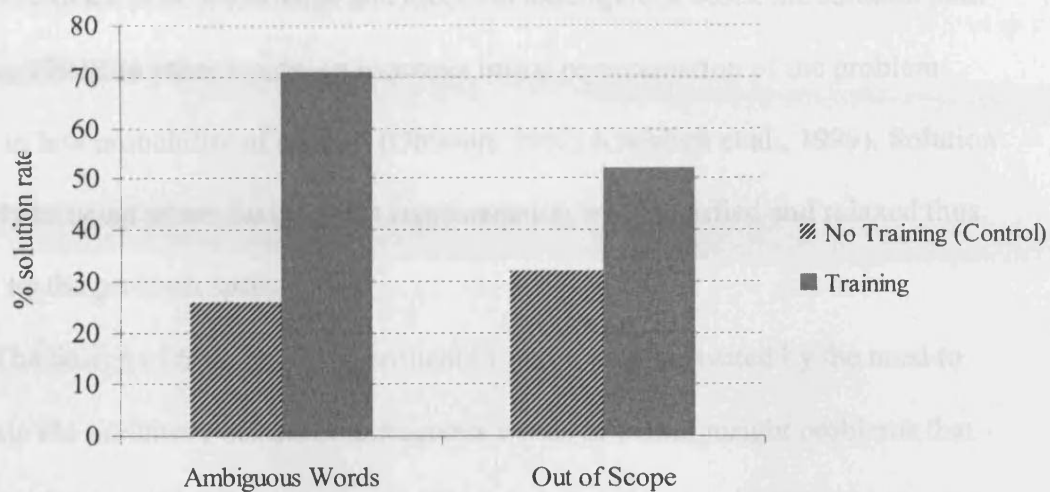


Figure 4.2. Percentage solution rate by conditions and problem category

In conclusion, training to be aware of and identifying ambiguous words improved the solution rates for novel exemplars of the ambiguous word category. However, the transfer of training effect was not categorically specific as transfer was also observed to one of the three out of scope problems. The data provided support, albeit not unequivocal, for heuristic-based training in promoting positive transfer to novel exemplars of the trained categories, and thus partially supported theoretical predictions of transfer effects (e.g., Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901).

4.4. General Discussion

The goal of this chapter was to investigate training to promote positive transfer on verbal insight problems through specific training. Two experiments were reported that investigated the effects of heuristic-based training on verbal insight problem solving. In both experiments, no training (control) participants found it difficult to solve test problems, thus supporting the assertion that presentation of a problem

triggers irrelevant prior knowledge and incorrect assumptions block the solution path (Ohlsson, 1992). In other words, an incorrect initial representation of the problem resulted in low probability of success (Ohlsson, 1992; Knoblich et al., 1999). Solution was likely to occur when the incorrect representation was identified and relaxed thus opening up the problem space.

The design of training in Experiments 1 and 2 was motivated by the need to ameliorate the inhibitory effects of ambiguous words in verbal insight problems that have a high probability of triggering a habitual conceptualisation leading to impasse. That is, training aimed to direct solvers to identify and consider the alternative meaning of ambiguous words in order to prevent the most common meaning of words from constraining problem solving and, consequently, solution attainment. However, the nature of the training varied slightly between Experiments 1 and 2 as training in the former focused on identifying ambiguous words and names within problem specifications whereas the latter focused on identifying ambiguous words only. The training was therefore specific in nature in that positive transfer was predicted to occur only to test problems belonging to trained categories. Indeed, the results of Experiments 1 and 2 found that transfer was specific to the test problems such that a greater number of problems were solved that belonged to trained categories. However, in Experiment 1, participants found ambiguous name problems easier to solve. A reason for this may be that the problem space is greater when searching for ambiguous words rather than ambiguous names that can be linked with animals.

Positive transfer was also observed in Experiment 2, with participants solving 71% of problems belonging to the ambiguous words category. However, the training effect was not specific as training improved performance on both the ambiguous word and out of scope categories. One possibility for this finding is that as training

encouraged participants to consider alternative interpretations of words within the problem specification, this heuristic may have encouraged participants to consider alternative solutions for all problems. However, analysis by problems indicated that successful performance was observed for only one out of three out of scope problems.

Experiment 2 also found that the length of the verbal insight problem did not predict problem solubility. Thus, the insight problems were difficult regardless of their length and performance on original and lengthened versions of the test problems did not significantly differ, thus suggesting that lengthened problems did not increase the search through the problem space to attain the solution. On the other hand, it was possible that the original problems in this experiment needed to be lengthened by more than 12 words to observe an effect. This suggests that the manipulation of word length was not sensitive in this experiment, which should have been detected from the results of the pilot study but was overseen due to experimenter error.

The approach to training in Experiments 1 and 2 was specific to test problems as it identified categories of verbal insight problem that were solvable by the use of heuristics. The success of training in both experiments is consistent with the success of similar training, for example, in mathematical problem solving (Schoenfeld, 1979). However, although training equipped problem solvers with heuristics to solve both ambiguous word and ambiguous name problems in Experiment 1, solution rates for both were far from perfect even after training, being 59% and 39% respectively. One would expect that the search for an ambiguous word would involve a process of trial and error with each word in the problem statement until an alternative meaning is found. In contrast, an ambiguous name could easily be found and the search for an appropriate animal is dependent on the problem solver's vocabulary. The solution rate for the ambiguous word problems in Experiment 2 was 71%, compared to 59% in

Experiment 1, which was not statistically significant. It was possible that the training effects were partially dependent on the problems that were selected for testing.

However, as training was not completely successful, it suggests room for improvement. Verbal protocols in Experiment 1 indicated that participants used the heuristics yet the modest solution rates suggest unexplained sources of difficulty. It can be speculated that a strong incorrect representation of the problem inhibited the identification of the ambiguous word or name, and its alternative meaning. This may have been further exacerbated due to uncertainty concerning which heuristic would lead to problem solution.

It is important to acknowledge limitations which may have affected the conclusions drawn from the experiments reported in this chapter. The choice of problems in Experiments 1 and 2 were limited due to lack of problems concerning the chosen heuristics. This meant that the same training and test problems could not be used although efforts were made to keep training as similar as possible. Furthermore, one of the practice problems (practice problem 1) in Experiment 1 unintentionally contained both an ambiguous word as well as a name, although the latter was not relevant to the training. This may have misled participants into thinking that the problems corresponding to the name category also contained ambiguous words. However, verbal protocol data suggests that participants attempted at first to generate solutions associated with names, and when they felt they had exhausted all possibilities they attempted to search for ambiguous words in a problem. This did not conform to the training sequence which addressed ambiguity of words before ambiguity of names. Further, is it not known whether the above strategy was adopted by some participants because they had a smaller problem space for names. Finally,

another limitation concerns interpretation of the training effects. It remains to be determined over what periods of time such training effects would persist.

In conclusion, verbal insight problems are difficult to solve because they involve restructuring or reorganising one's initial representation of the problem (Ohlsson, 1992). This poses a challenge not only to the problem solver as the shift in representation does not occur spontaneously but also to the design of training in terms of how restructuring can be facilitated to solve several problems that are novel to training problems. The two experiments reported in this chapter demonstrated how training can be successfully designed to make modest improvements on verbal insight problem solving. Experiments 1 and 2 concerned learning the heuristics for different types of constraint and training in categories of problems, which resulted in positive transfer to test problems that also belonged to the trained categories. Thus, the experimental results supported previous theoretical formulations of specific transfer (e.g., Gick & Holyoak, 1980; Thorndike & Woodworth, 1901). The next chapter will explore specific training to facilitate transfer to another type of insight problem that require solvers to consider novel functions of objects to reach solution.

CHAPTER 5

Specific training for functional fixedness verbal insight problems

Experiments 1 and 2, reported in *Chapter 4*, demonstrated that heuristic-based training can facilitate transfer to novel exemplars belonging to trained categories of verbal insight problems. However, as insight problems are not homogeneous (Bowden et al., 2005; Isaak & Just, 1995), such training is unlikely to be helpful in facilitating performance on insight problems which involve the use of objects in unfamiliar ways to attain solution. These are said to involve functional fixedness or functional fixity (e.g., Birch & Rabinowitz, 1951; Duncker, 1945; Maier, 1931). This chapter provides a brief discussion of research on functional fixedness to illustrate that in solving both functional fixedness problems and verbal insight problems, problem solvers encounter the same difficulties. That is, both types of problems trigger habitual conceptualisations that induce an incorrect representation and restructuring is required to solve such problems. A novel approach to specific training is proposed and tested in Experiment 3 that encourages divergent thinking to facilitate transfer to functional fixedness problems.

5.1. Divergent thinking and representational change in solving functional fixedness problems

Functional fixedness problems require an unfamiliar item/object to be used in a novel way to reach the solution (e.g., Duncker, 1945; Maier, 1931, 1945). As past experiences can inhibit general problem solving (James, 1890; Reason, 1990), it is not surprising that in solving functional fixedness problems, our experiences and knowledge of objects can lead to encoding of objects in terms of their typical

functional attributes (Keane, 1989), and an unwarranted impasse may result that can only be overcome through representational change (Ohlsson, 1992). This is consistent with what Reason (1990) proposed. That is, in cognitively underspecified situations solvers tend to default to high frequency responses. To overcome functional fixity, a specific form of representational change is required in that the solution generated must include a novel, unusual function for an object mentioned in the problem specification (Dunker, 1945; Maier, 1931). Therefore, similar difficulties are encountered in solving both verbal insight problems, studied in *Chapter 4*, and functional fixedness problems, as it is the unhelpful reliance on past habits which impedes performance on both types of problems. The remainder of this section provides a brief discussion of the evidence in support of functional fixity to demonstrate the importance of divergent thinking in overcoming functional fixity.

A number of experiments have attested to the existence of functional fixedness (Adamson, 1952; Birch & Rabinowitz, 1951; Duncker, 1945; Frank & Ramscar, 2003; Maier, 1931). For example, Maier (1931) found that participants were unable to generate a novel function for pliers in the two-string problem until the experimenter provided a hint. Duncker's (1945) candle problem is perhaps more famously known for demonstrating functional fixedness. Participants were set the task of attaching a candle to a wall so that it can burn upright, with only some matches and a box of drawing pins. Duncker found that participants tried to attach the candle directly to the wall with the drawing pins, or to glue it to the wall by melting it. It was concluded that participants were 'fixated' on the box's normal function of holding the drawing pins and could not reconceptualise it as a platform to solve the problem.

Adamson (1952) replicated experiments using functional fixedness problems conducted by Duncker (1945) due to limited original data. Participants in the

experimental 'pre-utilisation' condition experienced initial functions for objects in the problems which were intended to inhibit problem solutions (i.e. boxes in the candle problem were presented with items in them so their function as a container was prevalent). Control participants were given the problems without any pre-utilization so no prior functions for items were displayed (i.e. boxes were presented empty to participants). The solution rate for the experimental condition was more than twice than that of the control condition. These findings provided further support for functional fixity and suggested that those who utilise an object for a particular function in the past will have greater difficulty solving problems that require a novel function for that object. That is, past experiences reinforce the typical function of an object which has a pervasive effect on problem solving.

Both Adamson (1952) and Duncker (1945) used the same situation for the pre-utilisation and new task, therefore it was difficult to determine what aspect of the tasks was most difficult for participants. Furthermore, there was no control for the experience participants had with objects prior to the experiment (Mayer, 1992). To overcome these criticisms, Birch and Rabinowitz (1951) adapted Maier's (1931) two-string problem. They presented two experimental conditions with different pre-tasks while a control condition received no pre-task. Findings demonstrated that the situational context influenced what function is generated for a given object. In particular, participants' previous experience (gained during the pre-task) led them to use objects in a certain way and functional fixedness did not allow them to perceive the objects to be used for another purpose. Thus, the context of a problem can also have a strong inhibitory effect in solving functional fixedness problems, as a habitual response results in the formation of an incorrect representation which consequently prevents the solver from generating a novel or unusual function for an object.

More recently, Frank and Ramscar (2003) found that performance on Duncker's (1945) candle problem was dependent on participants' lexical representation of the concept 'box' rather than the instructional manipulation i.e., when noun phrases such as 'box of matches' were underlined, or only when nouns such as 'box' were underlined, or when the same instructions used in the original experiment were employed. They suggested that overcoming functional fixity in the candle problem is dependent on having a flexible representation of the word 'box' in order to realise that the box can serve as a platform to attach the candle to the wall.

The above studies illustrate that functional fixity can prevent divergent, flexible thinking i.e., thinking that entails generating numerous solutions to a problem (e.g., different uses for a brick), which may be an underlying process in insight problem solving (DeYoung et al., 2008; Gilhooly, Fioratou, Anthony, & Wynn, 2007; Gilhooly & Murphy, 2005). DeYoung et al. (2008) suggested that breaking frame (similar to breaking out of functional fixedness; Duncker, 1945) and divergent thinking were two broad cognitive abilities that both independently predicted insight. That is, breaking frame is necessary to avoid persevering with an incorrect problem representation, while divergent thinking is necessary to generate elements of a new representation. Further, flexibility also independently predicted insight, thus suggesting that the ability to switch between perspectives may be an important aspect of divergent thinking.

Chrysikou (2006) utilised training in functional fixedness problems but tested performance across a range of different test problems which included verbal insight, visuo-spatial and functional fixedness problems. Training entailed participants completing an Alternative Categories Task (ACT), which is a variation of the Unusual Uses Test (Christensen & Gilford, 1958). The task was presented in the form of a

questionnaire and required participants to generate up to six common categories for 12 items including a shoe and a fork. For example, it was stated that for 'shoe', the common category was 'item used as a footwear'. The Alternative Categories with Critical Items Task (ACT-C) was similar to ACT but participants in this condition were also given items that were critical to solving the test problems e.g., 'box' which is of importance to solving Duncker's (1945) Candle problem. Although the training appeared to be specific to solving functional fixedness problems, a training effect (57%) was found across all test problems even when the critical item was not included in the task. It appears that the training encouraged divergent thinking, which is particularly important in solving functional fixedness problems, although in this case it encouraged participants to assess alternative interpretations in triggering representational change, as suggested by the positive training effect.

The above studies support the Gestalt (Köhler, 1924; Wertheimer, 1945) and representational change (Knoblich et al., 1992, 2001; Ohlsson, 1992) perspectives of insight problem solving as past experiences trigger stereotypical responses, which in this case is a typical function of an object, that impede problem solving. Perhaps more importantly, the results draw attention to the role of restructuring or representational change in overcoming functional fixity. In other words, a switch in representation is needed to access low probability hypotheses that tend to be overridden by default, high probability hypotheses that in most situations are correct (Patrick et al., 1999). For example, in Maier's (1931) two-string problem, participants were unable to restructure their initial interpretation of the function of the pliers to serve as a pendulum weight without experimenter intervention (Adamson & Taylor, 1954). Similarly, solvers were unable to consider the box in the candle problem to serve as a platform to hold the candle (Dunker, 1945). It could be argued that that these

problems were made more difficult because real objects were provided within the problem context. Consequently, this may have strengthened the habitual response concerning the typical function of a particular object, therefore making it harder for the solver to overcome functional fixity. On the whole, research indicates that training to overcome functional fixity needs to be devised that encourages divergent thinking such that participants generate several uses for familiar objects in the hope of facilitating restructuring. Hence, the common function for an object, that is usually unhelpful, is abandoned early on during problem solving.

5.2. Experiment 3

One suggestion for training to facilitate divergent thinking is to devise training that encourages solvers to systematically generate several functions for objects in functional fixedness problems. In other words, the training encouraged solvers to re-encode objects in terms of atypical functions (Ohlsson, 1992). As with the training in Experiments 1 and 2, the training in the present experiment also cued the type of problem with the difference being that a process was introduced which focussed participants on the functionality of objects that is relevant to solving functional fixedness problems. This approach to training has never been investigated before, particularly where the problems were presented verbally like the verbal insight problems in this thesis.

Consequently, training was designed that involved the following steps: participants' were directed to select an object mentioned in the problem specification and to generate a function for that object. Next, participants were required to consider whether that function could be used to solve the problem. If not, then participants repeated the above process for that item until they were unable to produce anymore

functions. Another item was selected and the above process was repeated as many times as necessary until a solution was generated. The above approach to training was simple as it critically cued the solver to systematically consider several functions for an object, therefore encouraging both divergent thinking as well as averting the solver from fixating on the typical function of an object in a functional fixedness problem.

The aim of experiment 3 was to test the efficacy of this training by evaluating performance on two categories of insight problems: functional fixedness problems and verbal insight problems (that do not involve functional fixity) that were out of scope to training. However, unlike the original experiments that investigated functional fixity using real objects (e.g., Duncker, 1945; Maier, 1931), the functional fixedness problems in this experiment were presented verbally in order to keep in line with the out of scope, verbal insight problems. Further, by presenting functional fixedness problems verbally, it could be argued that the items were presented in a more neutral context than the original experiments, therefore weakening the effects of functional fixity. It was predicted that transfer would be restricted to the trained category of test problems as suggested by theoretical formulations (e.g., Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901), and no transfer would take place on problems that were out of scope to the training. The think aloud methodology was utilised in this experiment to keep in line with the methodology of Experiments 1 and 2.

5.2.1. Method

Participants

Twenty-four first year undergraduate psychology students from Cardiff

University participated in this experiment as partial fulfilment of course requirements³. Ages ranged between 18 to 21 years ($M = 19.13$, $SD = .90$).

Materials

As in Experiments 1 and 2, the Word Association task (Cohen, 1975), which served as a filler task, was completed by no training (control) participants.

Training Programme

Training was designed to facilitate performance on verbal versions of functional fixedness insight problems and involved the following two stages:-

1. Introduction to training to consider alternative functions of objects. This entailed:
 - a) A short explanation of how the training works.
 - b) A worked example of how to use the training to solve a problem.
2. Practice in using the training to solve problems. This entailed:
 - a) Solving one problem with experimenter guidance and prompting.
 - b) Solving one problem under test conditions.

For Stage 1a, a written explanation of how the training works was provided.

The aim of the training was to encourage participants to select items within a problem statement and to systematically consider the uses and functions of each object. After selecting and generating a function of an item, participants considered whether that function served as a solution for the problem. If the solution could not be used, participants were prompted to generate another function for the object and then to consider the plausibility of that solution. This process was repeated until all possible functions for an object was exhausted. Participants then selected another item and

³ All materials were pilot tested using 10 participants, which revealed that four-minutes per problem were sufficient for participants to read and solve test problems. As no changes were required, datasets for these participants comprised part of the final sample.

repeated the above process. Hence, this interactive process of selecting an item, generating a function, and checking the plausibility of the function as a solution was repeated until a solution was generated that could be used to solve the problem.

In Stage 1b, an example of how the training can be applied to solve the following problem was presented:-

Several wooden poles, clamps, and string have been made available. The task is to hang the string from the ceiling to the floor without defacing the ceiling.”

(Based on Maier, 1945)

The item ‘clamps’ was selected first to illustrate the use of the training. Possible functions of the clamps include using them to hold the string, to connect the poles together, and to use the clamps to attach things in position. One solution was to clamp the string to the ceiling. However, this solution is not usable as, according to the problem statement, there is nothing to clamp the string to the ceiling. Next, possible functions of the wooden poles were considered including using it to hold the string or to fit the poles between the floor and the ceiling vertically. A solution was to tie the string to the top end of the wooden pole and to place it against the ceiling. However, this solution is also unusable as the string is hanging against the wooden pole and not against the ceiling. Next, alternative functions of the wooden poles and clamps in combination were considered. These include using the clamp to attach two poles together vertically, using the clamps to form a right angle with the poles, using the clamps to form an arch, and to use the wooden pole to hang the string from. A possible solution to solve the problem was to tie the string around a pole and to firmly fix the pole against the ceiling by forming an arch structure using the clamps with two other poles.

Next, participants were required to practise using the training to solve the following problem (Stage 2a):-

A piece of white cardboard with four black squares fastened to it is to be hung from a ring fixed to the ceiling. On the table in the room are the following objects available: paper, a pen, a ruler and some paperclips. How could the cardboard squares be hung on the ring?

The solution was to bend one paper clip to form a hook from which to hang the large square.

In the final part of the training, participants were required to solve the following problem under test conditions (Stage 2b):-

Three cords are to be hung side by side from a wooden ledge. On the table in the room there is paper, pencils, tinfoil, two short screw-hooks and a hand powered screwdriver. How could the three cords be hung up?

The solution was to screw the two screw-hooks using the hand powered screwdriver and then to hang the chords on the hooks.

Test Problems

Six test problems were used: three functional fixedness verbal insight problems and three out of scope verbal insight problems (see Appendix E for original and altered problems). The functional fixedness problems were: the Two-String problem (Maier, 1931), the Candle problem (Chrysikou, 2006; Duncker, 1945) and the Hatrack problem (Maier, 1945). As past research has presented these problems in both pictorial and verbal format, changes were made for the purpose of this experiment and problems were presented verbally to match the presentation of the out

of scope problems. Some names for items in the problems were exchanged for more common terms to increase comprehension.

The three out of scope verbal insight problems were: the Charlie problem (Chrysikou, 2006; Weisberg, 1995), the Fake Coin problem (Chrysikou, 2006; Ansbarg & Dominowski, 2000) and the Prisoner and Rope problem (Chrysikou, 2006; Isaak & Just, 1995).

Design

The independent variables were condition (training and no training) and problem category (functional fixedness verbal insight problems and out of scope verbal insight problems). Participants were randomly allocated to a condition and completed all problems from each category. Presentation of the test problems was randomised. The dependent variable was whether the problems were solved.

Procedure

Participants read the introduction to the experiment and then were given practice in 'thinking aloud', as in Experiments 2 and 3. After completing the think aloud training, participants in the control condition completed the Word Association task before completing test problems. Participants in the training condition completed the training programme beforehand. Participants were given a six-minute time limit⁴ for each problem and if they were silent for a period of time, the experimenter again used two non-directive prompts. Verbalisations were recorded continuously. After completing each test problem, participants were required to rate how familiar they were with that problem on a 5-point scale. Participants were not given solution

⁴ Chrysikou (2006) allocated eight minutes for each problem, whereas Experiments 2 and 3 gave a four minute time limit; therefore the mean time of six minutes was used as a limit in this Experiment, which the pilot study revealed was sufficient for problem solving.

feedback. Finally, participants were debriefed and asked not to reveal information about the experiment to others. The duration of the experiment was approximately one hour.

5.2.2. Results and Discussion

No participants rated the test problems above three on the 5-point familiarity rating scale, therefore data for all 24 participants were included in the analyses. An exploratory data analysis did not reveal any violations of homogeneity and normality of the data, therefore ANOVA was conducted to test the hypothesis. It was predicted that transfer would be specific and therefore positive transfer would be restricted to the trained category of functional fixedness verbal insight problems.

A 2 (condition: training and no training) x 2 (problem category: functional fixedness and out of scope) ANOVA revealed that more problems were solved in the training than in the no training (control) condition, ($F(1, 22) = 7.62, MSE = 3.00, p < .05$) although this effect interacted with problem category, ($F(1, 22) = 6.60, MSE = 3.00, p < .05$). Simple main effect analyses indicated that training was better than the control condition in the functional fixedness problem category ($p < .01$) and that there was no difference in the out of scope, non-functional fixedness category, as predicted. In fact, it was observed that performance on the latter category was the same in both the control and training conditions. Solution rates are given in Table 5.1.

Table 5.1. Effect of training on solution rate of problem categories (three problems in each category)

Problem Category	Condition			
	No Training (control)		Training	
	Mean	SD	Mean	SD
Functional Fixedness	0.17	0.39	1.17	0.84
Verbal Insight	0.67	0.65	0.67	0.65

Figure 5.1 reports percentage solution rate between conditions and problem category. Solution rate for the no training condition was taken as a baseline measure. It can be seen that training raised the solution rate from 6% (no training) to 39% (training) for functional fixedness problems whereas the solution rate remained at 22% in both the no training and training conditions for the out of scope, verbal insight problems. Thus, a facilitation of 33% was observed for the trained category.

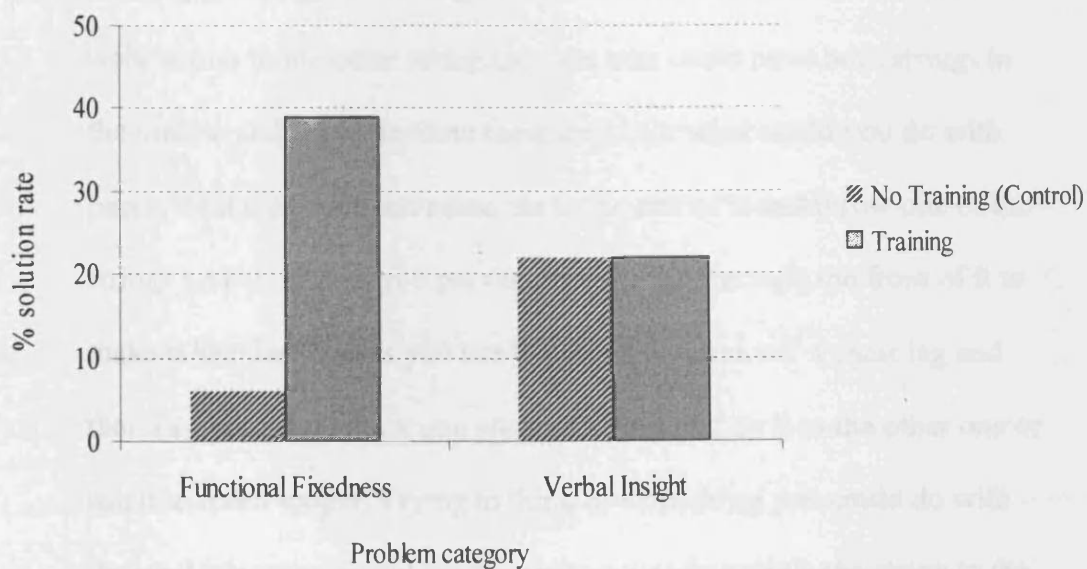


Figure 5.1. Percentage solution rate by conditions and problem category

Verbal protocols were not formally analysed as the primary aim of Experiment 3 was to design and test training to facilitate performance on functional fixedness

problems. However, a tentative observation that was made from inspection of the verbal protocols was that participants in the training condition indeed deliberated on different uses for objects in the functional fixedness problems, which enabled them to continue thinking of solutions after reaching an impasse as the problem space was broadened. Below is an excerpt that is representative of how training encouraged reasoning to a successful solution in the Two-String problem:

“The task is to reach one string while holding the other. So lets assume you’re standing directly underneath the piece of string and you put your hand on it and you cant reach the other one in a stretch...um, if you put the chair in between the two pieces of string in a central point, from the middle you would be able to reach one and then the other because the lengths of string would meet in the middle I assume....Or use the drawing pins to pin the string in the chair and hold one there and then walk across to the other string and then you could have both strings in the middle and could tie them together....Or what could you do with paper? Make a paper aeroplane, tie to the end of it and throw one of the strings across. Unless you put the drawing pin through the front of it to make it heavier. Unless you use the pliers to break off a chair leg and that way you could hook one piece of string and tie it to the other one or use it to reach longer. Trying to think of something you could do with the jar. Unless you could use the jar as a weight and tie the string to the jar and then swing it, but then it would swing back. But it would be easier. So you could swing it to the centre.”

The participant initially fixated their focus on the chair before considering several different ways of using it to reach the string. After referring to the training, they

changed their representation and considered alternative functions for other items in the problem such as the paper which lead them to the correct solution. There is a comparable difference in the structure of how this participant solved the problem in comparison to the excerpt below of a participant in the control condition:

“It says it’s impossible to reach them but it doesn’t say if that is not using anything. Um, you could stand, you could maybe, I don’t know, stand on the chair and try and reach it. Or move one string and pin it so that it is easier to get to and then stand on the chair and get it from where you pinned it so it’s nearer and then hold on to it. I don’t know. I don’t know how to use the pliers, paper or the jar but yeah, I would do something like that.”

This example is representative of how participants in the control condition exhausted ideas quickly. They were more likely than the training condition to accept the first solution they generated i.e., fixating on the typical function of an object such as using the chair to reach the string, which was often incorrect. Furthermore, when they reached an impasse, they were more likely to give up problem solving. This supports Reason (1990) in that control participants were likely to default to high-frequency responses that were incorrect in solving functional fixedness problems.

In conclusion, training in systematically considering alternative functions of objects as possible solutions to functional fixedness verbal problems improved solution rates on novel exemplars of this problem category. As predicted, positive transfer on functional fixedness problems was observed. Indeed, training encouraged the specific process of generating alternative functions for objects in the problems, which supports Keane (1989) who suggested that a problem solver needs to think

beyond previously encoded functional attributes of objects to overcome an impasse. However, although a positive training effect was observed, it is not clear whether it was the content or the process of training that resulted in improved performance on functional fixedness verbal insight problems.

5.3. General Discussion

The goal of this chapter was to investigate specific training to promote positive transfer on functional fixedness verbal insight problems. For this purpose, Experiment 3 utilised training that cued the functionality of items in a problem statement. In other words, the approach to training explicitly drew solvers' attention to the function of items that is critical to the problem solution. Hence, divergent thinking was encouraged as the solver was required to generate an exhaustive list of possible functions for individual items. It was expected that solvers in the no training (control) condition would initially interpret items in terms of their common functions that would trigger an incorrect representation of the problem, which would inhibit problem solving. Training was based on the notion that the solver could be trained to systematically consider the different functions of objects and to check the plausibility of each function as a solution to the problem. This process of considering different functions of objects should shift the faulty representation and cue re-encoding, as suggested by Ohlsson (1992) and this, in turn, should facilitate solution to problems belonging to that problem category.

The results suggest that training indeed facilitated performance to novel exemplars of the trained category as solution rate was raised from 6% (no training) to 39% for functional fixedness problems. However, this is a modest training effect, and suggests that other sources of difficulty may impede performance. As predicted,

transfer did not occur to non-functional fixedness problems i.e., verbal insight problems that were out of scope to the training. One possible reason is that the training directed the problem solver's attention to objects in a problem and generating object-related solutions which was not applicable to solving the out of scope problems.

There are some unresolved issues concerning the present experiment. Firstly, it is unknown whether the training would be as beneficial if original versions of the functional fixedness problems were used. However, it could be argued that by presenting functional fixedness problems verbally, the problems were made a little easier in contrast to the original versions of the problems (e.g., Duncker, 1945; Maier, 1931). This is because the original problems presented objects that were highly associated with the context in which they were presented in whereas in this experiment, the objects were presented in an abstract form within the problem specification, thus the assumption concerning the function of an object may have been less difficult to overcome. Of course, this requires further investigation. Secondly, it remains to be determined over what periods of time such training effects would persist. Thirdly, it is not clear from these results whether it was the process or the content of the training that led to increased solution rate on problems for the trained category. However, despite the aforementioned issues, it is important to note that this experiment was the first attempt at assessing problem solving of functional fixedness verbal insight problems where past research presented these problems in both pictorial and verbal format (Dunker, 1945; Maier, 1931, 1945).

In conclusion, the results demonstrated that it is possible to facilitate transfer to functional fixedness problems, which are considered a different class of problems compared to verbal insight problems. Both functional fixedness and verbal insight

problems are generally difficult to solve because of their apparent familiarity with past problems, which results in problem solving being based on a stereotypical response or assumption that is incorrect. Therefore, solution attainment for such problems is critically dependent on the problem solver abandoning their assumptions and considering a novel way of viewing the problem. The present results are promising and provide modest support for previous theories of transfer (e.g., Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901). However the challenge remains to devise generic training that promotes positive transfer to novel exemplars of verbal insight problems. This is addressed in *Chapters 6 and 7*.

CHAPTER 6

Inconsistency checking as a form of generic training for verbal insight problems

Chapters 4 and 5 reported experiments that utilised specific training to promote transfer on insight problems that were exemplars of the trained categories. Thus, Experiments 1 and 2 (reported in *Chapter 4*) provided training in heuristics that were applicable to solving verbal insight problems that either contained ambiguous words or ambiguous names. Further, the training in Experiment 3 (reported in *Chapter 5*) was also specific in that positive transfer was facilitated to a category of problems involving functional fixity. Thus, the training in these experiments aimed to improve performance on categories of verbal insight problems. Therefore, a major challenge is to devise a more generic form of training that facilitates positive transfer to not only to the problem for which it was designed but also to a range of novel verbal insight problems. In other words, training should be applicable to solving any verbal insight problem irrespective of the implicit constraint that is associated with it.

This chapter begins with a brief summary of the research that has investigated generic training (reviewed in detail in *Chapter 3, section 3.3.2*) to facilitate performance on a range of insight problems. A new form of generic training is introduced in Experiment 4 which aimed to guide problem solvers in recognising and reinterpreting their implicit assumptions that form part of a faulty representation which, in turn, impedes verbal insight problem solving.

6.1. Inconsistency checking as a mechanism for representational change

The aim of generic training is to improve performance across a range of insight problems that were associated with different constraints. This section provides

a summary of the findings of key generic training studies on insight problem solving (e.g., Ansburg & Dominowski, 2000; Chrysikou, 2006; Wicker et al., 1978), which highlight that the nature of the assumptions or constraints associated with insight problems vary so much that it is difficult to design generic training for verbal insight problem solving. In addition, past generic studies suggest that solvers must overcome such assumptions to attain insight without specifying how this can be achieved.

As reviewed earlier in *Chapter 3 (section 3.3.2)*, the earliest study cited that attempted some form of generic training for verbal insight problem solving was by Wicker et al. (1978). Two experiments were conducted which revealed that training in a reformulation strategy that encouraged problem solvers to keep changing their view of the problem to avoid making unnecessary assumptions was sufficient in raising the solution rate to 68% (Experiment 2) when compared to a visualisation condition that required participants to form a detailed visual representation of the problem. Training entailed practice with solution feedback on eight problems in Experiment 1 and three problems in Experiment 2. Although the results are promising, the study incurred methodological criticisms. That is, the reformulation strategy was not specified in detail and participants were not given practice in identifying and rejecting incorrect assumptions. Further, a list of the training and test problems was not provided which makes it difficult to assess the transfer relationship between the conditions. Finally, verbal protocols were not collected, therefore it was not known to what extent training was applied during problem solving. The above criticisms are addressed in the experiment reported in this chapter.

Ansburg and Dominowski (2000) employed a similar approach to generic training as Wicker et al. (1978) to facilitate performance on novel verbal insight problems. They found that strategic instructions in reading a problem carefully and

interpreting it in alternative ways resulted in a training effect of 48% (Experiment 2). It was also found that the number of practice problems attempted during training had little effect on test performance (Experiment 3). It was concluded that highlighting the relevance of training to problem solvers along with practice during the training phase was critical in observing a training effect (Experiment 4). As with the findings of Wicker et al, the above results also draw attention to the importance of considering alternative interpretations of a problem in facilitating representational change.

The approach to training by Chrysikou (2006) appeared to be specific to functional fixedness problems as it provided practice in generating alternative uses for items. However, training was successful in facilitating performance across a range of different types of insight problem, including verbal problems, and a facilitation effect of approximately 60% was reported. A possible explanation for this finding is that training provided participants with a general skill to avoid relying on their first interpretation which is often incorrect and that considering alternative solutions was likely to lead to the correct solution to a problem.

In summary, the training employed by Wicker et al. (1978), Ansburg and Dominowski (2000), and Chrysikou (2006) encouraged solvers to consider alternative interpretations. In particular, the training by Wicker et al. and Ansburg and Dominowski made solvers aware that the first interpretation is often incorrect and to consider alternative interpretations, thereby helping solvers to identify and correct their faulty representations. Indeed, Gestalt theory (e.g., Köhler, 1924; Wertheimer, 1945) and theories of representational change (Knoblich et al., 1999; Ohlsson, 1992) propose that insight problems are difficult because the problem specification triggers a faulty representation which leads to an impasse. The faulty representation needs to be abandoned through representational change that enables the solver to overcome the

impasse, which consequently allows the solver to resume problem solving (Ohlsson, 1992).

A possible reason why the training effects reported by the above generic studies were not closer to 100% is that the issue of the different assumptions or constraints associated with verbal insight problems was not addressed. That is, the idiosyncratic nature of the stereotypical assumptions that are triggered by verbal insight problems have largely been ignored by past training studies in this area. Hence, the training in this experiment adopted a generic approach similar to that of Wicker et al. (1978) and Ansburg and Dominowski (2000) rather than the more specific training tailored to the solution of specific categories of insight problems that was investigated in Experiments 1 to 3. Further, unlike past training generic studies that were summarised earlier, the present experiment specified a mechanism for overcoming implicit assumptions in verbal insight problem solving.

An important prerequisite for any such generic training is to first draw the trainee's attention to nature of the difficulties (Campione & Ambruster, 1985) associated with insight problems before providing the trainee with a solution strategy that can be applied to solve such problems. With regards to verbal insight problems, firstly, a problem solver must be made aware that such problems induce a stereotypical assumption that is often mistaken or incorrect. Secondly, training must provide a mechanism that allows the solver to identify the nature of these incorrect assumptions during problem solving. A suggestion for a possible mechanism is to focus participants' on identifying inconsistencies between their hypotheses or interpretations of a problem and the problem specification, as a means of assessing the plausibility of their hypotheses. Indeed, hypothesis generation is an important part of problem solving as it entails generating possible solutions that are consistent with the

problem statement (Patrick et al., 1999). This idea of inconsistency checking during problem solving is certainly not novel. In fact, inconsistency checking has been found to be a central process in hypothesis generation as it can occur spontaneously even if participants have not been asked explicitly to consider the plausibility of their hypotheses (Fisher, Gettys, Manning, Mehle, & Baca, 1983).

The mechanism of inconsistency checking used in the present experiment derives from a study by Patrick et al. (1999) in which nuclear power plant operators were trained to shift incorrect representations when trying to diagnose unusual faults by identifying the inconsistency between the actual symptoms and those associated with their incorrect hypothesised fault. This served as a cue to the problem solver to abandon their high frequency default representation and to extend their stereotypical single fault hypothesis into a more unusual multiple one. Patrick et al. suggested that such cognitive training that provided practice in resolving inconsistencies between a hypothesis and available symptoms should be generalisable to laboratory problem tasks.

It was expected that the identification of inconsistencies between a person's representation of a verbal insight problem and the problem specification should cue representational change, which in turn should make solution attainment easier, particularly if the inconsistency concerned the hypothesised constraint associated with a particular problem. For example, the Anthony and Cleopatra and Coming Home problems (see Appendix F) evoke initial representations that inappropriately include humans and night-time, respectively. By training people to make a careful comparison between the problem specification and their representation of the problem this should result in the identification of the stereotypical assumption that has been associated with their representation. Thus, by attempting to make a solver's representation of a

problem more explicit by comparing their interpretation of a problem with the problem statement, they are forced to question themselves, which has been found to be useful in other problem solving contexts such as design problem solving (e.g., Wetzstein & Hacker, 2004). However, for this method to be effective, it is important that solvers are given practice in doing this which is coupled with solution feedback (e.g., Anderson, 1983; Ansburg & Dominowski, 2000; Newell & Rosenbloom, 1981).

6.2. Experiment 4

The aim of Experiment 4 was to facilitate performance across a variety of verbal insight problems. The training was geared towards overcoming incorrect assumptions associated with verbal insight problems that lead the solver to form a faulty problem representation. In particular, participants were trained in recognising potential inconsistencies between their interpretations of a problem and the problem specification, in the hope of cueing representational change to facilitate solution. It was predicted that positive transfer would occur across all verbal insight test problems. Concurrent verbal protocols were collected to glean evidence of the nature of the hypotheses generated by participants. In addition verbal protocols might also reveal whether training was having the intended effect on the search for the solution.

6.2.1. Method

Participants

Twenty-four non-psychology students from Cardiff University took part in this study and were awarded payment for their participation⁵. Ages ranged from 18-23 years ($M = 20.50$ $SD = 1.38$).

⁵ All materials were pilot tested using four participants, which revealed that no changes were required. Therefore, datasets for these participants comprised part of the final sample.

Materials

Training programme

There were two main objectives of the training programme:-

1. To make participants' aware that an incorrect interpretation (or representation) of a problem may block the solution.
2. To provide practice at identifying inconsistencies between the problem statement and their interpretation of it.

For the awareness training, participants read through two examples of how this blocking effect might operate. For example:

Why are 1988 pennies worth more than 1983 pennies? (Sloane, 1992, p. 28)

In this problem people might assume that the numbers refer to years, which would then block the correct interpretation that the numbers refer to quantities of pennies. Therefore, 1998 pennies would be worth £19.88, which is more than £19.83.

The second part of the training involved four stages that focused on overcoming such blocking effects by providing increasingly independent practice. In the first stage, participants were presented with two different problems, each having two written interpretations that were inconsistent with the problem specification. Participants were required to identify these inconsistencies and, if they failed to do so, were prompted by the experimenter. An example of one problem was:

Archie and Ben were professional golfers and keen rivals. One day during a game, they had each scored 30 when Ben hit a bad shot. Archie immediately added 10 to his own score. Archie then hit a good shot and he had won the game. Why? (Sloane, 1992, p. 21)

Possible interpretation: 'Two friends were playing golf, they were both on 30 points, then one reached 40 points and won'.

The interpretation is inconsistent with the problem specification because it does not state that Archie and Ben were friends, nor that they were playing golf. The solution was that they were playing tennis.

The second stage of training required participants to read a similar problem to the above, to write down their own interpretation, and then to attempt identifying any inconsistencies between their interpretation and the problem specification. After this participants attempted to solve the problem. The experimenter prompted participants who were unable to complete any aspects of this. In the third stage of this training, participants were required to solve a standard problem without writing their own interpretation but being prompted, if necessary, by the experimenter. In the final stage, participants attempted to solve a problem without the experimenter prompting although solution feedback was given at the end.

Test problems

All participants attempted to solve four test problems (Anthony & Cleopatra, Bombs Away, Coming Home, and Unseen Walker) which were taken from Sloane (1992, see Appendix F). Each test problem was selected because it was hypothesised to involve a constraint that could block solution attainment.

Design

The independent variable was condition (training and no training). Participants were randomly allocated to a condition and completed all test problems. Presentation

of the test problems was counterbalanced using a four by four balanced Latin square. The dependent variable was whether the problems were solved.

Procedure

Participants read the introduction to the experiment and then were given practice in 'thinking aloud', as in previous experiments. After completing the think aloud training, participants in the no training (control) condition completed the test problems whereas participants in the training condition completed the training programme followed by the test problems whilst thinking aloud. Participants were given a five-minute time limit for each problem and if they were silent for a period of time, the experimenter again used two non-directive prompts. Verbalisations were recorded continuously. After completing each test problem, participants were required to rate how familiar they were with that problem on a 5-point scale (1 = very unfamiliar, 5 = very familiar). Participants were not given solution feedback. Finally, participants were debriefed and asked not to reveal information about the experiment to others. The duration of the experiment was approximately 40 minutes.

Qualitative categorisation and assessment of reliability

Verbal protocols were analysed to gather evidence that the inconsistency checking training procedure was responsible for any improvement in solution rates and also to identify the nature of the hypotheses generated between the training and no training conditions. A rigorous three-stage procedure was developed to identify and categorise hypotheses and to assess inter-coder reliability. The three stages were as follows:

1. Three coders (experimenter and two unrelated researchers) read the protocols and individually identified what they considered to be hypotheses/solutions for each of the four test problems completed by participants. A hypothesis or solution was defined as 'any verbalisation that attempts to answer the question posed in the problem statement'. Practice was given on a verbal protocol that was excluded from the final dataset to ensure that coders understood the task. Once all three coders had identified potential hypotheses, disagreements were discussed until a consensus was reached and a final list of agreed hypotheses was produced. A total of 382 hypotheses were initially identified by all three coders. Of these 382 hypotheses, 312 or 81.7% were identified by all three coders, 45 or 11.8% were identified by two coders, and 25 or 6.5% were identified by one coder. After discussion, 20 hypotheses were rejected, making the final total 362.

2. The second stage examined verbal protocols to determine whether participants engaged in training related behaviours. Firstly, two coders (the experimenter and a different coder to those in the previous stages) used the following criteria to identify any segment in the protocols that demonstrated that:

a) a participant re-read or paraphrased some part or the whole of the problem statement.

b) a participant explicitly questioned something within the problem statement.

c) a participant self-questioned their hypotheses i.e., self-reflection.

A total of 394 segments across the verbal protocols were initially identified of which 42 or 11% were identified by one coder and 352 or 89% was identified by both coders. After discussion, 12 segments were rejected, making the final

total 382. Next, the coders categorised the 382 segments into one of the above three types of training related behaviours. The Perrault and Leigh (1989) reliability index was 0.92 between the two raters, which was acceptably high (Gremler, 2004; Krippendorf, 1980).

3. The final stage involved the same coders, as in stage 1, who independently categorised the agreed 362 hypotheses into one of three categories: inconsistent, incorrect other, or correct. Inconsistent hypotheses were so called because they contradicted some information given in the problem statement. Thus, for example, in 'Bombs Away' the hypothesis that 'the release mechanism failed' is inconsistent because it was contradicted by the problem statement that 'the plane was in perfect condition and everything on it worked properly'. Incorrect other hypotheses were so called as they did not contradict any information in the problem specification but nonetheless were wrong. For example, in 'Anthony and Cleopatra', an incorrect other hypothesis was that death was due to natural causes. Correct hypotheses referred to the solution of the problem. Of the 362 hypotheses, it was found that 334 or 92.3% were categorised by all three coders using the same categories and the remaining 28 or 7.7% were categorised by two coders. The Perreault and Leigh (1989) reliability indices for the three pairs of coders were 0.95, 0.97, and 0.97, respectively, which were acceptably high (Gremler, 2004; Krippendorf, 1980).

6.2.2. Results and Discussion

No participants were familiar with test problems; therefore the final sample comprised 24 participants. An exploratory data analysis did not reveal any violations

of homogeneity and normality of the data, therefore a t-test was conducted, which revealed that more problems were solved in the training ($M = 1.58, SD = 1.00$) than in the no training condition ($M = .83, SD = 1.03$), $t(22) = 1.81, p < .05$ (1-tailed). Hence, training almost doubled the solution rate from the baseline of 21% (no training) to 40% in the training condition, thus illustrating a facilitation of 19% (see Figure 6.1).

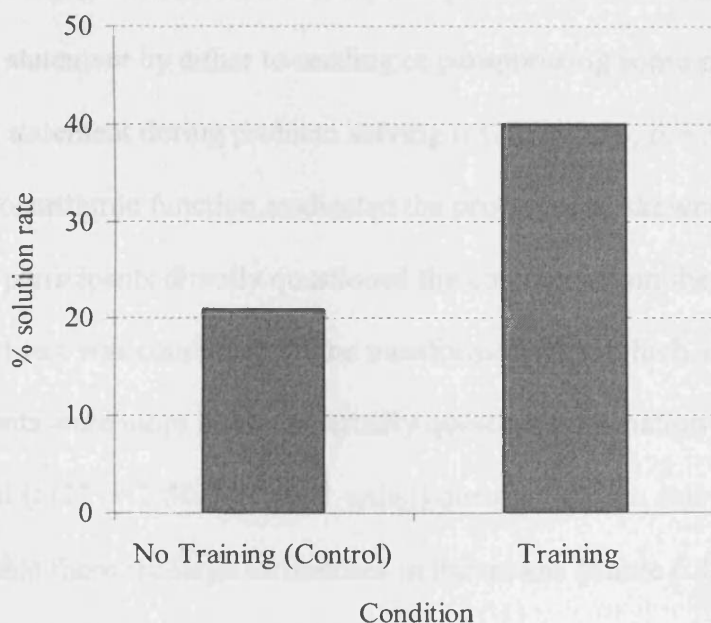


Figure 6.1. Percentage solution rate by condition

In order to confirm that the above transfer effect was as a result of the training procedures, the verbal protocols were analysed firstly to identify whether participants in the training condition engaged in training related behaviours and secondly, to determine if there were differences in the types of hypotheses generated by participants between the two conditions.

Analysis of verbal protocols

It was expected that participants in the training condition were more likely to refer to the problem specification and/or to question some of part of the specification as a result of training. Another possible consequence of training is that it may have prompted participants to engage in more reflective thinking i.e., self-reflection that entailed questioning their hypotheses.

A t-test revealed that trained participants were more likely to refer to the problem statement by either re-reading or paraphrasing some parts or the whole of problem statement during problem solving ($t(22) = 2.56, p = .009$, 1-tailed, Table 6.1). A logarithmic function eradicated the problems of skewness on data assessing whether participants directly questioned the content within the problem statement. Hence a t-test was conducted on the transformed data, which also revealed that trained participants were more likely to verbally question information in the problem statement ($t(22) = 2.50, p = .01$, 1-tailed) during problem solving (Table 6.1). It can be seen that there are large differences in the means (Table 6.1.) between the two conditions for the first two training related behaviours, which provides support for the use of the training procedure by trained participants. Table 6.2 provides examples of questions participants in the training condition verbalised during problem solving for each of the four test problems.

Finally, due to violations of skewness on self-reflection data which could not be eradicated with a logarithmic transformation, a Mann-Whitney U test was conducted which revealed no differences between conditions ($z = 3.98, p = .75$) in the engagement of self-reflection during problem solving (Table 6.1). The mean rank was 12.04 for no training condition and 12.96 for training condition. This suggests that there were no differences between conditions in participants explicitly questioning

their hypotheses during problem solving. In fact, the results indicate that very few hypotheses were generated that illustrated self-reflection. Examples of hypotheses that suggested trained participants had engaged in self-reflection for the Anthony and Cleopatra include ‘we assume that it is the famous Anthony and Cleopatra as they are in Egypt but they could be any old Anthony and Cleopatra’ and ‘we assume that the Anthony and Cleopatra relate to the death that happened thousand of years ago but may be I shouldn’t have assumed that’. This null finding is not surprising as self-reflection is a metacognitive activity which was not covered by the training in this experiment.

Table 6.1. Types of training related behaviours during problem solving

Types of training related behaviour	No Training (Control)		Training	
	M	SD	M	SD
1. Re-read or paraphrased some part or whole of the problem statement	7.75	5.82	15.67	8.98
2. Directly questioned something within the problem statement	1.83	1.95	5.83	6.79
3. Self-reflection on hypotheses	0.33	0.65	0.42	0.67

Table 6.2. Examples of questions verbalised by trained participants that were related to the problem statement

Test Problem	Examples of questions asked by participants
Anthony & Cleopatra	<ol style="list-style-type: none"> 1. How else could they [Anthony and Cleopatra] die? 2. What was in the bowl? 3. Did they [Anthony and Cleopatra] die in the villa?
Coming Home	<ol style="list-style-type: none"> 1. Why is there a car that doesn't have its headlights on racing down the road? 2. It doesn't say he has been drinking alcohol? 3. How did the driver of the car manage to see him?
Unseen Walker	<ol style="list-style-type: none"> 1. Why is he walking several miles on a busy Friday afternoon? 2. Is he walking on the pavement? 3. Why didn't anyone see him?
Bombs Away	<ol style="list-style-type: none"> 1. It doesn't say that it [the plane] was flying? 2. Where did the [the bombs] fall from? 3. Did it [the problem statement] say they [the plane] were carrying bombs?

The above findings demonstrate that what was learnt during training was indeed applied during problem solving as participants engaged in the types of behaviours that they were trained to carry out.

Next, verbal protocols were analysed to determine whether there were differences between conditions in the types of hypotheses participants had generated during problem solving. It was expected that participants in the control condition were more likely to generate hypotheses that contradicted some information available in the problem statement. Further, as a result of training, participants in the training condition were more likely to avoid or repress a hypothesis that was inconsistent with the problem statement and thus were less likely to generate inconsistent hypotheses but were more likely to generate 'other' hypotheses that were neither inconsistent nor correct. Due to violations of skewness on the number of inconsistent hypotheses generated in the no training condition, a logarithmic function was conducted which served to eradicate this problem. Contrary to predictions, a t-test revealed no

significant difference between the no training ($M = 2.75$, $SD = 2.09$) and training ($M = 3.33$, $SD = 1.67$) conditions in the generation of inconsistent hypotheses, ($t(22) = 1.212$, $p = .119$, 1-tailed). Further, no significant difference was found between the no training ($M = 11.92$, $SD = 3.77$) and training ($M = 9.75$, $SD = 4.12$) conditions in the generation of 'incorrect other' hypotheses, ($t(22) = 1.344$, $p = .096$, 1-tailed).

Possible reasons for the above null findings include that although the coding was highly reliable, the verbalisations did not fully reveal participants' thoughts during problem solving. Alternatively, the results may have been due to lack of power.

Differences in the approach to problem solving

Verbal protocols illustrated differences in the approach to problem solving between conditions. The following excerpt is representative of the difficulties participants in the no training had in solving the Coming Home problem in which the problem statement triggered the incorrect assumption that it is night-time, thus preventing the solver to realise that the driver was able to see the man because it was day-time.

"In this question, we know that the man had quite a few drinks and was on his way home...I assume that the driver might have some instinct that made him aware there might be a person on the road. Especially if the driver is a very careful person and at the time he was driving on the road, it was so dark, most probably he was not speeding as well. When it came to getting near to the man who was drunk, I think the driver must have sensed a feeling of human trace on the road...It made him stop immediately when there was danger....And because it was all in the dark and because the man was drunk, he might be walking from one end to the other, and somehow he got to see the movement,

shadow. I would think that that would make the driver of the car aware that there might be an animal or a car and it makes him alert...He might have good eyesight as well so that makes things easier for him rather than a person with blurred vision. Or just simply having the music loud...would contribute to him being able to see the drinking man.”

This excerpt is very different to the following excerpt from the training condition which demonstrates how the inconsistency checking mechanism was applied to attain solution in the Unseen Walker problem:

“No moon light but there could have been, there’s possibility of starlight.....But there could have been another source of illumination such as a house or there was starlight. Or there could have been some other method of seeing him such as a...some cars have got infrared filters in the windscreens but I think that’s quite unlikely. It says he’s been out drinking but it doesn’t imply alcohol.....He was walking down the middle of a deserted country road...which makes me think of narrow and windy so...If it was windy, the driver might not have seen him until the last minute because of a bend rather than because of everything being black...It doesn’t say it’s night-time though. So I think he’s walking in the middle of the day and there was a bend and whoever was dressed in black, the driver saw him anyway.”

The participant initially made the incorrect assumption that it was night-time and thus hypothesised alternative light sources that could explain how the driver saw the man. Then it was correctly identified that although the man had been drinking, he may not have been drinking alcohol which is often associated with night-time. The solution was attained shortly afterwards.

In conclusion, the results of the present experiment suggest that this form of generic training can be successful in raising people's awareness of stereotypical assumptions that hinder verbal insight problem solving, and that providing appropriate practice at identifying these wrong assumptions through an inconsistency checking mechanism can facilitate solutions to novel problems.

6.3. Conclusions

The approach to training in Experiment 4 was generic in that it was designed to facilitate performance on any verbal insight problem in spite of the type of assumption or constraint associated with that problem. Hence, the aim was to mitigate the difficulties encountered in solving any verbal insight problem by helping solvers to identify the assumption or constraint that blocks the solution path. It is of importance to note that the generic training in this Experiment 4 was different to the training of Wicker et al. (1978) and Chrysikou (2006). Both Wicker et al.'s reformulation strategy and Chrysikou's goal-derived categorisation training emphasised that the solver must overcome assumptions to attain solution although how this can be achieved was unspecified. This criticism was addressed by the training in Experiment 4 in which an explicit cognitively related process was provided that attempted to identify problem constraints. It was expected that the solver would initially generate an incorrect representation of a problem that was based on a stereotypical assumption that was triggered by the problem specification. Further, this incorrect representation would contain unhelpful additional information that was not detailed in the problem specification. Training was based on the idea that the solver could be trained to recognise this additional information by explicitly comparing their

interpretations with the problem specification, which would, in turn, facilitate recognition of the problem constraint or assumption and consequently the solution.

This training was clearly successful and this was substantiated by analysis of the verbal protocol analyses. This found that the participants in the training condition referred more frequently back to the problem statement and also explicitly questioned some information in the problem statement during problem solving. Therefore it is reasonable to conclude that the training effect was due to this trained procedure, which was termed inconsistency checking. It was unsurprising that there was no evidence of more reflective thinking in the trained condition as the training did not directly address this.

However the solution score was only at 40% even in the trained condition and so training clearly still needs to be improved. Indeed there was no significant difference between the training and control conditions in the generation of hypotheses that were inconsistent with the problem statement. One possible explanation for this finding is that the cognitive process of inconsistency checking was difficult for solvers to perform because it requires assessment of their representation of a problem, which may have increased cognitive load and demands on working memory. Therefore, a replication of this experiment would benefit an extra condition in which an external memory aid, such as a pen and paper to write down interpretations during problem solving, is utilised which may further make a person's representation more explicit and increase recognition of a discrepancy.

Generally the verbal protocol data provided some indirect support for Representational Change theories (Knoblich et al., 1999, 2001; Ohlsson, 1992) that suggest that a faulty representation blocks the solution path. Indeed, the test problems evoked an implicit constraint that affected hypothesis generation. In other words, once

the stereotypical assumption was triggered, it had a strong and pervasive effect during problem solving, particularly in the no training condition.

In conclusion, the results demonstrated that it is possible to facilitate performance on novel verbal insight problems through generic training, as reported by previous studies (e.g., Ansburg & Dominowski, 2000; Chrysikou, 2006; Wicker et al., 1978). The inconsistency checking mechanism embedded within problem solving aimed to identify faulty representations, and therefore served as a novel method for demonstrating restructuring within verbal insight problem solving. However, the results provided modest support for previous theoretical formulations of transfer (e.g., Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901). A suggestion to improve the training effect is to further break down the use of the inconsistency checking mechanism by directing participants to scrutinise interpretations of different parts of the problem statement, thus increasing the likelihood of identifying a faulty representation during verbal insight problem solving. Thus, an iterative approach to problem solving, which is similar to that used to solve functional fixedness problems in Experiment 3, is explored in Experiment 5, reported in *Chapter 7*.

CHAPTER 7

Enhanced training in inconsistency checking for verbal insight problems

The generic training employed in Experiment 4, in the preceding chapter, introduced an inconsistency checking mechanism as a way of operationalising restructuring in verbal insight problem solving. A training score of 40% was reported. The aim of the present experiment was to raise this effect by implementing an adapted version of the inconsistency checking training that was utilised in Experiment 4. In addition, an additional training condition was employed to determine whether instructing participants to explain and justify their solutions induced reflective thinking that served to improve performance.

7.1. An iterative process of inconsistency checking

The training in inconsistency checking was adapted for the present experiment in order to improve performance on verbal insight problems by encouraging participants to adopt a more systematic approach to using the mechanism to help them identify and correct the faulty representation to attain insight. The training in Experiment 5 was designed to encourage participants to systematically select different parts of a problem and to identify any inconsistencies between their interpretations concerning that part of the problem and the problem specification. In addition, once participants had generated an interpretation that was consistent with the problem specification, they were required to consider whether it could be used to solve the problem. Thus, an iterative process was introduced that explicitly encouraged participants to be more exhaustive in examining the problem specification and their interpretations thereby increasing the chance of identifying their incorrect

representation. This iterative approach to assessing the problem statement is similar to the training method that was implemented in Experiment 3, reported in *Chapter 5*, which was successful in facilitating performance on functional fixedness verbal insight problems for which the solution required an object to be used in a novel way. That is, participants were trained to solve functional fixedness problems by iteratively selecting an object mentioned in the problem specification, to consider all possible functions for that object and to systematically consider the plausibility of each function for an object as a possible solution to a problem.

Another issue that was addressed in the present experiment concerned the number of test problems used. It has been mentioned throughout this thesis that one of the problems for designing training for insight problem solving is that there is much variability between problems in terms of their associated assumption or constraint. Hence, the number of test problems was raised from four to seven in Experiment 5 in an attempt to better gauge the efficacy of inconsistency checking as method for inducing representational change.

7.2. Reflective verbalisations during problem solving

An alternative approach to the concurrent verbalisations used in this thesis, after Ericsson and Simon (1993) is 'reflective verbalisation', which requires a problem solver to critically evaluate their thinking by explaining themselves, and thus can give rise to reflection. Reflection of the solution searching process encourages relevant metacognitive activity and thus directs the problem solver to focus on aspects of the problem that are critical to solution attainment (Ansburg & Dominowski, 2000; Dominowski, 1990). For the purpose of training, it has been reported that asking participants to explain their thinking is in fact beneficial to problem solving (Ahlum-

Heath & DiVesta, 1986; Berry & Broadbent, 1984, 1987; Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010). This section briefly discusses research that has successfully induced reflective thinking, particularly during analytical problem solving, and then makes suggestions for how reflection can be trained and investigated with respect to verbal insight problem solving.

Wetzstein and Hacker (2004) cited a case study which compared verbal protocols of two participants who were instructed to think aloud. It was observed that one participant frequently asked questions, stated hypotheses and conclusions, and attempted to explain, evaluate and justify their behaviour whereas the other participant tended to be descriptive in their verbalisations. The former participant demonstrated greater problem solving performance. This suggested that the mere act of self question-answering dialogues or discussions during problem solving appears to encourage a solver to reflect on their thinking which can have positive benefits.

Support for verbalisations that encourage reflection has been reported for the well-defined Tower of Hanoi task (Ahlum-Heath & DiVesta, 1986). Ahlum-Heath and DiVesta found that when naïve problem solvers were required to state and justify each move they made before it was made during practice on two, three, four and five disk versions of the Tower of Hanoi problem, performance was consequently improved on the test problem involving six disks. It was also found that participants who were given practice without having to justify their moves were more likely to make excess moves than those in the justification condition. These results were interpreted to suggest that instructing solvers to state and justify their moves encouraged solvers to consider each move separately, to evaluate alternative moves for a given problem state and to reject moves that were unfavourable. Thus, the

process explicitly encouraged solvers to employ means-end analysis to explicate moves to reach the goal state.

Support for the positive effects of reflective verbalisations has also been reported for search tasks (Berry & Broadbent, 1984, 1987) and design tasks (Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010) which are complex and ill-defined. For example, Berry and Broadbent (1984) investigated three different methods for improving search performance on a diagnostic task. It was reported that a) verbal instruction on procedures and b) giving participants instructions to so say aloud their reasons for their actions were both ineffective in changing performance. However, when verbal instructions were combined with asking participants to justify their actions aloud, performance was improved. It was also reported that positive effects for the latter method were not only observed on the specific task that was trained, but also on a different task that required the same procedures. Thus, positive transfer was observed to a different context. This finding suggests that simply instructing participants to justify (or explain) their thinking during problem solving is sufficient to encourage reflection, which in turn, aids performance.

More recently, Winckelmann and Hacker (2010) demonstrated that question-based reflection improved the quality of design solutions in students and experts. In designing an artefact, the experimenter asked questions that encouraged participants to explain, justify and to evaluate their finished design. That is, the questions aimed to stimulate reflection on the solutions generated that participants could subsequently modify or revise. Examples of questions included “How does the proposed solution fulfil the required functions?” and ‘Which alternative solutions for the functions may exist?’. All groups (experts versus novices) demonstrated an improvement on solutions when compared to a control. Moreover, of the experts, those with less

experience showed better performance in comparison to those with higher experience, thus supporting the beneficial effects of reflective verbalisations.

The above research indicates that reflection during problem solving has been induced via two different methods. That is, by asking participants to explain their thinking which may also result in the participant engaging in self question-answer dialogues, or by allowing participants to ask the experimenter questions during problem solving. A more notable observation is that research investigating the effects of reflective verbalisations has focussed on non-insight problems which, unlike insight problems, do not require restructuring to reach solution (Ohlsson, 1992; Sternberg & Davidson, 1995; Weisberg, 1995). In other words, solution attainment for such problems is not dependent on the abandonment of an initial misrepresentation of the problem. Ormerod and Ball (2007) recently noted that studies that measure cognitive activity through the collection of verbal protocols often make no attempt to engage participants in reflecting about their performance. The qualitative analyses in Experiment 4 revealed that participants in the training condition were more likely to explicitly question information in the problem statements during problem solving although no evidence was found for a greater engagement in reflective thinking in the training condition. Therefore, an additional aim of Experiment 5 was to assess whether instructing participants to explain and justify their thinking during practice promotes reflective thinking, which in turn facilitates performance on verbal insight problems. No research has investigated this approach using verbal insight problems.

7.3. Experiment 5

The present experiment comprised of two training conditions (Enhanced Inconsistency Checking and Explanation and Justification conditions) and two control

conditions (Practice and No Training conditions, see Table 7.1) The Enhanced Inconsistency Checking training condition completed training in inconsistency checking on three practice problems that encouraged a systematic approach to identifying inconsistencies between interpretations of different parts of a problem and the problem specification. The Explanation and Justification condition also completed the same three practice problems as the first training condition. The difference between the two training conditions was that in the latter condition participants were required to explain and justify their thinking during practice whereas the former required participants to verbalise their thinking only. Solution feedback was provided during practice for both conditions.

The inclusion of the first control condition was based on research evidence suggesting that training incorporating practice coupled with feedback facilitates learning (Anderson, 1983), which is an important determinant of positive transfer not only in insight problem solving (Ansburg & Dominowski, 2000) but also in other domains including improving and maintaining performance on vehicle-related jobs (Komaki, Heinzmann, & Lawson, 1980), training of sprinters (Howell, 1956), and performance on perceptual tasks (Ludwig & Pieper, 1998). Thus, participants in the Practice condition also completed three practice problems and solution feedback was provided. Finally, participants in the No Training condition completed the same test problems as all the other conditions but did not receive practice beforehand. Table 7.1 provides a table of the design for Experiment 5 summarising the four conditions.

Table 7.1. Table of design for Experiment 5

Condition	Training/Practice Phase	Testing Phase
No Training (Control)	No training or practice given.	Completed seven test problems
Practice	Completed three practice problems and solution feedback was provided.	Completed seven test problems
Explanation & Justification	<ol style="list-style-type: none"> 1. Completed three practice problems and solution feedback was provided. 2. Instructed to explain and justify thoughts during the training. 	Completed seven test problems
Enhanced Inconsistency Checking Training	<ol style="list-style-type: none"> 1. Completed inconsistency checking training on three practice problems and solution feedback was provided. 2. Instructed to verbalise thoughts during the training 	Completed seven test problems

It was predicted that the Enhanced Inconsistency Checking Training would facilitate the greatest performance followed by the Explanation and Justification condition, then the Practice condition with the No Training condition last. This is because the Enhanced Inconsistency Checking training introduced an exhaustive search process in considering discrepancies between several interpretations and the problem specification, thereby increasing the likelihood of recognising and changing the faulty representation that is associated with a given problem. The next best performance should be observed for the Explanation and Justification condition as it encourages participants to explicitly examine their thoughts, thereby indirectly encouraging careful scrutiny of solutions, which in turn may draw participants' attention to the faulty problem representation. It was expected that performance should be significantly higher for the former condition because training is explicitly geared towards helping participants identify the misrepresentation of a problem unlike the latter condition. Thirdly, solution rate should be higher for the Practice condition than the No Training condition because practice with solution feedback on problems

similar to test problems should facilitate performance (Anderson, 1983) in comparison to no practice as is the case in the No Training condition.

7.3.1. Method

Participants

Forty-eight psychology students from Cardiff University took part in this study and were awarded payment for their participation⁶. Ages ranged between 18 to 35 years ($M = 19.27$, $SD = 2.46$).

Materials

Training Programme

The Enhanced Inconsistency Checking training programme was similar to that used in Experiment 4 (see Appendix G). The differences were that participants were instructed to select a part of a problem and to identify any inconsistencies between their interpretations and the problem specification. Further, when they had generated a consistent interpretation, they were required to consider the plausibility that their interpretation could lead to a possible solution to a problem. This iterative process was repeated until the solution was attained. An explanation in how to use the training was illustrated using the Sid Shady problem and then guided practice was given on the Barney Dribble problem followed by practice under test conditions on the Coffee problem (see Appendix H).

Explanation and Justification

Participants in this condition were required to explain and justify their

⁶ All materials were pilot tested using 16 participants, which revealed that four-minutes per problem were sufficient for participants to read test problems and to problem solve. As no changes were required, datasets for these participants comprised part of the final sample.

solutions during completion of practice problems. The practice problems (Sid Shady, Barney Dribble, Coffee problem) were the same as those used for the above condition.

Practice and No Training

Participants in the Practice condition completed the three practice problems as both of the above conditions before the completion of test problems whereas the No Training condition completed test problems without receiving any practice.

Test Problems

Seven test problems were selected from Ansborg and Dominowski (2000). These were: Pear tree, Dr Apple, Train, Directory, Antique coin, Professor Bumble and Light (see Appendix H).

Design

The independent variable was condition with four levels (Enhanced Inconsistency Checking Training, Explanation and Justification, Practice and No Training). Participants were randomly allocated to a condition and completed all seven test problems that were randomly presented. The dependent variable was whether the problems were solved.

Procedure

Participants read the introduction to the experiment and then were given practice in 'thinking aloud', as in previous experiments. After completing the think aloud training, participants in the No Training (control) condition completed the test problems. Participants in the Practice and Explanation and Justification conditions

first completed practice problems except those in the latter condition were required to explain and justify their solutions during the practice phase. Participants in the enhanced training condition completed the training programme prior to completing test problems. Participants were given a four-minute time limit for each problem and if they were silent for a period of time, the experimenter again used two non-directive prompts ('What are you thinking?' and 'Please keep talking'). Verbalisations were recorded continuously. After completing each test problem, participants rated their familiarity with the problem on a 5-point scale (1 = very unfamiliar, 5 = very familiar). Participants were not given solution feedback. Finally, participants were debriefed and asked not to reveal information about the experiment to others. The duration of the experiment was approximately one hour.

Qualitative categorisation and assessment of reliability

As in Experiment 4, verbal protocols were analysed to gather evidence that the inconsistency checking training procedure was responsible for any improvement in solution rates and also to identify the nature of the hypotheses generated between the conditions. A similar procedure to the one adopted in Experiment 4 was employed in coding verbal protocols. The results of the coding are reported below.

1. Three coders (experimenter and two unrelated researchers) read the protocols and individually identified each hypothesis/solution for each test problem. A total of 816 hypotheses were initially identified by all three coders. Of these 816 hypotheses, 786 or 96.3% were identified by all three coders, 6 or 0.7% were identified by two coders, and 24 or 2.9% were identified by one coder. After discussion, 27 hypotheses were rejected, making the final total 789.

2. The second stage involved the same coders independently categorising the agreed 789 hypotheses into one of three categories (inconsistent, incorrect other, or correct), as was the case in Experiment 4. Thus, inconsistent hypotheses were those that contradicted some information given in the problem statement. Incorrect other hypotheses were those that did not contradict any information in the problem specification but nonetheless were wrong. Correct hypotheses referred to the solution of the problem. Of the 789 hypotheses, it was found that 786 or 99.6% were categorised by all three coders using the same categories, 2 or 0.3% were categorised by two coders, whereas 1 or 0.1% were categorised by only one coder. The Perreault and Leigh (1989) reliability indices for all three pairs of coders were 0.99 respectively, which were acceptably high (Gremier, 2004; Krippendorf, 1980).

7.3.2. Results and Discussion

It was predicted that transfer would be positive and that the Enhanced Inconsistency Checking training condition would facilitate the highest solution rate because the approach was much more systematic as well as exhaustive in comparison to the inconsistency checking training in Experiment 4. The next best performance should be observed for the Explanation and Justification, Practice and No Training conditions, respectively.

An exploratory data analysis did not reveal any violations of homogeneity and normality of the data, and therefore a one-way ANVOA was conducted on total frequency of solution, which revealed significant differences between the conditions, ($F(1, 44) = .61, MSE = .04, p < .001$). Bonferroni adjustment for multiple comparisons revealed that the Enhanced Inconsistency Checking Training condition

performed significantly better than the No Training (control) ($p < 0.001$) and Practice ($p < 0.01$) conditions (Table 7.2). Thus, practice with solution feedback is not sufficient in facilitating representational change as participants require training in how to identify the incorrect representation. The results also revealed that the Explanation and Justification condition also performed significantly better than the No Training (control) condition ($p < 0.01$), thus suggesting that encouraging participants to reflect on their thinking during problem solving facilitated performance on verbal insight problems. Counter to predictions, there was no difference between the Enhanced Inconsistency Checking and Explanation and Justification conditions. This null finding may possibly be due to lack of power.

Table 7.2. Solution rates across conditions (out of seven problems)

Condition	Mean	SD
No Training (Control)	2.08	1.68
Practice	2.75	1.14
Explanation & Justification	4.08	1.51
Enhanced Inconsistency Checking Training	4.75	1.14

Figure 7.1 provides the percentage solution rate among the training conditions. The solution rate for the No Training condition was 30%, which served as a baseline measure. It can be seen that the Enhanced Inconsistency Checking Training raised solution to 68% (i.e., 38% facilitation), whereas the Explanation and Justification condition raised solution to 58% (i.e., 28% facilitation). Further, a t-test was conducted on proportion solution data after checking assumptions of homogeneity and normality, which revealed that performance in the Enhanced Inconsistency Checking condition ($M = .68$, $SD = .16$) was significantly better than performance in the inconsistency checking training in Experiment 4 ($M = .40$, $SD = .25$), ($t(22) = 3.23$, $p < .01$, 1-tailed). This further validates that introducing an iterative process of

inconsistency checking in the present experiment was indeed beneficial. In other words, unlike Experiment 4, participants in Experiment 5 were more exhaustive in searching for and in identifying inconsistencies between their interpretations and the problem statement, which in turn increased the likelihood of recognising an incorrect representation, which is critical to achieving insight.

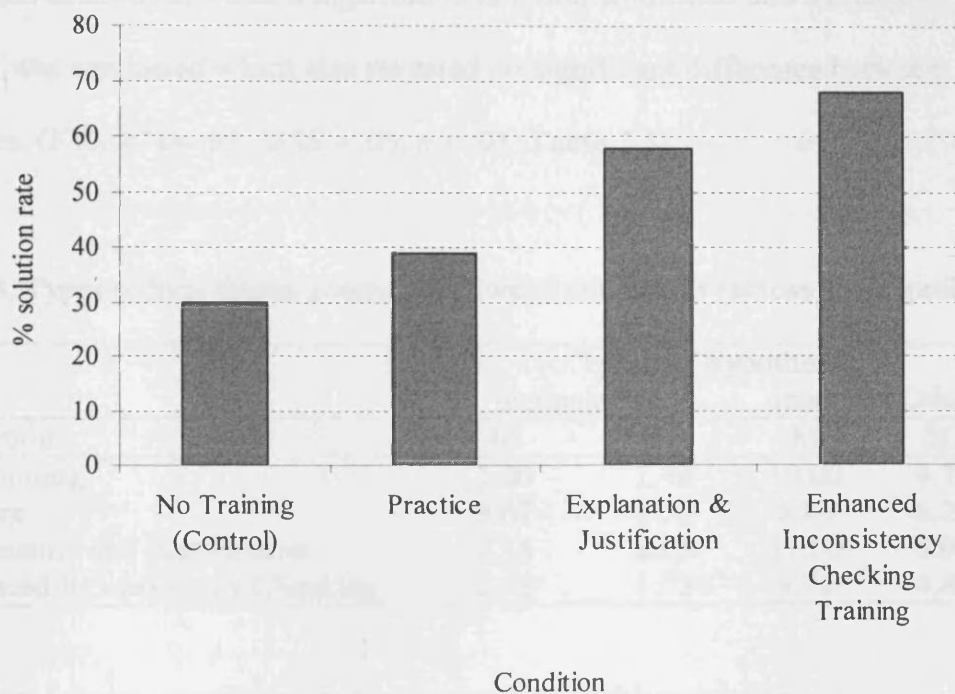


Figure 7.1. Percentage solution rate by condition

Analysis of inconsistency identification

As in Experiment 4, it was expected that the enhanced training in inconsistency checking would reduce the number of inconsistent hypotheses generated by encouraging participants to check the consistency between their representation/interpretation of the problem and the information within the problem specification. Thus, as a result of training, participants may avoid or repress inconsistent hypotheses that contradicted some information in the problem

specification) and instead generate more incorrect other hypotheses which are neither inconsistent nor incorrect. An exploratory data analysis did not reveal violations of homogeneity and normality on the inconsistent hypotheses data, therefore a one-way ANOVA was conducted, which revealed no difference between conditions in the generation of inconsistent hypotheses, ($F(3, 47) = .94, MSE = 2.72, p > .05$, Table 7.3). The data on 'incorrect other' hypotheses was skewed across the practice and explanation conditions, which a logarithmic function eradicated and a one-way ANOVA was conducted which also revealed no significant difference between conditions, ($F(3, 47) = .65, MSE = .03, p > .05$, Table 7.3).

Table 7.3. Types of hypotheses generated between conditions (across seven problems)

Condition	Types of hypotheses			
	Inconsistent		Incorrect Other	
	M	SD	M	SD
No Training	3.00	1.48	10.00	4.75
Practice	3.67	1.50	5.96	6.20
Explanation and Justification	2.75	2.05	11.00	5.03
Enhanced Inconsistency Checking	2.58	1.73	8.75	4.83

Differences between conditions in the approach to problem solving

The verbal protocols provided additional evidence that enhanced training in inconsistency checking was applied during problem solving. Comments such as 'It doesn't say in the problem...so it's not consistent' and 'now I'll go through the problem again' suggested that participants' were checking for inconsistencies between their interpretations and the problem statement. For example, the following excerpt demonstrates how a trained participant iteratively used the inconsistency mechanism to reach the solution for the Light problem in which participants needed to overcome the assumption that it was not night-time.

“Um.....turned off the light in his bedroom and managed to get into bed before the room was dark. I'm not really sure if the room is his bedroom, so he could possibly be getting into bed in another room.....If the bed is ten feet from the light switch.....I'm not ever so sure that's got anything to do with it at the minute. A young boy turned off the light in his bedroom and managed to get into bed before the room was dark..... possibly there was a light source elsewhere, outside.....Got into bed before it was dark..... The light may have possibly had a setting on it that stayed on for a few seconds after.....The bed is ten feet from the light switch. I'm not sure what or how that ten feet will be-- maybe he could turn it off from there I'm not sure, whilst he's in bed...Is it really *consistent* with the problem? It is possible that the light has some sort of setting if it is the main light. I'm assuming that it was the main one....Maybe he went.....oh no there were no other contraptions.....so that the possibility of turning it off from the bed probably isn't *consistent* with the problem.....Turned off the light in his bedroom. At the moment I'm thinking.....a solution...probably something to do with the light having a setting..... If the bed is ten feet, it's quite far I think isn't it from the light switch...and he used no wire....So a young boy, he turned off the light in his bedroom.....before the.....unless the bed's in another room but then.... he managed to get into bed before the room was dark.....possibly we shouldn't assume that it's the night time.....possible solution yeah it might be that. He may have turned off the light in his bedroom and it wouldn't have made any difference if it was daytime outside and therefore he managed to get into bed before the room was dark. That's possibly *consistent* with the problem so if there was no wires, strings or contraptions

and he can't reach the light switch.....That may be a young boy turned off the light in his bedroom and managed to get into bed before the room was dark...I think the answer is that the boy did turn off the light in his bedroom and got into bed before the room was dark after it was still sort of daytime and there was light coming possibly from a window."

Below is an excerpt which demonstrates how a participant in the Explanation and Justification condition reached the solution for the Light problem.

"Well he wouldn't have been able to reach the light from his bed so he would have had to have got out of bed to turn it off and he didn't have any contraptions. He'd have more than likely had to turn the light off himself.....but he turned the light off and would still be able to get back into bed before the room was dark. Um so.....but he didn't use any wires or strings or other contraptions.....So he would probably have had to physically turn the light off.....but if he turned the light off then you would have thought that the room would go into instant darkness unless.....it was a dimming light bulb....*But if it's not night time then it will be light outside so even when he switches the light off the light from the window would come into the room so he'd still be able to see in which case he would be able to find his way back to bed as it wouldn't be dark.....in which case he doesn't have to use any wires, or strings or other contraptions because he could have actually physically walk to the light switch to turn it off and walk back and be able to see where he was going because of the light shining in through the window.*"

It can be seen that the participant spent quite some time explaining what they were thinking and justifying their solutions, in particular, towards the end when they had reached the correct solution to the problem.

In conclusion, the results of Experiment 5 provide further support for generic training in raising people's awareness of stereotypical assumptions that hinder verbal insight problem solving. In particular, the results support those reported in Experiment 4 and suggest that providing people with detailed practice in identifying their assumptions and incorrect interpretations in a systematic manner through an inconsistency checking mechanism is successful in producing a strong training effect. In addition, a novel finding was that asking participants to explain and justify their solutions had a positive effect on problem solving performance.

7.4. General Discussion

The aim of Experiment 5 was to adapt the inconsistency checking training that was employed in Experiment 4 in the hope of raising performance on verbal insight problems. It was expected that an iterative approach to inconsistency checking training that encouraged participants to be more systematic in comparing their interpretations concerning different parts of a problem with the problem specification would increase the likelihood of recognising the faulty representation for a given problem. It was found that the Enhanced Inconsistency Checking training raised the solution rate from 40% (Experiment 4) to 68% (Experiment 5), thus illustrating a facilitation of 28%. This result lends greater support for inconsistency checking as a mechanism for inducing representational change in verbal insight problem solving. It is also possible that this result was due to reflective thinking that was indirectly promoted by the trained mechanism. Indeed the analyses of the training related behaviours in Experiment 4 suggested that participants did engage in some form of reflective thinking by verbalising explicit questions that concerned the information in the problem specification.

As was the case in Experiment 4, analysis of the frequency with which inconsistent hypotheses were generated by participants in Experiment 5 did not reveal a difference due to training. One possible explanation for the null results is that it was possible that although the coding of the verbal protocols was highly reliable, the verbalisations may not have fully revealed these cognitive related activities.

It was unclear from the results in Experiment 4 whether the training effect was due a) to inconsistency checking alone; b) the reflective thinking induced by the mechanism as evidenced by the questions verbalised by participants; or c) due to a combination of both inconsistency checking and reflective thinking. Research has demonstrated that reflective verbalisations encourage problem solvers to critically evaluate their solutions, which in turn direct the solver to important aspects of a problem that are critical to solution (Ahlum-Heath & DiVesta, 1986; Berry & Broadbent, 1984, 1987; Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010). In particular, instructing participants to explain or justify their solutions has been demonstrated to aid problem solving in search tasks (Berry & Broadbent, 1984, 1987) and design tasks (Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010) although no research has been conducted using verbal insight problems. Therefore, an additional condition was employed in the present Experiment in which participants were instructed to explain and justify their thoughts during training. The results revealed a 58% training score for this condition, thus a facilitation of 28% was observed when compared to the No Training condition. This suggests that the practice in explaining and justifying solutions without training in how to overcome incorrect representations appeared to improve performance on verbal insight problems, which is consistent with research that demonstrate the benefits of using reflective verbalisations (Ahlum-Heath & DiVesta, 1986; Berry & Broadbent, 1984, 1987;

Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010). A possible explanation for this finding is that by asking participants to explain their responses indirectly promoted inconsistency checking, thereby increasing the likelihood of drawing participants' attention to information that they may not have noticed initially that is critical to identifying faulty representations.

A significant difference in performance was expected between the Enhanced Inconsistency Checking and Explanation and Justification conditions but this was not the case. It could be argued that, at a more general level, that both conditions give rise to some reflective thinking that drew participants' attention to critical information related to an incorrect representation. However, the key difference between the two conditions was that in the former a detailed mechanism was specified that could essentially be used to overcome a range of incorrect assumptions or representations. In contrast the Explanation and Justification condition did not specify particular mechanisms but this nevertheless had a training effect. Further research is needed to determine how these general instructions to explain and justify any proposed solution facilitated verbal insight problem solving.

Although not directly comparable, the training score of 68% is in line with the generic studies of Wicker et al. (1978) and Ansburg and Dominowski (2000), who reported effects of 63% and 60% respectively. Indeed, inconsistency checking encouraged solvers to assess their interpretations of a problem in identifying their faulty representations, which was also the aim of the studies mentioned above although it was not explicitly stated as such. The difference between the present training and those employed by Wicker et al. and Ansburg and Dominowski is that the former encouraged participants to systematically examine the problem statement in greater depth by comparing interpretations concerning different parts of the problem

with the problem statement such that the likelihood of identifying a faulty representation was increased. In addition, participants were asked to explicitly consider the plausibility of their hypotheses in solving the problem. These steps broke down the problem solving process, thus providing participants with greater guidance in reaching the solution.

The results also provide indirect support for theories of insight problem solving which stipulate that an incorrect representation triggered by an insight problem inhibit problem solving, and that solution is likely to be attained once the problem is re-represented (e.g., Ohlsson 1992; Knoblich et al., 1999, 2001). Indeed, participants in the No training and Practice conditions found this difficult to achieve. However, as training did not raise the solution rate to 100%, it is clear that other sources of difficulty were present that were not overcome through training. This again highlights the powerful inhibitory effect of a familiar but incorrect representation on problem solving and the future need to investigate how training can be further improved.

CHAPTER 8

A measure of constraint activation in verbal insight problem solving

Past studies have tended to focus on the processes underlying insight problem solving and, at present, there is no means of identifying whether or not a person is constrained in their representation of the problem other than by asking them to solve the problem. Thus, representational change is inferred if a person is able to solve a problem. In contrast, lack of representational change is inferred when a person is unable to solve a problem. This is particularly problematic as there are large individual differences both within and between problems which make it difficult to not only predict whether a person could solve a particular problem but also whether or not a particular training manipulation is effective without testing a person with the actual problem. Therefore, a mechanism is needed that is capable of measuring constraint activation without requiring a person to attempt a problem to determine whether their representation is constrained. This chapter begins with a discussion of the issues concerning the measurement of representational change. It then introduces an experiment that aimed to test a novel method of predicting the degree of constraint activation with a given verbal insight problem.

8.1. Representational change during insight problem solving

Ash, Cushen, and Wiley (2009) identified three issues or difficulties facing the empirical investigation of restructuring or representational change during insight problem solving. The first issue concerned the difficulty of selecting problems to test theories of representational change. A variety of problems have been classed as insight problems including matchstick arithmetic problems, rebus problems, and word

riddles, which make it difficult to identify what is common between these problems. The second issue concerns the lack of appropriate controls or comparison conditions. Often successful performance on a target problem is compared with those who were unsuccessful. However, this type of comparison provides no evidence for representational change during the solution attempt. Further, studies that include a comparison condition tend to vary the difficulty of the baseline problem. Thus, differences in the solution rate between the target and control problems may arise due to problem difficulty rather than the need for representational change. The third issue concerns lack of differentiation between spontaneous restructuring versus the addition of new information in facilitating restructuring. Ash et al. defined spontaneous restructuring as changes in the representation due to the assimilation of new information, acquisition of new skills or because environmental events cued different memory traces. It was argued that some studies failed to differentiate between successful solvers and unsuccessful solvers who were given solution feedback. Thus, solution feedback may help unsuccessful solvers to comprehend the steps required to solve a problem and therefore their experiences are likely to be different to those who independently reached the solution.

The criticisms outlined by Ash et al. (2009) were taken into consideration in designing a method that could indicate whether a problem solver's representation of an insight problem was constrained. Thus, only verbal insight problems were selected for testing in both the present and previous experiments reported in this thesis to overcome the issue of variability between the types problems studied. However, verbal insight problems also vary in terms of their associated constraints and their influence on problem solving, which the current experiment aims to shed some light on. In order to address the second issue of lack of appropriate control noted by Ash et

al., some of the experiments in this thesis included verbal protocol analysis, which revealed differences in problem representations between trained and control participants. Finally, the issue of solution feedback influencing performance was addressed by not providing participants with solution feedback when they solved test problems.

8.2. The Lexical Decision Task (Mayer & Schvaneveldt, 1971)

Verbal insight problems are deemed difficult to solve because they deliberately mislead readers to make inferences from particular words or themes within the problem (Ansburg & Dominowski, 2000; Kershaw & Ohlsson, 2004). A robust technique for studying the activation of semantic associations to a given word, referred to as a 'prime', is Meyer and Schvaneveldt's (1971) Lexical Decision Task (LDT) (e.g., Neely, 1991; Shelton & Martin, 1992; Hermans, Houwler, & Eelen, 2001; Perea & Rosa, 2002). Experiment 6 aimed to test whether the LDT could be used to identify whether a person's representation of an insight problem is constrained. Typically, participants are presented with several strings of letters and asked to classify whether each is a 'word' or a 'nonword'. Before performing the actual LDT, participants are 'primed' with a certain stimulus. A consistent finding is that participants are faster to respond to a target word i.e., a semantic associate to the primed word, when they are first shown a semantically related prime. Hence, participants are faster to classify 'nurse' as a word when it is preceded by 'doctor' than when it is preceded by 'butter'. The semantic priming effect demonstrated by LDTs suggest that if a problem solver is constrained after reading a verbal insight problem, then they are likely to respond faster to a target word that is semantically

representative of the problem constraint. An example of a verbal insight problem to illustrate this point is the Unseen Walker problem:

‘A man walked home after having been out drinking. He walked down the middle of a deserted country road. There were no streetlights to illuminate the road and there was no moonlight. He was dressed all in black. Suddenly a car that did not have its headlights on came racing down the road. At the last moment, the driver of the car saw the man and swerved to avoid him. How did he manage to see him?’

When attempting to solve this problem, it is often incorrectly inferred that it is night-time when in fact it is day-time, hence why the driver was able to see the man. It follows then that priming resulting from reading the problem will result in the problem solver being constrained and thus more likely to respond faster to the target word ‘night’, which is representative of the problem constraint or inference that is likely to be drawn when this problem is encountered for the first time.

Yaniv and Meyer (1987) found that participants were unable to retrieve answers to factual questions, yet they were quick to respond to the answers in a subsequent LDT, thus suggesting that inaccessible stored information primes the later recognition of the information. This suggests that if a problem solvers’ representation of a problem is constrained, then they are likely to respond faster to a word that is semantically representative of this representation.

To summarise, it is proposed that the LDT can offer a novel way of empirically measuring constraint activation to a fine-grain level through variations in reaction time to a target word after reading a verbal insight problem. First, it was predicted that a priming group required to read a verbal insight problem will exhibit a faster reaction time when subsequently tested with a target word (that is associated

with an incorrect representation of that problem) on a LDT in comparison to a group that is required to read a neutral passage of text. Thus, a faster response to the target word would represent increased constraint activation with respect to the insight problem. Second, participants who were identified as being constrained as a result of the LDT are unlikely to be able to solve the problem afterwards.

8.3. Pilot study

As verbal insight problems differ in terms of their associated constraints (Isaak & Just, 1995), it was important to first identify a selection of verbal insight problems and a potential target word that was perceived representative of the constraint associated with each problem. If participants in the pilot study associated the selected target word strongly with the problem, then it could be considered as a strong semantic representative of the problem constraint, which could subsequently be used in the LDT to measure the degree of constraint activation for a given problem.

8.3.1. Method

Participants

Forty-four undergraduate students from Cardiff University participated on a voluntarily basis. Ages ranged between 19 to 48 years ($M = 24$, $SD = 5.70$).

Materials and Design

Seven verbal insight problems were selected and randomly presented (see Appendix I). Using a semantic differential scale scored 1 (weak) to 7 (strong), participants rated the extent to which they associated each problem with each of two words that were selected by the experimenter: a target word that represented the

problem constraint and a non-target word that was related to the problem solution. For example, after reading the Unseen Walker problem participants were asked to rate the extent to which they associated the problem with the words 'night' (the target word) and 'day' (the non-target word). For each problem, the two words were rated one after the other and presentation was counterbalanced. For the Unseen Walker problem, if a participant was constrained and an inappropriate representation was generated, then 'night' would be associated strongly and 'day' weakly. Conversely, if participants were not constrained i.e., they did not incorrectly infer that it was night-time, they were more likely to strongly associate 'day' than 'night' with the problem.

Procedure

Each participant was presented with the seven problems in a booklet, where one page was used per problem. Instructions on how to rate the problems, using the semantic differential scale described in the preceding section, was provided at the start of the booklet, which took approximately 10 minutes to complete. Participants were debriefed at the end of the study.

8.3.2. Results

Table 8.1 reports the means and SDs for the target and non-target words for all seven problems. The average rating for the two words for each problem was assessed and when a score of five or above was given to the target word it was considered a strong association and thus a strong representative of the problem constraint. Target words for five out of the seven problems had an average rating of five or above and t-tests revealed significant differences in rating between target and non-target words for each problem (all $ps < .05$). Therefore, when the target word was associated strongly

with the problem, the non-target word was given a low rating. The Jumped and Tan problems were not used for further experimentation in this experiment.

Table 8.1. Mean rating and SD for target and non-target words for each problem

Problem	Word		
	Target	Non-Target	
Unseen Walker		<i>Night</i>	<i>Day</i>
	Mean	6.14	1.91
	SD	0.97	1.25
Murder		<i>Men</i>	<i>Women</i>
	Mean	6.19	1.86
	SD	1.28	1.06
Professor Bumble		<i>Driving</i>	<i>Walking</i>
	Mean	5.91	1.91
	SD	1.38	1.17
Shot		<i>Killed</i>	<i>Photographed</i>
	Mean	5.63	2.12
	SD	1.59	1.55
Guide		<i>Person</i>	<i>Maps</i>
	Mean	5.02	2.67
	SD	1.93	2.03
Jumped		<i>Outside</i>	<i>Inside</i>
	Mean	4.21	2.65
	SD	1.96	1.93
Tan		<i>Naked</i>	<i>Clothed</i>
	Mean	4.23	3.12
	SD	2.05	1.78

8.3.3. Discussion

The assumption behind the pilot study was that participants would associate each problem with a target word selected by the experimenter to be representative of the problem constraint. In contrast, it was expected that participants would associate the non-target word, which is required for the solution of problems, less strongly to the problem. As predicted, participants associated the problems strongly with the target words which suggested that the incorrect stereotypical inferences were drawn. That is, the target words selected for five out of the seven verbal insight problems

were representative of the problem constraint and thus appropriate for use in the experiment (see Appendix J).

8.4. Experiment 6

The pilot study identified five verbal insight problems and target words suitable for assessing constraint activation for each problem. Using these problems and target words, Experiment 6 aimed to utilise the LDT in a novel way to determine constraint activation for a given problem.

The experiment was conducted in three stages for each insight problem. In the first stage, participants in the priming condition read a verbal insight problem whereas those in the no priming condition read a neutral passage of text. In the second and third stages, participants in both conditions followed the same procedure. Thus, in the second stage, participants completed the LDT corresponding to the verbal insight problem presented to participants' in stage one of the priming condition and in the final stage, participants were given the problem to solve. Two predictions were made. Firstly, if a participant's representation of a problem is constrained, they were more likely to make an incorrect inference after reading the problem (in the priming condition) and thus respond faster to the target word in the LDT. After completing the LDT, participants were asked to solve the problem. Secondly, if reading the verbal insight problem triggers constraint activation, then participant's performance on the problem is likely to be impaired resulting in no solution. This would suggest that reaction data for target words could provide a means of predicting solution or not in advance of asking a participant to solve a verbal insight problem.

8.4.1. Method

Participants

Fifty-six first year undergraduate psychology students from Cardiff University participated in this experiment as partial fulfilment of course requirements⁷. Ages ranged between 18 to 25 years ($M = 18.71$, $SD = 1.23$).

Apparatus and Stimuli

Verbal insight problems, a passage of neutral text (for control participants), and the LDTs were presented on a computer screen. Reaction time (in milliseconds/ms) to target words was measured and verbalisations recorded via a microphone.

Prior to completing the corresponding LDT for each of the five insight problems identified by the pilot study for experimentation, each problem was displayed on the screen for either 50 s (seconds: Unseen Walker, Murder and Shot problems) or 30 s (Professor Bumble and Guide problems). Thus, the presentation times for problems were adjusted to minimise the amount of extra time participants had to re-read shorter problems. Participants in the no-priming condition were presented with a neutral passage of text instead of the verbal insight problem for the same length of time before the LDT. These passages were matched on number of words with the corresponding verbal insight problem and the content was unrelated to the target word or verbal insight problem (see Appendix K).

Five experimental LDTs and one practice LDT were used. The stimuli used for each experimental LDT consisted of six words (i.e., five words and the target word) and six non-words. All words were matched as much as possible to the target

⁷ All materials were pilot tested using 20 participants. As no changes were required, datasets for these participants comprised the final sample.

word on mean reaction time and length in letters according to the English Lexicon Project (ELP: Balota et al., 2002), which is a database of behavioural and descriptive characteristics of over 40 thousand words and nonwords. Nonwords were also selected from Balota et al. in order to match their characteristics to the words as much as possible. However, as reaction times for nonwords are generally higher than that of words (Balota et al., 2002), nonwords were matched only on length in letters with the words. All LDT sets are provided in Appendix L.

The presentation of the target word was randomised from fourth to ninth position in the LDT to reduce the likelihood of participants anticipating the position of the target word.

Test Problems

Five verbal insight problems were used: the Unseen Walker problem, the Murder problem, the Professor Bumble problem, the Shot problem, and the Guide problem (see Appendix K).

Design

Participants were randomly allocated to one of two conditions: no-priming (control) or priming (experimental) condition. The five verbal insight problems/five passage of texts and corresponding LDTs that included the target word associated with the problem constraint were randomly presented. Table 8.2 outlines the design followed for presentation of the first verbal insight problem, which was repeated for the remaining four test problems. The dependent variables were participants' reaction time (in milliseconds) to target words and whether they correctly solved the subsequent problem.

Table 8.2. Table of design for Experiment 6

Priming Condition		
Stage one	Stage Two	Stage Three
Read verbal insight problem one	Completed LDT that included the target word that is representative of the problem constraint for verbal insight problem one.	Given verbal insight problem one to solve.
No Priming condition		
Stage one	Stage Two	Stage Three
Read neutral passage of text that is equivalent in length to verbal insight problem one.	Completed LDT that included the target word that is representative of the problem constraint for verbal insight problem one.	Given verbal insight problem one to solve.

Procedure

Participants completed the experiment individually in a quiet room.

Participants read the introduction to the experiment and then were given practice in ‘thinking aloud’, as in previous experiments. Participants then completed a practice LDT. Participants’ were seated approximately 50cm from the computer screen. All LDTs commenced with the presentation of the fixation marker, +, at the centre of the screen that remained visible for .25 s. Then the first word/nonword in the LDT appeared and remained on the screen until the participant responded by either pressing the *A* key for ‘word’ or the *L* key for ‘nonword’.

After completing the practice LDT, participants in the priming condition read the first verbal insight problem whereas participants in the no-priming condition were presented with a neutral passage of text. Once the problem or passage of text disappeared off the screen, participants completed the corresponding LDT which included the target word. Immediately upon completion of the LDT, the verbal insight

problem was presented for four minutes and participants' in both conditions were required to think aloud during problem solving. Thus, participants in the priming condition were presented with the same verbal insight problem before and after the LDT whereas participants in the no-priming condition viewed the verbal insight problem once i.e., after completing the LDT. Verbalisations were recorded continuously. If participants were silent for a period of time, the experimenter used two non-directive prompts: 'What are you thinking?' and 'Please keep talking'. After completing the problem, participants rated the familiarity on a scale ranging from 1 (very unfamiliar) to 5 (very familiar) in order to ensure participants were not familiar with the problems. Participants were not given solution feedback. The above procedure was repeated for the remaining four test problems. Finally, participants were debriefed and asked not to reveal information about the experiment to others. The duration of the experiment was approximately 45 minutes.

8.4.2. Results and Discussion

No participants were familiar with the test problems; therefore data for all 56 participants were included in the analyses. Due to violations of skewness on total number of problems solved and all reaction time data, all variables were transformed using a logarithmic function that served to eliminate these problems and analyses were performed on transformed data.

It was hypothesised that participants in the priming condition were likely to be constrained after reading a verbal insight problem. Hence, participants' who were constrained were 1) more likely to display a faster reaction time to the target word in the LDT, and 2) remain constrained by failing to correctly solve the problem after completing the LDT.

To test prediction one, a t-test was conducted between conditions on average reaction time for the target word across five problems, which revealed that participants in the priming condition ($M = 621.50$, $SD = 101.09$) were faster to respond to the target word than participants in the no priming condition ($M = 679.14$, $SD = 134.95$), ($t(54) = 1.83$, $p = .04$, 1-tailed). This indicates that participants who had read the verbal insight problem prior to completing the LDT were more likely to respond faster to the target word associated with that problem, thus illustrating that participants representation of the problem was constrained which consequently resulted in a faster response to the target word that was representative of the problem constraint.

The second prediction of Experiment 6 was that a faster reaction to target words would suggest constraint activation, and that this would be further corroborated when participants demonstrated an inability to solve the problem after completing the LDT. To test this prediction, a Pearson's correlation was conducted between average reaction time for the target word across five problems and the total number of problems solved out of five. The results revealed a weak, negative correlation between average reaction time data and solution frequency data, ($r(56) = -.23$, $p = .05$, 1-tailed). This suggests that a faster reaction to the target word was indeed associated with a non-solution result.

Although the above finding provides support for constraint activation, it was possible that participants were simply faster to react to the target word as a result of demand characteristics. To test this prediction, a difference was first calculated between the baseline reaction times available for the (non-target) words and the average reaction time for the target word, and a Pearson's correlation was conducted between this difference and the total number of problems solved out of five. The

correlation was not significant ($r(56) = -.14, p > .05$, 2-tailed), and thus is not in accordance with the previous finding. This suggests that the LDT was not predictive of constraint activation in verbal insight problem solving in the present experiment.

8.5. General Discussion

The LDT was employed as a technique for measuring the activation of an incorrect or faulty representation for individual verbal insight problems that was subsequently likely to constrain problem solving performance. It was hypothesised that reaction time data could be used to infer constraint activation and consequently, the likelihood of reaching the solution to a problem. Hitherto, it is the first time an established method such as the LDT has been utilised in this way to try and measure constraint activation in verbal insight problem solving. The results indicated that participants who had read a verbal insight problem were subsequently faster to respond to a target word, which was considered semantically representative of problem constraint, in a LDT. Furthermore, reaction time to target words was predictive of whether a participant was able to solve the problem after completing the LDT. However, this latter finding was not supported when reaction times for non-target words were taken into consideration. Therefore, it was possible that demand characteristics encouraged participants' to respond quickly to target words rather than it being the influence of constraint activation per se.

Nevertheless, this study was the first attempt at measuring constraint activation in verbal insight problem solving. The LDT was utilised in a novel way to demonstrate that if a problem solver's holds the incorrect representation of a problem, then they are likely to respond faster to a word that is representative of the problem constraint, and subsequently performance on the problem is likely to be impaired. It is

unknown from these results what constitutes a fast reaction time which is indicative of constraint activation as further research is required using a larger set of insight problems.

This study however is not without limitations. To standardise the features of the target word and 'words' in the LDT, the lexical features that were considered were length in letters and average reaction time. Another lexical feature that may have contributed to the variability in reaction time data was word frequency i.e., how common a target word is. Research by Balota and Chumbley (1984) found that word frequency was highly related to lexical decisions. That is, high-frequency or common words were recognised more quickly than low-frequency words and this aspect was not controlled in the present study.

Another factor that may have reduced the effects of priming is that once participants in the priming condition had learnt that they were required to solve the problem after completing the LDT, it was possible that they had rehearsed the problem solution during completion of the LDT. Thus, participants in the priming condition may have focussed on generating possible solutions to the problem whilst completing the LDT, which subsequently affected reaction time to the target word.

In summary, the LDT was a successful method to some extent in predicting constraint activation for a given verbal insight problem. However, further research is required to determine whether the LDT may serve as an accurate measure of constraint activation in verbal insight problem solving. The advantage of this method is that it provides the possibility of predicting whether a problem solver is likely to be constrained in terms of representation without an explicit attempt being made at solving the problem. This information might be useful in reducing the likelihood of

employing problems that produce floor or ceiling effects, which is problematic across training studies.

CHAPTER 9

General Discussion

The primary aim of this thesis was to design and implement training to overcome implicit assumptions in solving verbal insight problems. This chapter begins with an overview of the experimental findings and their contribution to the field of problem solving. This is followed by discussion of the limitations of the research and recommendations are made for future work.

9.1. Overview of experimental findings

The first four experimental chapters of this thesis (*Chapters 4 to 7*) tested novel training programmes that were designed to improve verbal insight problem solving. The approach to training was either specific or generic. Thus, specific training aimed to effect change for a particular category of insight problems which shared common features, whereas the aim of generic training was to facilitate performance on any verbal insight problem in spite of the constraint associated with it. *Chapter 8*, the final experimental chapter, reported an experiment that was different in that it aimed to measure whether a person's representation was constrained prior to asking them to solve a problem. The experimental findings are discussed below with reference to past research on insight problem solving.

Specific training studies on insight problem solving have tended to focus on one test problem thereby limiting the generalisability of training (e.g. Burnham & Davis, 1969; Chronicle et al., 2001; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981). This limitation was addressed by the specific training programmes tested in this thesis as a wider range of verbal insight problems was tested by categorising

problems that were solvable through the application of simple heuristics. Experiment 1 (reported in *Chapter 4*) trained participants in using heuristics to solve two categories of verbal insight problems: those that contained an ambiguous word or an ambiguous name in the problem specification. For the ambiguous word category, training raised the solution rate from 26% (no training condition) to 59%, whereas for the ambiguous names category, the solution was raised from 4% (no training condition) to 39%. As the results may have been influenced by the choice of test problems for the ambiguous word category, Experiment 2 further tested the heuristic-based training for problems with ambiguous words only. An additional aim of the experiment was to test whether problem length was a determinant of problem success. That is, whether longer problems increased the search for the solution through the problem space, thus resulting in poorer performance. No effect of problem length was found although training raised solution from 26% (no training condition) to 71%, thus performance was improved substantially. In general, Experiments 1 and 2 demonstrated positive transfer for trained categories, which was consistent with the success of similar training (e.g., Schoenfeld, 1979).

Experiment 3 was slightly different in that training aimed to improve performance on functional fixedness insight problems that were presented verbally, unlike past research which has presented such problems both pictorially and verbally (Adamson, 1952; Birch & Rabinowitz, 1951; Duncker, 1945; Frank & Ramscar, 2003; Maier, 1931). Success on both functional fixedness problems and verbal insight problems is dependent on overcoming an unhelpful stereotypical response. In the case of the former, a typical function for an object is triggered that prevents the generation of an unusual function that is required for solution. The approach to training in Experiment 3 was novel in that it encouraged participants to think divergently and yet

systematically by considering several functions for objects mentioned in the problem statement thereby increasing the probability of solution attainment. Indeed, the results provided evidence of positive transfer to test problems that were exemplars of the functional fixity problem category only, as training raised performance from 6% (no training condition) to 39%.

Collectively the results of Experiments 1 to 3 provide modest support for theories of transfer (Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901). Further, although training successfully improved performance on the trained categories of problems, it can nevertheless be argued that the focus of training was narrow. Therefore the challenge was to design training that facilitated solutions on any verbal insight problem irrespective of the assumption or constraint associated with it. This was the aim of the two generic training experiments reported later in this thesis, which aimed to address the variability in constraints that is found among verbal insight problems.

As was the case in the generic training studies of Wicker et al. (1978) and Ansburg and Dominowski (2000), the generic training in Experiments 4 and 5 in this thesis also aimed to firstly increase participants' awareness of the possibility of incorrect assumptions associated with insight problems. Indeed, past research has highlighted the importance of drawing trainee's attention to the nature of the difficulties associated with problems before providing a strategy to overcome such difficulties (Campion & Ambruster, 1985). The key difference was however that the training in Experiments 4 and 5 went further to specify and give practice in using a cognitive process to help participants identify and correct their incorrect assumptions, which has never been attempted by past training research in this area.

The first experiment, Experiment 4 (reported in *Chapter 6*), introduced training in using the inconsistency checking mechanism, which in theory could be applied to solve any verbal insight problem. This mechanism attempted to induce participants to scrutinise their interpretations and assumptions by comparing them with the problem specification in the hope of facilitating recognition and restructuring of an inappropriate representation. As practice with feedback is a prerequisite for promoting positive transfer for many tasks (e.g., Anderson, 1983; Ansborg & Dominowski, 2000; Newell & Rosenbloom, 1981), participants were provided with practice in using this mechanism to try and identify their incorrect representations. Further, training and test problems shared similarities to promote positive transfer (Anderson, 1983; Gick & Holyoak, 1980; Thorndike & Woodworth, 1901) as verbal insight problems were used that all involved only one constraint that needed to be overcome through restructuring. Experiment 4 revealed that training raised performance by 19% from the no training condition (21%) to 40%. Further, verbal protocol analyses suggested that what was learnt during training was indeed applied during problem solving.

The aim of Experiment 5 was to raise the training score of 40% by introducing an iterative approach to using inconsistency checking in the hope of improving the likelihood of participants recognising their faulty representation. Hence, participants were trained to systematically identify discrepancies between interpretations of different parts of a problem and the problem specification. Indeed, a similar approach to problem solving was successful in overcoming functional fixity in the specific training tested in Experiment 3 in this thesis. An additional condition was employed in which participants were instructed to explain and justify their solutions during practice to facilitate reflective thinking as research on search and design tasks has

demonstrated the positive benefits of such instructions (Berry & Broadbent, 1984, 1987; Wetzstein & Hacker, 2004; Winckelmann & Hacker, 2010). The results revealed greater support for inconsistency checking as a training score of 68% was found, hence a significant improvement from the training score of 40% reported in Experiment 4. Further, instruction to explain and justify solutions during practice resulted in a 58% training score. The difference in performance between these two conditions was not significant although the former had specified a mechanism for identifying the incorrect representation unlike the latter. Further research is needed to determine why instruction to justify solutions had such a positive effect on problem solving. The results also demonstrated that practice with feedback, which is important for facilitating positive transfer (Anderson, 1983; Ansburg & Dominowski, 2000; Newell & Rosenbloom, 1981), was not sufficient to improve solution rate without the provision of training that gives explicit practice in how to solve such problems. Collectively, the findings of Experiments 4 and 5 suggest that it is possible to facilitate positive transfer to a range of verbal insight problems that require some implicit assumption to be overcome.

The final experimental chapter (reported in *Chapter 8*) was different to the above training experiments. At present it is not possible to identify which verbal insight problems induce a powerful incorrect representation without testing a participant on the problem. Therefore, Experiment 6 introduced the Lexical Decision Task (LDT; Hermans et al., 2001; Meyer & Schvaneveldt, 1971; Neely, 1991; Perea & Rosa, 2002; Shelton & Martin, 1992) as a novel method for identifying when participants were constrained by an incorrect problem representation prior to asking the participant to solve the problem. The results indicated an overall effect of priming. That is, when compared to participants who read a neutral piece of text, participants

who had read a verbal insight problem prior to completing the LDT were faster to respond to the target word (that was semantically representative of the problem constraint or assumption). Although the results suggested that a faster reaction time was associated with weaker solution rates when participants were asked to solve the problem immediately upon completion of the LDT, further analyses indicated that this finding may have been due to demand characteristics. Nevertheless, the results of Experiment 6 have important implications for training research as it suggests that the LDT may be used as a means of predicting solution or no solution in advance of problem solving. Further research is therefore required to validate use of the LDT as a predictive technique.

9.2. Some implications

Consistent with past research on insight problem solving, performance data for no training participants across the training experiments in this thesis demonstrated that insight problems are generally difficult. Different approaches to training were designed that were informed by theories of insight problem solving and transfer, which produced problem solving success. The results make two broad contributions which are discussed below.

The research in this thesis did not set out to explicitly test any specific theory of insight problem solving that was reviewed in *Chapter 2*. Nevertheless, the findings lend indirect support for Representational Change Theories of insight (e.g., Knoblich et al., 1999, 2001; Ohlsson, 1992) that argue insight problem solving is constrained by an incorrect representation of the problem. Indeed, verbal protocol data suggested that participants' solutions to problems were in line with their incorrect interpretation of the problem. Researchers such as Jones (2003) and Öllinger et al. (2006) tested both

Representational Change Theory and the Progress Monitoring Theory (MacGregor et al., 2001), which the present research did not set out to test. However, it was unlikely that participants used a lookahead heuristic, as advocated by the Progress Monitoring Theory, as this heuristic is particularly suited to solving move problems. The aim of the training programmes was to facilitate representational change and performance data provided indirect support for this.

The current thesis also makes practical contributions to the study of verbal insight problem solving by providing examples of training that successfully facilitated performance. When faced with a problem situation it is common practice for the problem solver to default to habitual conceptualisations, which, in most circumstances, is helpful in reaching the solution. However, such habitual conceptualisation can be unhelpful in unusual problem situations, and in some cases, detrimental as illustrated by the David-Besse (1985) and Three Mile Island (1979) incidents. The present research demonstrated how training can be designed to overcome these habitual conceptualisations. That is, solvers must first be made aware of how past habits impede problem solving before guided practice is given in overcoming these unhelpful responses using deviant scenarios. Further, practice in overcoming these unhelpful responses can take place either through specific mechanisms (Experiments 1 to 3) or through a generic mechanism such as inconsistency checking (Experiments 4 and 5). Of course, the benefits of such training are greater when it is useful in a variety of different problem situations where a novel interpretation of a problem is critical to solution. Indeed, results of the latter training studies (Experiments 4 and 5) in this thesis provided support for the cognitive process of inconsistency checking, which has been successfully utilised to break habitual

barriers in reasoning about unusual faults within the nuclear industry (Patrick et al., 1999).

9.3. Limitations

Ash et al. (2009) recently challenged whether restructuring plays any role in problem solving. It cannot be concluded with 100% certainty that representational change occurred as a result of training in the present experiments although verbal protocols provide some indirect evidence of this and performance data was indicative of a change in representation leading to the correct solution. Further, verbal protocols not only gave some clues as to the nature of the cognitive processes during problem solving but also could be used to ascertain whether the training was used by participants during problem solving. Although past research suggests that the think aloud procedure does not overshadow thought processes when the guidelines set by Ericsson and Simon (1993) are followed (Chrysikou, 2006; Fleck & Weisberg, 2004; Gilhooly, Fioratou, & Henretty, 2010), it cannot be ascertained whether this was true in the present experiments as a silent control condition was not included. Despite much justification for the utility of verbal protocols, it is important to note that they may not be as veridical as we might hope them to be.

Another limitation is that it is difficult to disentangle the different aspects of training that resulted in improved performance in Experiments 1 to 5. Thus the effect of the subject matter of the training (e.g., heuristics) cannot be separated from the processes involved e.g., practice.

A major problem concerns the variability of the constraints associated with the problems and their variation in difficulty. Although the aim of the experiments was to establish overall training effects, this variability no doubt affected performance on

individual problems where some problem constraints may have been easier to overcome than others. To some extent this problem variability will also contribute to the different levels of performance across the training studies and training effects do depend to some extent on the nature of the problems used.

Finally, a possible limitation that may explain some of the null effects concerns the issue of lack of power due to small sample sizes. Further, it must be noted that whilst steps were taken to minimise participants' familiarity with test problems through the collection of familiarity ratings, there was still the possibility that some participants were able to solve test problems based on past experience.

9.4. Recommendations for future research

According to the Cognitive Load Theory (Sweller, 1998; Sweller, Van Merriënboer, & Paas, 1998), working memory is very limited when handling new information. Hence, when the processing of information requires the solver to organise, contrast or compare, the solver is only able to consider two or three items simultaneously. Further, new information is lost within 15 to 30 seconds if it is not rehearsed (Artino, 2008). This suggests that the requirements of training may have resulted in intrinsic cognitive load i.e., the mental work load required for problem solving, therefore impinging on working memory. In fact, cognitive overload may have been an issue across all five training experiments reported in this thesis. For example, participants may have found it difficult to monitor which heuristics they had applied in attempting to solve test problems or which words they had assessed in identifying the ambiguous words in Experiments 1 and 2. Also, in Experiment 3, participants were encouraged to produce an exhaustive list of functions for items in a problem which may have impinged on working memory as it required participants to

remember all the functions they had considered. Perhaps the demands on working memory were greatest in Experiments 4 and 5. That is, inconsistency checking required by the training may have been difficult for participants to perform because of the need to attempt to examine one's own representation of the problem and then to compare it with the problem specification.

One suggestion for reducing cognitive load is through the use of an external memory aid (e.g., pen and paper). Research evidence indicates that it is necessary to provide an external memory aid for problem solving to facilitate the selection of an appropriate solution path (Cary & Carlson, 1999). In addition, it has been suggested that people are likely to use an external memory aid (e.g., to make notes) when they believe the benefits of using it outweigh the costs of not doing so (Cary & Carlson, 2001). Rigney (1978) also suggested that tasks such as note taking act as an 'orienting task' as it 'orients' the trainee to develop appropriate learning strategies. Future research needs to ensure that training does not fail due to inappropriate memory demands, and therefore should include a condition in which a memory aid could be used in order to reduce and evaluate the effect of such cognitive workload. Further, in relation to the generic training studies, asking participants to make notes during problem solving may facilitate a more precise and detailed comparison between the external written representation and the problem specification.

One of the difficulties experienced in design training for verbal insight problem solving was the limited choice of problems. Cunningham and MacGregor (2006) proposed rebus puzzles as candidates for insight problem solving which require further investigation. A rebus combines verbal and visual clues to create a familiar phrase (e.g., you *just* me = 'just between you and me'). It is argued that rebus problems require restructuring because past experiences of reading result in the

solver focusing on syntax rather than the semantic meaning. The solution to a rebus puzzle ignores grammar and instead the spatial relationships of components in a problem must be given a verbal interpretation. Furthermore, our past experiences of reading ignore print characteristics such as bold letters or capital letters whereas such characteristics are important to the solution of rebus problems (e.g. PUNISHMENT = 'capital punishment'). The advantage for using rebus problems is that the level of difficulty can be manipulated (MacGregor & Cunningham, 2008) in terms of the number of implicit assumptions involved. Many studies have attempted to control for problem difficulty by comparing performance on 'insight' problems with 'non insight' problems that were matched on overall solution rates (e.g., Gilhooly & Murphy, 2005; Jung-Beeman et al., 2004; Lavric et al., 2000; Metcalfe & Wiebe, 1987; Schooler et al., 1993). However, rebus problems are similar to verbal insight problems because past experiences trigger an incorrect cognitive representation and restructuring of the representation is required for solutions. To further shed light on the role of representational change in insight problem solving, future research could adapt the training designs tested herein using rebus problems that involve one implicit assumption, as did the verbal insight problems employed in this research.

Finally, it is unknown to what extent the training effects are durable as participants were tested immediately after completion of the training programmes. The success of training will be better determined if future research tested participants after a longer duration elapsed between time of training and time of testing.

9.5. Conclusions

The challenges that verbal insight problem solving pose to the training designer is that such problems do not have a clear state and that the overall nature of

the constraints associated with these problems vary greatly, which may explain why there is a paucity of training research in this area. The experiments reported in this thesis provide convincing evidence that habitual responses triggered by verbal insight problems are possible to overcome through training that induces a shift in representation. The implication of these results is that domains which require problem solvers to interpret situations in an unusual way would benefit from training similar to that developed in this thesis, to help avoid the constraints imposed by problem solvers' past experiences. In particular, a psychological shift in representation can be induced by developing awareness of incorrect representations which is coupled with practice in identifying and correcting these faulty representations. In conclusion, the current thesis made some important advances in research in this area.

REFERENCES

- Adamson, R.E. (1952). Functional Fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology*, 44, 288-291.
- Adamson, R.E., & D.W. Taylor (1954). Functional fixedness as related to elapsed time and to set. *Journal of Experimental Psychology*, 47, 122-126.
- Adeyemo, S.A. (2003). The effects of training on insight problem solving. *Psychology and Education: An interdisciplinary Journal*, 40(2), 51-55.
- 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.
- Ansburg, P.I., & Dominowski, R.L. (2000). Promoting insightful problem solving. *Journal of Creative Behavior*, 34, 30-60.
- Anderson, J.R. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.
- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology*, 60, 451-474.
- Ash, I.K., Cushen, P.J., & Wiley, J. (2009). Obstacles in investigating the role of restructuring in insightful problem solving. *The Journal of Problem Solving*, 2(2), 6-41.

- Ash, I K., & Wiley, J. (2006). The nature of restructuring in insight: An individual-differences approach. *Psychonomic Bulletin and Review*, *13*(1), 66-73.
- Ansburg, P.I. (2000). Individual differences in problem solving via insight. *Current Psychology*, *19*, 143-146.
- Ansburg, P.I., & Dominowski, R.L. (2000). Promoting insightful problem solving. *Journal of Creative Behavior*, *34*(1), 30-60.
- Artino, A.R., Jr. (2008). Cognitive load theory and the role of learner experience: An abbreviated review for educational practitioners. *Association for the Advancement of Computing in Education Journal*, *16*(4), 425-439.
- Baddeley, A. (1992). Working memory. *Science*, *31*, 556-559.
- Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340-357.
- Balota, D.A., Cortese, M.J., Hutchison, K.A., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2002). *The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords*. Retrieved July 16, 2008 from www.elexicon.wustl.edu/.
- Bassok, M., & Holyoak, K.J. (1989). Information transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 153-166.
- Berry, D.C., & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, *36A*, 209-231.

- Berry, D.C., & Broadbent, D. E. (1987). Explanation and verbalization in a computer-assisted search task. *Quarterly Journal of Experimental Psychology*, 39(A), 585-609.
- Birch, H.G., & Rabinowitz, H.S. (1951). The negative effect of previous experience on productive thinking. *Journal of Experimental Psychology*, 41, 121-125.
- Bohannon, N.M. (1994). *Insight problem solving: An examination of the process and training of insight*. Unpublished masters thesis.
- Burnham, C.A., & Davis, K.G. (1969). The nine-dot problem: Beyond perceptual organisation. *Psychonomic Science*, 17, 321-323.
- Campione, J., & Ambruster, B. (1985). Acquiring information from texts: An analysis of four approaches. In J. Segal, S. Chipman, and R. Glaser (Eds.), *Thinking and learning skills. Volume 1. Relating instruction to research* (pp. 317-359). Hillsdale, NJ: Lawrence Erlbaum.
- Cary, M., & Carlson, R.A. (1999). External support and the development of problem solving routines. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 1053-1070.
- Cary, M., & Carlson, R.A. (2001). Distributing working memory resources during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 836-848.
- Cheung, S.K., Li, Y.L., Kong, Y.C., Wong, W.Y., & McBride-Chang, C. (2005). Getting an 'Aha!' with different types of hints. *Korean Journal of Thinking and Problem-Solving*, 15, 47-56.
- Christensen, P.R., & Guilford, J.P. (1958). *Creativity/Fluency Scales*. Beverly Hills, CA: Sheridan Supply.

- Chronicle, E.P., MacGregor, J.N., & Ormerod, T.C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 14-27.
- Chronicle, E.P., Ormerod, T.C., & MacGregor, J.N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *Quarterly Journal of Experimental Psychology*, 54A(3), 903-919.
- Chrysikou, E.G. (2006). When shoes become hammers: Goal-derived categorization training enhances problem-solving performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 935-942.
- Chrysikou, E.G., & Weisberg, R.W. (2004). Insight problem solving as goal-derived, ad hoc categorization [Abstract]. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Conference of the Cognitive Science Society*, 1540.
- Cohen, E.D. (1975). *C. G. Jung and the Scientific Attitude*. New York: Philosophical Library.
- Cunningham, J.B., & Macgregor, J.N. (2008) Training insightful problem solving: Effects of realistic and puzzle-like contexts. *Creativity Research Journal*, 20(3), 291-296.
- Davidson, J.E. (2003). Insights about insightful problem solving. In J.E. Davidson and R.J. Sternberg (Eds.), *The psychology of problem solving* (pp. 149-175). New York: Cambridge University Press.
- Dayton, T., Durso, F.T., & Shepard, J. D. (1990). A measure of the knowledge reorganization underlying insight. In R.W. Schvaneveldt (Ed.), *Pathfinder*

- network associative networks: Studies in knowledge organization* (pp. 267–277). Norwood, NJ: Ablex Publishing.
- DeBono, E. (1967). *New think: The use of lateral thinking in the generation of new ideas*. New York: Basic Books.
- DeYoung, C.G., Flanders, J.L., & Peterson, J.B. (2008). Cognitive abilities involved in insight problem solving: An individual differences model. *Creativity Research Journal*, 20, 278-290.
- Dominowski, R.L. (1990). Problem solving and metacognition. In K.J. Gilhooly, M. T.G. Keane, R.H. Logie and G. Erdos (Eds.). *Lines of thinking: Reflections on the psychology of thought* (Vol. 2, pp. 313-328). Chichester, England: Wiley.
- Dominowski, R. (1994). *Insight and instructions*. Paper presented at the 11th Annual conference, Cognitive Psychology Section, British Psychological Society, New Hall, Cambridge.
- Dominowski, R.L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg and J.E. Davidson (Eds.), *The nature of insight* (pp 33-62). USA: MIT Press.
- Dorfman, J., Shames, V.A., & Kihlstrom, J.F. (1996). Intuition, incubation, and insight. In G. Underwood (Ed.), *Implicit cognition in problem solving* (pp. 257-296). Oxford, England: Oxford University Press.
- Dow, G.T., & Mayer, R.E. (2004). Teaching students to solve insight problems: Evidence for domain specificity in creativity training. *Creativity Research Journal*, 16(4), 389-402.
- Duncker, K. (1945). On problem-solving. *Psychological Monographs*, 58, ix, 113.
- Durso, F.T., Rea, C.B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science*, 5, 94-98.

- Dvir, T., Eden, D., Avolio, B.J., & Shamir, B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, 45, 735-744.
- Edkins, G.D. (2002). A review of the benefits of aviation human factors training. *Human Factors and Aerospace Safety*, 2(3). 201-216.
- Egan, D.E., & Greeno, J.G. (1974). Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning, and problem solving. In L.W. Gregg (Ed.), *Knowledge and Cognition* (pp. 43-103). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-281.
- Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data*. (revised edition) (original edition published in 1984). Cambridge, MA: Bradford books/MIT Press.
- Fisher, S.D., Gettys, C.F., Manning, C., Mehle, T., & Baca, S. (1983). Consistency checking in hypothesis generation. *Organizational Behavior and Human Performance*, 31, 233-254.
- Fleck, J.I., & Weisberg, R.W. (2004). The use of verbal protocols as data: An analysis of insight in the candle problem. *Memory and Cognition*, 32(6), 990-1006.
- Frank, M.C., & Ramscar, M. (2003). How do presentation and context influence representation for functional fixedness tasks? [Abstract]. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Boston, MA.
- Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306-355.

- Gilhooly, K.J., & Fioratou, E. (2009). Executive processes in insight and non-insight problems: An individual differences approach. *Thinking and Reasoning*, *16*, 355-376.
- Gilhooly, K.J., Fioratou, E., Anthony, S., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology*, *98*, 611-625.
- Gilhooly, K.J., Fioratou, E., & Henretty, N. (2010). Verbalisation and problem solving: Insight and spatial factors. *British Journal of Psychology*, *100*, 81-93.
- Gilhooly, K.J., & Johnson, C.E. (1978). Effects of solution word attributes on anagram difficulty: A regression analysis. *Quarterly Journal of Experimental Psychology*, *30*, 57-70.
- Gilhooly, K.J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking and Reasoning*, *11*, 279-302
- Goldstein, I.L., & Ford, J.K. (2002). *Training in organizations* (4th edn.). Belmont, CA: Wadsworth.
- Greer, T., & Dunlap, W.P. (1997). Analysis of variance with ipsative measures. *Psychological Methods*, *2*, 200-207.
- Gremler, D.D. (2004). The Critical Incident Technique in Service Research. *Journal of Service Research*, *7*(1), 65-89.
- Howell, M.L. (1956). Use of force-time graphs for performance analysis in facilitating motor learning. *Research Quarterly*, *27*, 12-22.
- Isaak, M.I., & Just, M.A. (1995). Constraints on thinking in insight and invention. In R.J. Sternberg and J.E. Davidson (Eds.), *The nature of insight* (pp. 281-325). Cambridge: MIT Press.

- James, W. (1890). *The principles of psychology*. Cambridge: Harvard University Press.
- Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1017-1027.
- Judd, C.H. (1908). The relation of special training to general intelligence. *Educational Review*, 36, 28-42.
- Kaplan, C.A., & Simon, H.A. (1990). In search of insight. *Cognitive Psychology*, 22, 374-419.
- Keane, M. T. (1989). Modeling 'insight' in practical construction problems. *Irish Journal of Psychology*, 11, 202-215.
- Kershaw, T.C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 3-13.
- Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1534-1555.
- Knoblich, G., Ohlsson, S., & Raney, G.E. (2001). An eye movement study of insight problem solving. *Memory and Cognition*, 29(7), 1000-1009.
- Köhler, W. (1924). *Die physischen Gestalten in Ruhe und im stationären Zustand: Eine naturphilosophische Untersuchung*. Erlangen: Verlag der Philosophischen Akademie.
- Komaki, J., Heinzmann, A.T., & Lawson, L. (1980). Effect of training and feedback: Component analysis of a behavioral safety program. *Journal of Applied Psychology*, 65, 261-270.

- Krippendorff, K. (1980). *Content analysis. An introduction to its methodology*. Beverly Hills: Sage.
- Ladkin, P. (1996). *Extracts from UK AAIB Report 4/90 on the 8 January 1989 accident of a British. Midland B737-400 at Kegworth, Leicestershire, England*.
- Lavric, A., Forstmeier, S., & Rippon, G. (2000). Differences in working memory involvement in analytical and creative tasks: An ERP study. *NeuroReport*, 11(8), 1613-1619.
- Lea, R.B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 1469-1482.
- Ludwig I., & Pieper. W. (1998). Effects of practice and feedback on the detection of embedded figures. *Perception*, 27, ECVF Abstract Supplement.
- Lung, C.T., & Dominowski, R.L. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 804-811.
- Luo, J., & Nicki, K. (2003). Function of hippocampus in insight problem solving. *Hippocampus*, 13(3), 316-23.
- MacGregor, J.N., & Cunningham, J.B. (2008). Rebus puzzles as insight problems. *Behavior Research Problems*, 40, 263-268.
- MacGregor, J.N., Ormerod, T.C., & Chronicle, E.P. (2001). Information processing and insight: a process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 176-201.

- Maier, N.R.F. (1931). Reasoning in humans II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, 12, 181-194.
- Maier, N.R.F. (1933). An aspect of human reasoning. *British Journal of Psychology*, 24, 144-155.
- Maier, N.R.F. (1945). Reasoning in humans: III. The mechanisms of equivalent stimuli and of reasoning. *Journal of Experimental Psychology*, 35, 349-360.
- Mayer, R.E. (1990). Problem solving. In M.W. Eysenck (Ed.), *The Blackwell dictionary of cognitive psychology*. Oxford, England: Blackwell.
- Mayer, R.E. (1995). The search for insight: Grappling with Gestalt psychology's unanswered questions. In R.J. Sternberg and J. Davidson (Eds.), *The nature of insight* (pp. 3-32). Cambridge: MIT Press.
- Mayer, R.E., Dyck, J.L., & Cook, L.K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology*, 76, 1089-1105.
- Metcalf, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory and Cognition*, 15(3), 238-246.
- Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner and G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, Hove, London: Lawrence Erlbaum Associates.

- Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Lawrence Erlbaum.
- Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- O'Connor, P., Flin, R., & Fletcher, G. (2002). Methods used to evaluate the effectiveness of flight crew CRM methods in the UK aviation industry. *Human Factors and Aerospace Safety*, 2, 235-255.
- Ohlsson, S. (1984a). Restructuring revisited I: Summary and critique of the Gestalt theory of problem solving. *Scandinavian Journal of Psychology*, 25, 65-78.
- Ohlsson, S. (1984b). Restructuring revisited II: An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology*, 25, 117-129.
- Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M.T. Keane and K.J. Gilhooly (Eds.), *Advances in the psychology of thinking* (pp. 1-44). New York; London; Toronto; Sydney; Tokyo; Singapore: Harvester Wheatsheaf.
- Öllinger, M., & Knoblich, G. (2009). Psychological research on insight problem solving. In *Recasting Reality Wolfgang Pauli's Philosophical Ideas and Contemporary Science* (pp. 277-300). Atmanspacher, Harald; Primas, Hans (Eds.) Springer.
- Ormerod, T.C., & Ball, L.J. (2007). Qualitative methods in cognitive psychology. In C. Willig and W. Stainton Rogers (Eds.), *The Sage Handbook of Qualitative Research in Psychology*. London: Sage Publications Ltd.

- Ormerod, T.C., MacGregor, J.N., & Chronicle, E.P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 791-799.
- Patrick, J. (1992). *Training: Research and practice*. London; San Diego; New York; Boston; Sydney; Toronto; Tokyo: Academic Press.
- Patrick, J., Grainger, L., Gregov, A., Halliday, P., Handley, J., James, N., & O'Reilly, S. (1999). Training to break the barriers of habit in reasoning about unusual faults. *Journal of Experimental Psychology: Applied*, 5(3), 314-335.
- Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, 66, 180-194.
- Perfetto, G.A., Bransford, J.D., & Franks, J.J. (1983). Constraints on access in a problem solving context. *Memory and Cognition*, 11, 24-31.
- Perreault, W.D., & Leigh, L.E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of Marketing Research*, 26, 135-148.
- Pfeiffer, T. (2004). Problem solving with a simple transformation problem with and without continuous external memory support. *European Journal of Cognitive Psychology*, 16(4), 555-572.
- Reason, J. (1990). *Human error*. Cambridge, England: Cambridge University Press.
- Rigney, J.W. (1978). Learning strategies: A theoretical perspective. In H.F. O'Neil, Jr. (Ed.), *Learning strategies* (pp. 164-205). New York: Academic Press.
- Salas, E., & Cannon-Bowers, J. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52, 471-499.
- Salas, E., Burke, C.S., Bowers, C.A., & Wilson, K.A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *The Journal of the Human Factors and Ergonomics Society*, 43(4), 641-674.

- Scheerer, M. (1963). Problem solving. *Scientific American*, 208(4), 118-128.
- Schoenfeld, A.H. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal of Research in Mathematics Education*, 10, 173-187.
- Schoenfeld, A. H. (1980). Heuristics in the classroom. In S. Krulik and R.E. Reys (Eds.), *Problem solving in school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Schooler, J.W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122(2), 166-183.
- Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, 16(1), 141-148.
- Shelton, J.R., & Martin, R.C. (1992). How semantic is semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1191-1210.
- Shepherd, A., Marshall, E.C., Turner, A., & Duncan, K.D. (1977). Diagnosis of plant failure from a control panel - a comparison of three training methods. *Ergonomics*, 20, 347-381.
- Sio, U.N., & Rudowicz, E. (2007). The role of an incubation period in creative problem solving. *Creativity Research Journal*, 19(2-3), 307-318.
- Sio, U.N. & Ormerod, T.C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, 135, 94-120.
- Sloane, P. (1992). *Lateral thinking puzzles*. New York: Sterling Publishing Co. Inc.
- Smith, S.M. (1995). Getting into and out of mental ruts: In R.J. Sternberg and J.E. Davidson (Eds.). *The nature of insight* (pp. 229-251). Cambridge: MA: MIT Press.

- Sternberg, R.J., & Davidson, J.E. (Eds.). (1995). *The nature of insight*. Cambridge: MA: MIT Press.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.
- Thorndike, E.L., & Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8, 247-261.
- Thomas, J.C. (1974). An analysis of behavior in the hobbits-orcs problem. *Cognitive Psychology*, 6, 257-269.
- Toms, M., & Patrick, J. (1989). Components of fault-finding: Symptom interpretation. *Human Factors*, 31(4), 465-483.
- Wallas, G. (1926). *The Art of Thought*. New York: Harcourt Brace.
- Weisberg, R.W. (1992). Metacognition and insight during problem solving: Comment on Metcalfe. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 426-432.
- Weisberg, R.W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R.J. Sternberg and J.E. Davidson (Eds.), *The nature of insight* (pp. 157-196). Cambridge, MA: MIT Press.
- Weisberg, R.W., & Alba, J.W. (1981). An examination of the alleged role of 'fixation' in the solution of several 'insight' problems. *Journal of Experimental Psychology: General*, 110(2), 169-192.

- Weisberg, R.W., & Alba, J.W. (1982). A problem solving is not like perception: More on Gestalt theory. *Journal of Experimental Psychology: General*, 111(3), 326-330.
- Wertheimer, M. (1945). *Productive thinking*. Chicago, IL: University of Chicago Press.
- Wetzstein, A., & Hacker, W. (2004). Reflective verbalization improves solutions - The effects of question-based reflection in design problem solving. *Applied Cognitive Psychology*, 18, 145-156.
- Wicker, F.W., Weinstein, C.E., Yelich, C.A., & Brooks, J.D. (1978). Problem-reformulation training and visualization training with insight problems. *Journal of Educational Psychology*, 70(3), 372-377.
- Winer, R.I. (2002). How to use Jung's Word Association Test. *Neurocare*. Retrieved February 24, 2009, from: <http://www.gesher.org/Neurocare/Medical%20Education/Jung.html>.
- Winckelmann, C., & Hacker, W. (2010). Question-answering-technique to support freshman and senior engineers in processes of engineering design. *International Journal of Technology and Design Education*, 20(3), 305-315.
- Yaniv, I., & Meyer, D.E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 187-205.

APPENDICES

Appendix A

Test problems - Experiment 1

Category: Ambiguous words

Problem A: MARRIED

A man who lived in a small town in the United States married 20 different women of the same town. All are still living and he never divorced any of them. In this town polygamy is unlawful; yet he has broken no law. How is this possible?

Hypothesised constraint: The man married each woman himself.

Solution: The man is a vicar/priest who married couples together.

Problem B: GUIDE

A mountain climber in the Himalayas took along with him two mountain guides. After a few hours, one of the guides fell into a deep crevasse. The climber and the other guide continued the climb and did not raise the alarm. Why?

Hypothesised constraint: The guide was human like the mountain climber.

Solution: The guide is inanimate such as a book or map.

Problem C: KING & QUEEN

Two sisters along with a large group of people watched as the queen attacked the king. No one said anything. Why?

Hypothesised constraint: The King and Queen were royalties.

Solution: The King and Queen were chess pieces and the game had reached checkmate.

Category: Ambiguous names associated with animals

Problem D: ANTHONY & CLEOPATRA

Anthony and Cleopatra are lying dead on the floor in an Egyptian villa. Nearby is a broken bowl. There are no marks on their bodies and they were not poisoned. Not a person was in the villa when they died. How did they die? (Sloane, 1992, p. 13)

Hypothesised constraint: Anthony and Cleopatra were human.

Solution: Anthony and Cleopatra were goldfish. They died when their bowl was knocked over by a clumsy guard dog.

Problem E: MR JONES

Mr Jones broke his leg on Saturday afternoon. He was immediately attended to by expert medical practitioners, and suffered no other injury. Sadly, he died later that day as a result. Why?

Hypothesised constraint: Mr Jones was human.

Solution: Mr Jones was a race horse who had to be put down after breaking his leg.

Problem F: JASON

Jason is lying dead. He has a piece of metal across his back and some food in front of him.

Hypothesised constraint: Jason was human.

Solution: Jason was a mouse that got caught in a mouse trap when attempting to eat the cheese in the trap.

Category: Out of Scope

Problem G: BOMBS AWAY

One night during the Second World War, an allied bomber was on a mission over Germany. The plane was in perfect condition and everything on it worked properly. When it had reached its target, the pilot ordered the bomb doors to be opened. They opened. He then ordered the bombs to be released. They were released. But the bombs did not fall from the plane. Why should this be so? (Sloane, 1992, p. 8)

Hypothesised constraint: The plane was flying the right way up.

Solution: The plane was flying upside-down.

Problem H: ROPE

A prisoner was attempting to escape from a tower. He found in his cell a rope that was half long enough to permit him to reach the ground safely. He divided the rope in half, tied the two parts together, and escaped. How could he have done this?

Hypothesised constraint: The rope was cut width ways therefore it remained the same length.

Solution: The rope was unravelled and two pieces were tied together.

Problem I: SONS

A woman had two sons who were born on the same hour of the same day of the same year. But they were not twins, and they were not adopted. How could this be so?

Hypothesised constraint: The sons were twins.

Solution: The sons were two of a set of triplets.

Appendix B

Experiment 2 – Training Programme

The training programme was similar to that used in Experiment 2 that covered training in a category of problem concerning ambiguous words and involved the following two stages:-

1. Solving problems that contained ambiguous words. This entailed:
 - a) awareness of ambiguous words in a problem and provision of the first heuristic.
 - b) practice in using this heuristic to solve a problem containing ambiguous words.
2. Practice solving problems containing ambiguous words.

For awareness training in ambiguous words (Stage 1 a), participants read an example of a problem that contained an ambiguous word and were given the heuristic 'If you cannot make sense of the problem then search for and identify any ambiguous word(s) and its alternative meaning(s)':-

A man is found dead in the arctic with a pack on his back. How did he die?

In this problem people might assume that the pack refers to a back pack. Participants were told that 'pack' was the ambiguous word and asked to think of alternative meanings for this word and to solve the problem. Participants then read that if the word 'pack' was thought of as a group of animals, then it would be the case that there was a pack of wolves on the man's back that were eating him. This interpretation would explain how the man died.

Next, participants were required to practise using the heuristic to solve a problem containing ambiguous words (Stage 1b):-

A man walked into a bar, and before he could say a word, he was knocked unconscious. Why?

In this problem the word 'bar' is ambiguous in meaning. The man had walked into a metal bar rather than a drinking bar.

In the final part of the training participants were presented with two problems (Stage 2). Participants were required to apply the training they had received to identify any words that were ambiguous and to think aloud whilst solving the problems. The problems were:

Problem 1

Two sisters along with a large group of people watched as the queen attacked the king. No one said anything. Why?

Problem 2

One morning a woman's earring fell into a cup that was filled with coffee, yet her earring did not get wet. How could this be?

Problem 1 contained the ambiguous words 'queen' and 'king', and thus referred to chess pieces, hence why the sisters and the group of people said nothing. Problem 2 contained the ambiguous word 'coffee' which refers to coffee granules rather than liquid coffee, hence why the earrings did not get wet.

Appendix C

Experiment 2 - Test Problems (Original version)

Category: Ambiguous words

Problem A: GUIDE

A mountain climber in the Himalayas took along with him two mountain guides. After a few hours, one of the guides fell into a deep crevasse. The climber and the other guide continued the climb and did not raise the alarm. Why?

Hypothesised constraint: The guide was human like the mountain climber.

Solution: The guide is inanimate such as a book or a map.

Problem B: SHOOT

A woman shoots her husband. Then she holds him under water for over 5 minutes. Finally, she hangs him. But 5 minutes later they both go out together and enjoy a wonderful dinner together. How can this be?

Hypothesised constraint: Shoot means kill him.

Solution: She took a photograph of him and developed it.

Problem C: COIN

While on safari in the wild jungles of Africa, Professor White woke one morning and felt something in the back pocket of her shorts. It had a head and a tail but no legs. When White got up she could feel it move inside her pocket. White however showed little concern and went about her morning rituals. Why such a casual attitude toward the thing in her pocket?

Hypothesised constraint: An animal e.g., a monkey as 'head and tail' is misleading.

Solution: A coin

Problem D: ISLAND

A woman is dead on an island and there is nothing else on the island with her. How did she die?

Hypothesised constraint: 'Island' means a tropical island.

Solution: She was hit by a car and is on a traffic island.

Category: 'Out of scope'

Problem E: WATER TOWER

A painter was hired to repaint the water tower for the township of Dubbo. The tower was located just off a busy street in the downtown area. Every day for 2 weeks the man painted diligently, but was never seen working by anyone and no change was noticed on the tower. At the end of the two weeks, the painter was thanked and paid a large sum by the city. Why?

Hypothesised constraint: He painted outside the building.

Solution: He painted the inside of the tower.

Problem F: TWINS

Marsha and Marjorie were born on the same day of the same month of the same year to the same mother and the same father - yet they are not twins. How is that possible?

Hypothesised constraint: The sons were twins.

Solution: They are triplets.

Problem G: CAPTAIN

Captain Scott was out for a walk when it started to rain. He did not have an umbrella and he wasn't wearing a hat. His clothes were soaked yet not a hair on his head got wet. How could this happen?

Hypothesised constraint: He has hair.

Solution: He is bald.

Appendix D

Experiment 2 - Test Problems (Lengthened version)

Category: Ambiguous words

Problem A: GUIDE

A mountain climber in the Himalayas took along with him two mountain guides. He had been looking forward to his climb for a long time. In his bag he took rations of food, flasks of coffee and water and a first-aid kit. After a few hours, one of the guides fell into a deep crevasse. The climber and the other guide continued the climb and did not raise the alarm. Why?

Hypothesised constraint: The guide was human like the mountain climber.

Solution: The guide is inanimate such as a book or map.

Problem B: SHOOT

A woman shoots her husband. Then she holds him under water for over 5 minutes. Finally, she hangs him. She made herself a cup of sweet tea and cleaned the kitchen, whilst he remained where she had left him. But 5 minutes later they both go out together and enjoy a wonderful dinner together. How can this be?

Hypothesised constraint: Shoot means kill him.

Solution: She took a photograph of him and developed it.

Problem C: COIN

While on safari in the wild jungles of Africa, Professor White woke one morning and felt something in the back pocket of her shorts. It had a head and a tail but no legs. The night before she had simply made herself supper, changed into her night clothes and went to bed. She didn't remember leaving anything in her pocket. When White got up she could feel it move inside her pocket. White however showed little concern and went about her morning rituals. Why such a casual attitude toward the thing in her pocket?

Hypothesised constraint: That an animal, such as a monkey, is in her pocket because of the misleading 'head and tail'.

Solution: A coin

Problem D: ISLAND

A woman is dead on an island and there is nothing else on the island with her. The young woman was wearing jeans, and a red shirt and had been looking forward to her day so was excited before she died. How did she die?

Hypothesised constraint: That island means a tropical island

Solution: She was hit by a car and is on a traffic island.

Category: 'Out of scope'

Problem E: WATER TOWER

A painter was hired to repaint the water tower for the township of Dubbo. In the small quaint town, there were a few coffee shops and clothing boutiques, a chapel, and a public library. The tower was located just off a busy street in the downtown area. Every day for 2 weeks the man painted diligently, but was never seen working by anyone and no change was noticed on the tower. At the end of the two weeks, the painter was thanked and paid a large sum by the city. Why?

Hypothesised constraint: He painted outside the building.

Solution: He painted the inside of the tower.

Problem F: TWINS

Marsha and Marjorie were born on the same day of the same month of the same year to the same mother and the same father - yet they are not twins. On the day they were

born the sun was shining through the hospital window; the view looked over the local park, which was filled with trees. How is that possible?

Hypothesised constraint: The daughters were twins.

Solution: They are triplets.

Problem G: CAPTAIN

Captain Scott was out for a walk when it started to rain. He did not have an umbrella and he wasn't wearing a hat. Other people around him started to run for cover under local shop fronts and bus shelters. His clothes were soaked yet not a hair on his head got wet. How could this happen?

Hypothesised constraint: He has hair.

Solution: He is bald.

Appendix E

Experiment 3- Original format of functional fixedness problems and alterations made for the experiment

Original 'visual' format of functional fixedness problems	Altered version used for experiment
Training Problems	
<p><i>String problem:</i> "Each S was told that we wished to hang a couple of strings from the ceiling but that we could not deface the ceiling and therefore a little construction was required" (Maier, 1945, p. 351)</p>	<p>Several wooden poles, clamps, and string have been made available. The task is to hang the string from the ceiling to the floor without defacing the ceiling.</p>
<p><i>Paperclip problem:</i> "A piece of white cardboard with four black squares fastened to it is to be hung on an eyelet screwed into the low ceiling. On the table lie paperclips, among other things." (Duncker, 1945, p. 87)</p>	<p>A piece of white cardboard with four black squares fastened to it is to be hung from a ring fixed to the ceiling. On the table in the room are the following objects available: paper, a pen, a ruler and some paperclips. How could the cardboard squares be hung on the ring? The word 'eyelet' was changed to 'ring' for better understanding. Objects were selected from a list provided by Duncker (1945).</p>
<p><i>Gimlet problem:</i> "Three cords are to be hung side by side from a wooden ledge. On the table lie, among many other objects, two short screw-hooks and the crucial object: a gimlet." (Duncker, 1945, p. 86)</p>	<p>Three cords are to be hung side by side from a wooden ledge. On the table in the room there is paper, pencils, tinfoil, two short screw-hooks and a hand powered screwdriver. How could the three cords be hung up? The term 'gimlet' was replaced with 'hand powered screwdriver'. Objects were selected from a list provided by Duncker (1945).</p>
Test problems:	
<p><i>Two-String problem:</i> "The experiment was carried on in a large room which contained many objects such as poles, ringstands, clamps, pliers, extension cords, tables and chairs. Two cords were hung from the ceiling, and were of such length that they reached the floor. One hung near a wall, the other from the centre of the room. The subject was told, "Your problem is to tie the ends of those two strings together." He soon learned that if he held either cord in his hand he could not reach the other. He was then told that he could use or do anything he wished." (Maier, 1931, p.182)</p>	<p>In a room two strings are hanging from the ceiling. The distance between them makes it impossible to reach one string while holding the other. The task is to reach one string while holding the other. A variety of objects are available, including a chair, paper, a pair of pliers, drawing pins and a jar. How may the two strings be tied together? Objects selected were based on those in a picture presented in Isaak and Just (1995).</p>
<p><i>Candle problem:</i> "On the door, at the height of the eyes, three small candles are to be put side by side. On the table, lie, among many other objects, a few tacks, three little pasteboard boxes." (Duncker, 1945, p. 86)</p>	<p>Your goal is to attach a candle to a wall so that it can burn upright. You have available a candle, some matches and a box of drawing pins. How would you solve the problem? 'Book of matches' was changed to 'some matches' and 'box of tacks' was changed to 'box of drawing pins' to avoid confusion.</p>
<p>The verbal form was based on the problem</p>	

presented in Chrysikou (2006):

“Your goal is to attach a candle to a wall so that it can burn upright. You have available a candle, a book of matches and a box of tacks. How would you solve the problem?” (Chrysikou, 2006, p. 937)

Hat Rack Problem:

“The S was ... asked to construct a hat rack in a certain spot. The available equipment consisted of two 1 inch X 2 inch poles (one 6 and the other 7 ft. long) and a 3-inch 'C clamp. A hat rack was then defined as something sturdy enough to properly hold a heavy coat and hat.” (Maier, 1945, p. 352)

Using two poles and a clamp, build a hat rack which is sufficiently stable to support a heavy coat and a hat. The opening of the clamp is wide enough so that both poles can be inserted and held together securely when the clamp is tightened.

Out of scope problems:

Charlie problem:

Dan comes home from work and finds Charlie lying dead on the floor. Also on the floor are some broken glass and some water. Tom is in the room too. Dan takes one look around and immediately knows how Charlie died. How did Charlie die? (Weisberg, 1995)

Hypothesized constraint: Charlie is a human so think of various murder scenarios involving two humans (i.e., Tom shot or stabbed Charlie).

Solution: Charlie, Dan’s pet fish, died of lack of oxygen when Tom, Dan’s cat, knocked over the fishbowl, causing it to shatter and spill its contents.

Fake Coin problem:

A dealer in antique coins got an offer to buy a beautiful bronze coin. The coin had an emperor’s head on one side and the date 544 BC stamped on the other. The dealer examined the coin and realized it was a fake. How did he know the coin was phoney? (Ansburg & Dominowski, 2000)

Hypothesized constraint: Consider the elements on the coin to determine what might have been faked. Was bronze invented? The date and the emperor might not have matched.

Solution: Christ would not have been born, so a coin from that time would not have been marked BC.

Prisoner and Rope problem:

A prisoner was attempting to escape from a tower. He found in his cell a rope that was half long enough to permit him to reach the ground safely. He divided the rope in half, tied the two parts together, and escaped. How could he have done this? (Isaak & Just, 1995)

Hypothesized constraint: That the rope was cut in half across the width so there were two shorter pieces.

Solution: The rope was unravelled or cut in half vertically so there were two pieces half long enough which could be tied together.

Appendix F

Experiment 4 – Test problems

Problem A: UNSEEN WALKER

On a busy Friday afternoon, a man walked several miles across London from Westminster to Knightsbridge without seeing anybody or being seen by anybody. The day was clear and bright. He had perfect eyesight and he looked where he was going. He did not travel by any method of transport other than by foot. London was thronged with people yet not one of them saw him. How? (Sloane, 1992, p. 11)

Hypothesised constraint: The man was walking above ground along the streets.

Solution: The man was walking underground through the sewers.

Problem B: ANTHONY & CLEOPATRA

Anthony and Cleopatra are lying dead on the floor in an Egyptian villa. Nearby is a broken bowl. There are no marks on their bodies and they were not poisoned. Not a person was in the villa when they died. How did they die? (Sloane, 1992, p. 13)

Hypothesised constraint: Anthony and Cleopatra were human.

Solution: Anthony and Cleopatra were goldfish. They died when their bowl was knocked over by a clumsy guard dog.

Problem C: COMING HOME

A man walked home after having been out drinking. He walked down the middle of a deserted country road. There were no streetlights to illuminate the road and there was no moonlight. He was dressed all in black. Suddenly a car that did not have its headlights on came racing down the road. At the last moment, the driver of the car saw the man and swerved to avoid him. How did he manage to see him? (Sloane, 1992, p. 20)

Hypothesised constraint: It was night-time.

Solution: It was daytime.

Problem D: BOMBS AWAY

One night during the Second World War, an allied bomber was on a mission over Germany. The plane was in perfect condition and everything on it worked properly. When it had reached its target, the pilot ordered the bomb doors to be opened. They opened. He then ordered the bombs to be released. They were released. But the bombs did not fall from the plane. Why should this be so? (Sloane, 1992, p. 8)

Hypothesised constraint: The plane was flying the right way up.

Solution: The plane was flying upside-down.

Appendix G

Experiment 5 - Training Programme

The first part of the training was broken down into two stages. In Stage 1a, trainees were informed that problems are difficult to solve because of incorrect interpretations and that they will be trained in a mechanism to help them identify their incorrect interpretations in order to solve such problems. The mechanism was described as follows:

‘After reading the problem carefully, select a part of the problem and consider your interpretation of it. Then check whether your interpretation is consistent by comparing it with the information in the problem specification. If it is inconsistent, then select another part of the problem and repeat the above process. However, if you generate a consistent interpretation, then consider whether this interpretation can be used to solve the problem. If not, then repeat the above process until you have generated a solution to solve the problem.’

In Stage 1b, an example of how to apply the mechanism to solve the following problem was presented:-

Sid Shady works for a large construction company that was very concerned about employee theft. Someone tipped the company that Shady was the man to watch. Each night he passed through security with a wheelbarrow full of scrap lumber, discarded electrical wires and chunks of concrete. The security guards checked the contents daily but could find nothing of value. What was Shady stealing? (Ansburg & Dominowski, 2000)

To answer the question, a possible interpretation concerning what Shady was stealing is that something was concealed on his person. However, this is inconsistent with the problem statement as it does not explain the taking of the lumber, wires and concrete. Another interpretation was that Shady was using the lumber, wires and concrete for a useful purpose but this is also inconsistent as they are of no value and the guard was aware that he was taking them. An alternative interpretation was that Shady was concealing something within the contents of the wheelbarrow, which is also inconsistent because the statement states that the wheelbarrow is checked daily. The correct solution is that solution is stealing wheelbarrows, which is consistent because the statement states that Shady leaves with a wheelbarrow each night.

The second part of the training was also broken down into two stages. After reading the above example, participants were required to practise using the mechanism to solve the following problem (Stage 2a):-

Barney Dribble is carrying a pillow case full of feathers. Hardy Pyle is carrying three pillow cases the same size as Barney’s, yet Hardy’s load is lighter. How can this be? The solution was that Hardy’s cases were empty, and hence lighter.

In the final part of the training, participants were required to solve the following problem under test conditions (Stage 2b):-

A woman said to her husband “This morning, one of my earrings fell into my coffee. Even though my cup was full, the earring did not get wet.” How could this be true? The coffee cup contained coffee granules and thus the earrings did not get wet.

Appendix H

Experiment 5 – Practice and Test problems

Practice problems

Problem A: Sid Shady

Sid Shady works for a large construction company that was very concerned about employee theft. Someone tipped the company that Shady was the man to watch. Each night he passed through security with a wheelbarrow full of scrap lumber, discarded electrical wires and chunks of concrete. The security guards checked the contents daily but could find nothing of value. What was Shady stealing? (Ansburg & Dominowski, 2000)

Solution: Shady was stealing wheelbarrows.

Problem B: Barney Dribble

Barney Dribble is carrying a pillow case full of feathers. Hardy Pyle is carrying three pillow cases the same size as Barney's, yet Hardy's load is lighter. How can this be?

Solution: Hardy's cases were empty, and hence lighter.

Problem C: Coffee

A woman said to her husband "This morning, one of my earrings fell into my coffee. Even though my cup was full, the earring did not get wet." How could this be true?

Solution: The coffee cup contained coffee granules and thus the earrings did not get wet.

Test Problems

Problem A: PEAR TREE

A farmer in California owns a beautiful pear tree. He supplies the fruit to a nearby grocery store. The store owner has called the farmer to see how much fruit is available for him to purchase. The farmer knows that the main trunk has 24 branches. Each branch has exactly 6 twigs. Since each twig bears one piece of fruit, how many plums will the farmer be able to deliver?

Solution: None because plums do not grow on pear trees.

Problem B: DR APPLE

Shadow opened the door to Dr. Apple's office and surveyed the scene. Dr. Apple's head lay on his desk in a pool of blood. On the floor to his right lay a gun. There were powder burns on his right temple indicating that he was shot at close range. On his desk was a suicide note, and in his right hand was the pen that had written it. Shadow noted that death had occurred in the last hour. All of a sudden Dr. Apple's wife burst into the office and screamed "My husband's been shot!" She ran toward the body and saw the note and cried, "Why would he want to kill himself?"

Shadow replied "This was no suicide; it is a clear case of murder." How does Shadow know?

Solution: The gun would have been the last item in Dr Apple's hand if he had committed suicide.

Problem C: TRAIN

At 7 a.m., a train moving 90mph leaves Montreal heading for Toronto. At 8 a.m., a train running 110mph leaves Toronto heading for Montreal. Which train will be closer to Montreal when they meet?

Solution: When they meet, they would be at the same spot.

Problem D: DIRECTORY

There is a town in Northern Ontario where 5% of all the people living in the town have unlisted phone numbers. If you selected 100 names at random from the town's phone directory, on average, how many of these people selected would have unlisted phone numbers?

Solution: None because a phone directory contains listed numbers only.

Problem E: ANTIQUE COIN

A dealer in antique coins got an offer to buy a beautiful bronze coin. The coin had an emperor's head on one side and the date 544 BC stamped on the other. The dealer examined the coin and realized it was a fake. How did he know the coin was phoney?

Solution: Christ had not been born, therefore a coin from that time would not be marked BC.

Problem F: PROFESSOR BUMBLE

Professor Bumble, who is getting on in years, is growing absent-minded. On the way to a lecture one day, he went through a red light and turned down a one-way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behaviour?

Solution: Professor Bumble was walking.

Problem G: LIGHT

A young boy turned off the light in his bedroom and managed to get into bed before the room was dark. If the bed is ten feet from the light switch and he used no wires, strings or any other contraptions to turn off the light, how did he do it?

Solution: It was daytime.

Appendix I

Experiment 6 – Pilot study

Problem 1: UNSEEN WALKER (77 words)

A man walked home after having been out drinking. He walked down the middle of a deserted country road. There were no streetlights to illuminate the road and there was no moonlight. He was dressed all in black. Suddenly a car that did not have its headlights on came racing down the road. At the last moment, the driver of the car saw the man and swerved to avoid him. How did he manage to see him?

Solution: It was day-time (target word: night)

Problem 2: MURDER (74 words)

Acting on an anonymous phone call, the police raid a house to arrest a suspected murderer. They don't know what he looks like but they know his name is John and that he is inside the house. The police bust in on a carpenter, a lorry driver, a mechanic and an engineer all playing poker. Without hesitation or communication of any kind, they immediately arrest the engineer. How did they know who to arrest?

Solution: The engineer was the only male playing poker. (target word: men)

Problem 3: PROFESSOR BUMBLE (55 words)

Professor Bumble, who is getting on in years, is growing absent-minded. On the way to a lecture one day, he went through a red light and turned down a one-way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behaviour?

Solution: He was walking (target word: driving)

Problem 4: SHOT (75 words)

During the world fair a group of scientists were exhibiting their advances in genetic engineering. There were cross-breeds of various bulls, cows, and other domestic farm animals. Featured on the exhibit were several over-sized prized turkeys. One afternoon during the show, a woman walked up to the exhibit, shot the turkeys, and then ran out of the building. Although she was known to a number of people, nobody made any attempt to stop her. Why?

Solution: Women shot the turkeys with a camera. She was a journalist. (target word: killed)

Problem 5: JUMPED (60 words)

Mel Colly stared through the dirty soot-smearred window on the 26th floor of the office tower. Overcome with depression he slid the window open and jumped through it. It was a sheer drop to the ground. Miraculously after he landed he was completely unhurt. Since there was nothing to cushion his fall or slow his descent, how did he survive?

Solution: Jumped inside. (target word: outside)

Problem 6: GUIDE (42 words)

A mountain climber in the Himalayas took along with him two mountain guides. After a few hours, one of the guides fell into a deep crevasse. The climber and the other guide continued the climb and did not raise the alarm. Why?

Solution: Guide was a map (target word: person)

Problem 7: BEACH (61 words)

Rachel was sun bathing at a beach because she heard it was the best way to acquire an overall tan. The beach was full of people wearing nothing. Day after day she was out lying in the sun, yet couldn't manage the complete tan. What could she be doing wrong to prevent her from receiving the even tones she so desired?

Solution: She still had her bathing suit on. (target word: naked)

Appendix J

Experiment 6 – Test problems

Problem A: UNSEEN WALKER

A man walked home after having been out drinking. He walked down the middle of a deserted country road. There were no streetlights to illuminate the road and there was no moonlight. He was dressed all in black. Suddenly a car that did not have its headlights on came racing down the road. At the last moment, the driver of the car saw the man and swerved to avoid him. How did he manage to see him?

Target word: Night

Solution: It was day-time

Problem B: MURDER

Acting on an anonymous phone call, the police raid a house to arrest a suspected murderer. They don't know what he looks like but they know his name is John and that he is inside the house. The police bust in on a carpenter, a lorry driver, a mechanic and an engineer all playing poker. Without hesitation or communication of any kind, they immediately arrest the engineer. How did they know who to arrest?

Target word: Men

Solution: The carpenter, lorry driver, and mechanic were women. The engineer was the only male playing poker.

Problem C: PROFESSOR BUMBLE

Professor Bumble, who is getting on in years, is growing absent-minded. On the way to a lecture one day, he went through a red light and turned down a one-way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behaviour?

Target word: Driving

Solution: Professor Bumble was walking

Problem D: SHOT

During the world fair a group of scientists were exhibiting their advances in genetic engineering. There were cross-breeds of various bulls, cows, and other domestic farm animals. Featured on the exhibit were several over-sized prized turkeys. One afternoon during the show, a woman walked up to the exhibit, shot the turkeys, and then ran out of the building. Although she was known to a number of people, nobody made any attempt to stop her. Why?

Target word: Killed

Solution: Women shot the turkeys with a camera. She was a journalist.

Problem E: GUIDE

A mountain climber in the Himalayas took along with him two mountain guides. After a few hours, one of the guides fell into a deep crevasse. The climber and the other guide continued the climb and did not raise the alarm. Why?

Target word: Person

Solution: The guide is inanimate such as a book or map.

Appendix K

Experiment 6 - Neutral primes (length matched with the length of the corresponding experimental problem)

1. http://www.ninds.nih.gov/disorders/brain_basics/know_your_brain.htm (UNSEEN WALKER problem - 77 words)

The brain is undoubtedly the most complex part of the human body. For many centuries, scientists and philosophers have been fascinated by the brain, but until recently they viewed the brain as nearly incomprehensible. Now, however, the brain is beginning to relinquish its secrets. Scientists have learned a lot more about the brain in the last 10 years because of the increasing growth of research in neurological and behavioural science and the development of new research techniques.

2. <http://www.dmu.ac.uk/~jamesa/teaching/assessment.htm> (MURDER problem - 74 words)

It is the use of assessment which makes teaching into teaching. Assessment should not therefore be seen as an isolated process, but as integral to every stage of teaching. One point to note is that there is no such thing as an "objective test". Even when there's a high degree of standardisation, the choice of what things are tested and what constitutes a criterion of satisfactory performance is very much dependent on the assessor.

3.

<http://www.see.ed.ac.uk/~gerard/Management/art1.html?http://oldeee.see.ed.ac.uk/~gerard/Management/art1.html> (PROFESSOR BUMBLE problem - 55 words)

In a newspaper, the story is introduced in its entirety in a snazzy first paragraph. The next few paragraphs repeat the same information only giving further details to each point. The next section repeats the story, but developing certain themes within each of the sub-points. This is repeated until the reporter runs out of story.

4. <http://www.onepine.info/mcult.htm> (SHOT problem - 75 words)

A key role for culture is to differentiate the organisation from others and to provide a sense of identity for its members. Cultures do not have to be logical, in fact they seldom are and can appear quite haphazard and chaotic to the outsider. Culture can also have subgroups with varying agendas. A strong culture is one that is widely shared, and makes it clear what it expects and how it wishes people to behave.

6. <http://www.cinemas-online.co.uk/films/core/earthfacts/facts.htm> (Guide problem - 42 words)

The earth consists of several layers. The three main layers are the core, the mantle and the crust. The core is the inner part of the earth, the crust is the outer part and between the core and crust is the mantle.

Appendix L
Experiment 6 LDT sets

Practice LTD

Words	Mean RT	Nonwords
overpass	631.16	afoustic
initials	631.56	panknote
promises	632.03	dootlace
handlers	632.52	tecisive
majority	632.53	daughtir
forwards	633.09	crominal

1. UNSEEN WALKER problem

Words	Mean RT	Nonwords
jokes	531.58	avide
hired	540.09	lanks
necks	541.44	burge
lemon	541.70	narns
holds	542.47	vruly
night	540.39	Firch

2. MURDER problem

Words	Mean RT	Nonwords
hey	538.38	cof
ice	541.47	ald
fly	541.70	fub
red	542.76	bry
key	543.52	dut
men	540.55	bew

3. PROFESSOR BUMBLE problem

Words	Mean RT	Nonwords
recipes	673.29	addomen
folders	673.38	panquet
pigeons	673.42	sargain
rancher	673.76	millows
hugging	675.56	greadth
driving	673.44	breakop

4. SHOT problem

Words	Mean RT	Nonwords
hooked	545.79	arides
street	547.00	dalves
search	547.16	hanned
motors	550.71	ganter
second	553.58	faress
killed	551.50	darton

5. **OUTSIDE** problem

Words	Mean RT	Nonwords
likely	549.63	atmire
switch	549.97	shests
simply	551.35	phords
winter	552.24	glumisy
moment	554.37	pookie
outside	551.65	cozier

6. **GUIDE** problem

Words	Mean RT	Nonwords
labels	613.31	lastle
months	613.42	tasual
fibres	614.00	catnop
tricky	615.12	shamps
glossy	615.50	thapel
person	614.61	pheese

7. **BEACH** problem

Words	Mean RT	Nonwords
spice	591.09	theap
lodge	592.84	phoir
ovens	594.16	agout
spear	594.60	narns
jelly	595.47	bolks
naked	594.80	dands

