

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**METACOGNITIVE BELIEFS AND EXPERIENCES:  
BELIEFS, PREDICTIONS, MONITORING AND EVALUATIONS  
IN WELL-DEFINED AND INSIGHT  
PROBLEM SOLVING**

Thesis presented in partial fulfillment  
of the requirements for the degree  
of Master of Arts in Psychology  
at Massey University,  
Palmerston North,  
New Zealand

**SHANE PALMER**

**2002**

## ACKNOWLEDGEMENTS

The process of completing my research and thesis has provided its share of challenges, difficulties, and rewards. Although there is only one name on the cover, many people have been critical in providing their assistance and support to me during this project. I would like to offer my sincerest appreciation to you all:

First and foremost, to my primary supervisor Dr. Julie Bunnell, who, despite new responsibilities of her own, willingly and warmly provided her time and expertise in all aspects of my thesis, particularly on the conceptual and analytic issues, the administrative organisation (including help in obtaining the PSI, parking permits, research space and much else) and proofing of the write-up when it finally arrived! I am especially appreciative of Julie's understanding and compassion, and for her encouragement at all, including the most difficult, times.

Secondly, to my secondary supervisor Dr. John Podd, who provided much-appreciated back-up, and whose availability to provide help whenever needed was invaluable. John was particularly helpful in dealing with the analytic details, and in helping me to consider more clearly some of the conceptual issues that arose.

To other members of the School of Psychology: especially Harvey Jones for his technical and programming expertise in intuitively designing the computer program that was the basis for everything! Also, Malcolm Loudon for his calculation of the angular warmth measures. And Michael Donnelly for administrative support in procuring valuable research space and allowing me to pay participants!

There are a host of friends and fellow students who I thank for both intellectual and emotional support: Steve Stewart-Williams (our deep psychological and philosophical chats were enjoyable *and* useful!), and fellow Masters students along the way, Ben Heaven (having fun debating merits of quant. vs. qual., and a genuine good bloke!), Sharon Leadbetter (a true friend and inspiration since first year!), Toni Hyde, Nicola Addis, everyone hanging out in the grad room, and to all my friends – you know who you are!

My research would not have been possible without the participants who generously gave their time and concentration in taking part – despite being baffled by many problems! I thank you all for your willingness and effort, for making the research a great experience.

My family has given me exceptional and constant encouragement, and I thank them all for the support that made things easier to deal with. Thanks to my extended family: Rex, Val, Mark, Sam, and little Jevan. Special thanks to my mum, Susan, for all your years of love, support and guidance, and providing me with everything I needed to get this far in the first place!

And finally, to my wonderful wife Rosemary, who has been there with me through it all and survived! Thank you for unending support, encouragement, inspiration, much needed motivation through good and hard times – but mostly for your love, for being a special partner and friend, and making it all worthwhile.

## ABSTRACT

The present study attempted to replicate and extend understandings of differences in the metacognitive experiences of solving insight and well-defined problems. Insight often occurs with a sudden ‘Aha!’ reaction compared to the more continuous progress typical for well-defined problems. Thirty-two adults completed a within-subjects computer-based problem solving task involving sets of 8 insight and well-defined problems, while providing predictions, feeling-of-warmth monitoring, and evaluations of performance. A sub-sample completed a Problem Solving Inventory (PSI) to compare global and context-specific beliefs of ability. Predictions overestimated performance in both sets, but more so for insight than for well-defined problems. However, correlations between prediction and performance were not significant for either set. No consistent difference in monitoring was found; incremental patterns dominated insight and well-defined problems equally. Averaged evaluations mirrored the overestimation effects of the predictions, although distributions of confidence accuracy were similar across sets. However, interesting correlations were found between global PSI scores and the specific measures, for both problem types. Methodological differences between the present and earlier studies may account for the lack of problem set effects. Conceptual issues need to be addressed regarding definition of insight and verification of insight experiences, particularly if future research is to reconcile metacognitive and cognitive aspects of problem solving.



## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
ABSTRACT .....	iii
TABLE OF CONTENTS .....	iv
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
<b>PREFACE</b> .....	1
<b>INTRODUCTION</b> .....	3
<b>PROBLEM SOLVING</b> .....	3
1. Well-defined and Ill-defined Problems .....	4
2. Insight Problem Solving .....	5
a. Impasse in insight .....	8
b. Incubation in insight .....	9
<b>METACOGNITION IN PROBLEM SOLVING</b> .....	11
1. Metacognitive Knowledge and Executive Control .....	14
a. Metacognitive knowledge .....	14
b. Executive control: Monitoring and regulation .....	16
2. Metacognitive Experiences .....	17
3. Metacognitive Information Processing In Problem Solving .....	19
a. Identification and problem finding .....	19
b. Problem representation and solution prediction .....	20
c. Monitoring and evaluation .....	22
4. Metacognitive Monitoring And Verbalization .....	23
<b>METACOGNITIVE EXPERIENCE IN INSIGHT</b> .....	26
1. Discontinuous Versus Continuous Metacognition .....	27
a. Feelings-of-knowing .....	27
b. Feelings-of-warmth .....	29

2. Conscious Versus Unconscious Processing .....	32
<b>THE PRESENT STUDY</b> .....	36
Objectives .....	38
Hypotheses .....	38
<b>METHOD</b> .....	42
Participants .....	42
Materials .....	42
1. Problems .....	42
2. Problem Solving Inventory (PSI) .....	44
Apparatus .....	45
Procedure .....	45
1. Preliminary Procedures .....	45
2. Problem Presentation .....	46
3. Second Session .....	48
<b>RESULTS</b> .....	49
1. Performance .....	49
a. Solution accuracy .....	49
b. Time to solution .....	50
c. Individual performance .....	51
2. Predictions: Confidence and Predictive Accuracy .....	52
a. Confidence .....	52
b. Predictive accuracy .....	53
3. Metacognitive Monitoring: Feeling of Warmth Analysis .....	55
a. Mean FOW ratings across solution intervals .....	56
i. Warmth as a function of correctness in insight problems .....	57
ii. Warmth as a function of correctness in well-defined problems .....	57
iii. Warmth as a function of problem type .....	58
b. Insight/incremental pattern analyses .....	59
c. Differential and angular warmth correlations .....	62

4. Evaluations: Confidence and Evaluative Accuracy .....	63
a. Absolute evaluations .....	63
b. Relative evaluations .....	65
5. PSI Analysis: General Versus Specific Metacognitions .....	67
a. General problem solving confidence and inter-scale correlations ..	67
b. General versus specific metacognitions: Predictions .....	69
c. General versus specific metacognitions: Evaluations .....	69
<b>DISCUSSION</b> .....	71
PROBLEM SOLVING PERFORMANCE .....	71
PREDICTIONS .....	74
1. Predictive Confidence .....	74
2. Predictive Accuracy .....	75
a. Macro-predictive accuracy: Comparing estimate and performance proportions .....	75
b. Micro-predictive accuracy: Correlating predictions and performance .....	77
MONITORING: FEELINGS OF WARMTH .....	82
EVALUATIONS .....	88
1. Evaluation Accuracy .....	88
2. Utility Of Evaluation Measures .....	91
GENERAL VERSUS SPECIFIC METACOGNITIONS .....	93
1. PSI-Prediction and PSI-Evaluation Correlations .....	93
2. Differences Between PSI Scales .....	95
LIMITATIONS OF THE PRESENT STUDY .....	97
1. Methodological Limitations .....	98
2. Conceptual Limitations .....	100
a. Defining insight and insight problems .....	100
b. Verifying insight experiences .....	102
FUTURE RESEARCH .....	104
SUMMARY AND CONCLUSIONS .....	107

<b>REFERENCES .....</b>	<b>111</b>
-------------------------	------------

<b>APPENDICES .....</b>	<b>126</b>
-------------------------	------------

A Participation Forms .....	126
-----------------------------	-----

B Materials .....	134
-------------------	-----

C Additional Results Tables, Figures, And Notes .....	145
---	-----

LIST OF TABLES

Table 1. Mean (M) predictions for incorrectly and correctly solved problems in well-defined and insight sets ..... 53

Table 2. Mean (M) warmth ratings in last four intervals for correct and incorrect solutions in insight and well-defined sets, for problems solved within 4, 5, or 6 intervals only ..... 57

Table 3. Mean (M) warmth ratings in last four intervals for correctly solved insight and well-defined problems ..... 58

Table 4. Proportion of problems in insight set showing an insight or incremental pattern of warmth ratings, for low- and high-anchored protocols ..... 60

Table 5. Proportion of problems in well-defined set showing an insight or incremental pattern of warmth ratings, for low- and high-anchored protocols ..... 60

Table 6. Mean (M) absolute evaluations and correct solutions for well-defined and insight sets ..... 64

Table 7. Frequency of over-confident, accurately confident, and under confident absolute evaluations within well-defined and insight sets ..... 65

Table 8. Frequency of over-confident, accurately confident, and under confident relative evaluations within well-defined and insight sets ..... 67

Table 9. PSI scale means (M) and standard deviations (SD) ..... 68

Table 10. Correlations between PSI scale scores and prediction measures ..... 68

Table 11. Correlations between PSI scale scores and absolute and relative evaluations ..... 70

**LIST OF FIGURES**

Figure 1. Percentage of solutions for well-defined and insight problem sets .....	50
Figure 2. Mean number of FOW ratings made across problem set and solution accuracy .....	51
Figure 3. Mean prediction probability estimates and proportions of correct solutions for well-defined and insight sets .....	54
Figure 4. Distributions of relative evaluation ratings in well-defined and insight sets .....	66

## PREFACE

Problems of various kinds permeate most aspects of our everyday activities. Hence, problem solving is a fundamental and pervasive cognitive activity (Mayer, 1992), a necessary component in negotiating our daily lives. Solving these problems requires adequate understanding of what the problem situation entails, what steps must be taken to solve the problem, and knowledge of what strategies one may use to reach the goal of solution, as well as the ability to execute strategies to this end. Achievement of a desired solution often requires that one select the most appropriate and efficient strategies to fulfill the identified requirements, while regulating one's attempts at using these strategies in order to keep track of progress towards the goal, and identifying when the solution has been obtained. Furthermore, one's beliefs about problems and problem solving generally, together with both broad and context-specific beliefs about one's own competencies and abilities, may influence the course of one's solution efforts.

Metacognition may be one system through which personal beliefs and selective strategy application have a bearing on the accuracy of problem solving performance. Metacognition refers to a person's thinking about his or her thinking, through the higher-order processes of monitoring, regulating, and evaluating of ongoing cognitive processes (Flavell, 1978). As with other aspects of thinking, metacognition is considered to be a crucial influence in the efficiency and accuracy of people's problem solving activity. Theory and research in this area provide indications that problem solving processes are indeed facilitated by adequate metacognitive skills. While people differ in terms of the complexity and spontaneity of their metacognitive thinking, it appears that these skills can be enhanced through development and training (Hanley, 1995; Hayes, 1980; Simon, 1980). Therefore, metacognitive aspects of problem solving have both psychological and educational implications. The concept of metacognition is both meaningful and fruitful for our understanding of and attention to problem solving abilities.

The present study examines the relationships of metacognitive beliefs and experiences with the performance of problem solving activities. Both 'on-line' and 'off-line' assessments are used to assess the metacognitive knowledge and

experiences that participants have in relation to solving problems. On-line beliefs are measured in the form of predictions prior to solution attempts, monitoring during the solution process, and evaluations following completions of a problem set. Off-line beliefs are measured with the use of a Problem Solving Inventory (PSI) that assesses an individual's perceptions of his or her own general problem solving behaviours and attitudes. Furthermore, both on- and off-line beliefs are examined in relation to two general types of problems: well-defined and insight problems. Well-defined problems are typically solved in an incremental, step-by-step fashion towards a given goal. Insight problems typically encourage an obvious but incorrect method that leads to an impasse, which may be overcome by a sudden 'flash' of insight that quickly leads to the correct answer.

The following review examines the relevant literature in problem solving, particularly in relation to insight problems, and in metacognition, with attention on metacognitive beliefs and experiences in problem solving. Problem solving is discussed in terms of the commonly researched well-defined, ill-defined and insight problem structures. Insight is defined and discussed in the context of a classic problem solving model that relates the stages of preparation, impasse and incubation, illumination, and verification. Metacognition is discussed in terms of distinctions between knowledge, executive and procedural control, and affective experiences. These components are considered important to information processing models of problem solving through the metacognitive processes of identification, representation, planning, monitoring, and evaluating. Metacognitive representation and monitoring may be particularly important for solving insight problems; however, while research has demonstrated the positive effects of metacognition on well-defined problems, insight research has proven more complicated and controversial. Measures of subjective metacognitive experiences may increase our understandings of insight processes, although doubts remain as to whether insight involves rapid restructuring of knowledge, and whether unconscious or conscious processes are important in relation to metacognitive appraisals. These issues are debated, before an overview of the present study is presented.



## INTRODUCTION

### PROBLEM SOLVING

One needs only to consider briefly his or her daily life to realize the pervasive occurrence of problems across time and situation. Problems, differing in nature and severity, abound in whatever domains or contexts within which humans exist: for example, education, research, work place, home, leisure activities, and social relationships. In more demanding cases, problems may tax our abilities to handle the cognitive, emotional, and social demands required in our response to a situation, necessitating reliance on coping activities (Cassidy, 1999; Lazarus & Folkman, 1987). In all cases, the existence of a problem requires adaptation to some situation by negotiating obstacles or barriers that block our progress towards some goal; in short, problem solving.

Problems are distinguished from tasks; “mental demands for which the solution methods are known” (Doerner, 1979, cited in Jausovec, 1994) and that are executed solely from memory recall. Whether a given situation represents a task or a problem depends on the capacities and experience of the person in the situation. Most broadly, a problem exists for a person when “he wants something and does not know immediately what series of actions he can perform to get it” (Newell & Simon, 1972: p 72). Problem solving, then, involves goal-directed thinking aimed at overcoming the obstacles that hinder a person’s obtaining of some goal (Davidson & Sternberg, 1998). It is closely related to other cognitive activities, such as perception, attention, language comprehension, memory, decision-making, creative thinking and critical thinking (Swartz & Perkins, 1990). Together, these processes help us to engage and negotiate the situations and duties of daily life.

Conceptualisations of problems and problem solving are numerous (Jausovec, 1994). Newell and Simon’s (1972) ‘problem space’ model is perhaps the most formally explicit and generally applicable model of problems and their requisite solution processes. In this model, developed from an information-processing perspective, problem solution involves the interaction of a problem solver with a specific task environment. Solving processes are activated by the

identification and representation of a problem, followed by the selection and application of solution strategies. A person's representation is an internal 'mental model' of the external situation (Davidson, Deuser, & Sternberg, 1994). Problems are represented in terms of some 'problem space', incorporating an initial problem state, a desired goal state, operators or methods, and path constraints. The problem state is the point where problem solving begins, where one realises what problem exists and his or her desire to solve it. The goal state represents the endpoint or solution to be reached. Operators are the methods used to change the initial state and reach the goal state; path constraints include any rules or conditions that limit the operations used. The size of the problem space is determined by the amount of information covered in the representation, and hence the number of operations that can be applied towards an endpoint. Not all solutions may be considered desirable goals, however; search through a problem space may lead to an incorrect solution. Effective problem solution requires the application of operators that allow search through the problem space such that the size of the space is effectively reduced, until only the path to the desired goal remains. A desired solution usually requires either modification of the existing representation, or development of a new representation altogether. Central to the solving of problems, then, are the representations constructed of the situation and the strategies or operations applied to those representations.

### **1. Well-defined and Ill-defined Problems**

Newell and Simon's (1972) conception of problem space is particularly suited to problems that are well-structured. It may be less useful for other types of problem. The most common typology of problem structure distinguishes between well-defined and ill-defined problems (Gilhooley, 1988; Kitchener, 1983; Robertson, 1999). A well-defined problem exists if the elements of problem state, goal state, operators, and path constraints are clearly specified; for example, the problems in Appendix B are considered to be well-defined in nature. An ill-defined problem exists if any or all of these problem elements are vague or unspecified. For example, composing a poem, choosing a career or a marriage partner, finding means

to limit pollution are common ill-defined problems; a specific goal may not be defined, one of several feasible solutions may need to be chosen, or the path to a solution may not be easily specifiable. No doubt both problem types exist on a continuum, rather than as a strict dichotomy, of structural definition (Greeno, 1978). Nevertheless, while most research has focused on well-structured problem solving, it is widely acknowledged that most problems of daily life are ill structured in nature (Kahney, 1993; Kitchener, 1983; Kitchener & King, 1990).

The broad distinction between well- and ill-defined problems is not the only, nor indeed the most precise, taxonomy of problem types (see Jausovec, 1994 for a thorough review). Well-defined problems differ amongst each other in important respects, as do ill-defined problems. Therefore, a problem's characterisation may differ depending on the taxonomy that is used to classify it. Nevertheless, the well-defined versus ill-defined division emphasises a significant distinction in the classification of problems that is still considered a useful and meaningful distinction (Ashman & Conway, 1997; Matlin, 1998; Mayer, 1999; Robertson, 1999; Schraw, Dunkle & Bendixen, 1995).

## **2. Insight Problem Solving**

'Insight' refers to the sudden realization of a problem solution, often with a sudden change in one's understanding of the problem, and often with a 'flash' or a sense of surprise that prompts an 'Aha!' response (Davidson, 1995; Dominowski & Dallob, 1995; Metcalfe, 1986a, 1986b). The concept of insight has gained notoriety thanks to anecdotal evidence from biographies of historical figures who reputedly made astounding scientific discoveries or artistic creations through sudden insightful experiences; for example, Archimedes, Newton, and Darwin (Weisberg, 1986, 1993, 1995a, 1999). Insight is a particularly interesting form of problem solving to study given its purported links with creativity, its alleged role in many great discoveries, and because most people have experienced a 'flash' of insight at some point (Sternberg & Davidson, 1999). Also, the concept has historically been shrouded in a degree of mystery and controversy, given the numerous but difficult-to-verify explanations for its occurrence, often citing unconscious processes.

Systematic research into insight processes began with Gestalt psychologists who related the experience of insight in problem solving to perceptual processes involved in observing 'bistable' figures (for example, the Necker cube, the Janov duck-rabbit). These figures can be perceived in either of two forms, and perception of the figures often involves a sudden switch from one form to the other, with no apparent stable transition between the two forms. Gestalt theorists attributed this phenomenon to the holistic reorganization of the parts making up the figure, bringing meaningful order to the whole perceptual structure. In a similar fashion, insight in problem solving is achieved by 'seeing' a problem in a new way, or perceiving some coherent underlying structure (Mayer, 1999). Gestalt psychologists identified two broad types of problem solving: reproductive and productive (Kohler, 1969, cited in Dominowski, 1995). Reproductive thinking involves making use of previous experience, previously acquired knowledge or procedures in order to solve a current problem. The challenge is to identify the right knowledge or procedure to draw on. In contrast, productive thinking requires that one go beyond available knowledge, such that new procedures or knowledge be generated in order to achieve a solution. As Dominowski (1995) notes:

"Kohler argued that all problem solving concerns awareness of relations and that productive problem solving involves awareness of new relations among problem components. Understanding of these new relations, according to Kohler, is what is meant by insight" (p74).

Insightful productions are marked both by novelty, producing some idea or product that was not previously generated, and by functionality or value, with the new product fulfilling some purpose. Thus, insight processes are related to the wider domain of creativity and creative thinking, also providing a link between puzzle-based problems commonly studied in laboratory research and case studies of creative achievements with greater historical import. Insights may be novel and creative either historically, such as a new scientific discovery or invention that revolutionises how people interact the world, or personally, as when someone solves a puzzle they've never seen before; in the latter case, while a solution may be new

for an individual, many other people may have independently produced that same solution in the past (Robertson, 1999).

Classic Gestalt-based studies, while often lacking methodological rigour, often not satisfactorily replicated, and provoking vague explanations for insight, did introduce some intriguing concepts that have inspired modern perspectives on insight and creativity. These studies have paved the way for more rigorous research, extending our understanding of the processes involved.

The problems used to study insight in psychological research are generally defined by three criteria: they can be solved with little specialized knowledge, they commonly lead to an impasse in solution progress, and solution is attained suddenly with some new reorganization of knowledge accompanied by an 'Aha' experience (Dominowski, 1995; Schooler, Ohlsson, & Brooks, 1993). Typically, insight problems differ from ill-defined problems in that the former have specifiable problem states and goal states whereas the latter often do not. However, insight problems often do not have readily identified operations with which to reach a solution, in contrast to more well-defined problems. The key to solving many insight problems is in constructing an appropriate representation, rendering the solution obvious. The difficulty lies in the fact that presentation of the problem usually encourages an inappropriate representation, hence impeding solution.

The occurrence of insight has been set into a wider context of problem solving processes. Gestalt theories, particularly Wallas' (1926; cited in Robertson, 1999; Smith, 1995) classic model, propose four elements of insightful problem solving: preparation, impasse or incubation, insight or illumination, and verification. This model has enjoyed a modern resurgence in popularity, although specific aspects have been criticized. While most theorists generally accept that preparation and verification are elements of all problem solving, the concepts of impasse, incubation and illumination have been subject to some controversy. These latter concepts are assumed to be the defining processes in solution of insight-like problems.

### **a. Impasse in insight**

Solutions to insight problems are often characterized by a preceding impasse, or a period when the solver has no idea or direction of how to proceed. Typically, this is attributed either to the prior generation of an inappropriate representation of the problem or to an inability to generate potential strategies. The solver may realise that existing representations or strategies are not working, but be unable to produce any other useful ideas (Weisberg & Alba, 1981). Gestalt psychologists have demonstrated two factors that appear to promote impasses: mental set and functional fixedness. Stereotypy, or “mental set” fixation, refers to getting ‘stuck in a rut’, or the tendency to repeat previous strategies that have already proven to be unhelpful; however, one cannot escape the constraining influence of this set in order to try a more useful solution path (Davidson et al, 1994). Functional fixedness refers to the tendency to perceive and relate to an object only in terms of its usual function, even though using that same object for a different function can fulfil the requirements needed to solve a problem (Maier, 1931, cited in Ellen, 1982). In both cases, the inability to break away from inappropriate assumptions based on past experience can lead to impasses in solution attempts.

Two problem space conceptualisations, with differing implications for subsequent solution processes, have been proposed to explain impasses. First, impasses are viewed as searching through the wrong problem space or representation; solution thus requires generating a new, more appropriate representation, through some form of restructuring (Knoblich, Ohlsson, Haider, & Rhenius, 1999; Knoblich, Ohlsson, & Raney, 2001; Schooler & Melcher, 1995). Similarly, lateral thinking has also been construed as the ability to switch from one representation to another, rather than continuing to mine the depths of an unproductive approach: “Vertical thinking is digging the same hole deeper; lateral thinking is trying again elsewhere” (de Bono, 1967: p22). New representations may be generated, and impasses overcome, through several empirically supported processes including relaxation of constraints (inappropriate assumptions) and decomposition of perceptual chunks (Knoblich, et al., 1999), or selective encoding,



selective comparison, and selective combination of problem elements (Davidson, 1995; Davidson & Sternberg, 1984; Davidson et al, 1994).

Alternatively, impasses may result from employing the appropriate representation but not generating the correct strategy needed to navigate through the problem space in order to obtain the correct solution; for example, the search space may be sufficiently large that the correct path is difficult to find (Weisberg & Alba, 1981). However, the key to finding the correct path is to employ cued memory retrieval processes based on past attempts; practice and prior experience are helpful, insightful restructuring of existing knowledge is not necessary or helpful. This approach does not preclude a sudden solution; rather, even a sudden solution can occur without restructuring of the original representation. While the “insight” and “incremental” views may both be viable under different circumstances or problems, proponents of the incremental memory-search view tend to discredit the former, insight position.

#### **b. Incubation in insight**

Interestingly, researchers have demonstrated that people apparently overcome impasses and produce correct solutions following a period of incubation, or time taken away from mental work targeted on the problem (Mayer, 1995, 1999; Simon, 1966, cited in Robertson, 1999; Smith, 1995). This seems to contradict common sense; that is, *not* thinking about a problem seems to help a person solve it. The correct answer may appear as an insight either during this period of incubation or shortly after one resumes conscious solution attempts. Again, several explanations have been proposed, many focusing on unconscious mechanisms while others disavow any unconscious involvement. For example, Wallas (cited in Weisberg, 1993) implicated the unconscious recombination of old ideas to form new and more productive ideas; recent research provides some evidence for similar processes in terms of non-conscious spreading activation (Bowers, Farvolden, & Mermigis, 1995; Bowers, Regehr, Balthazard, & Parker, 1990; Ohlsson, 1992) and conscious selective combination (Davidson, 1995; Davidson et al, 1994). Breaks from a problem may allow for the substantial decay of an over-activated but inaccurate

representation (cf. mental sets) utilised prior to the break, such that returning to the problem allows one to overcome the fixation and develop a new representation that leads to solution (Simon, 1966, cited in Robertson, 1999; Smith, 1995).

Alternatively, terms or features in the problem presentation may implicitly cue non-conscious concepts in long-term memory related to the correct solution, priming a person to encode relevant information when it is experienced; related cues from the environment, even if attended to without awareness, may strengthen activation of the primed concepts in long-term memory to a degree that the appropriate concepts for solution suddenly appear in consciousness (Patalano & Siefert, 1994; Siefert, Meyer, Davidson, Patalano, & Yaniv, 1995; Yaniv & Meyer, 1987).

In contrast, Weisberg (1986, 1993) argues that many problems solved following an impasse are not accompanied by the sudden insight implied by the classic interpretation of incubation. He questions whether, given a solution that is generated without suddenness, incubation in the classical sense can be said to occur even if a break in progress is undertaken. Weisberg also suggests that in many cases of supposedly unconscious incubation people actually engage in sporadic, if brief, episodes of conscious “creative worrying” while concentrating on intervening activities. Subsequent progress towards a solution would likely be the result of these brief periods, even if the periods themselves were forgotten; if so, unconscious processes do not need to be implicated.

As with the experience of impasse, the precise processes occurring during periods of incubation may differ depending on the nature of the problems studied, the methods with which they are studied, and the context within which they are studied. Any or all of the above interpretations of incubation, or lack thereof, may be accurate under particular conditions; the task for researchers would then be to systematically determine under what conditions any particular set of processes are invoked. Until further research is conducted towards these ends, it would seem premature to dismiss out of hand any interpretation based on only selected readings of the literature.

Clearly, impasse, incubation and insight are disputed concepts. As with well-defined and ill-defined problems, insight-type problems come in many forms, and



can be distinguished in important respects (Weisberg, 1995b). Perhaps this is one reason why studies using insight-like problems have not yielded completely complementary results, and why our understanding of the processes involved in insightful solutions are incomplete. A greater appreciation of the processes and strategies, both cognitive and metacognitive, involved in the solution of insight and well-defined problems may refine our knowledge of the complexity of solution processes involved in these problems and in creativity more generally.

People obviously differ, individually and developmentally, in their abilities to solve problems (Brown, 1987; Jausovec, 1994; Kitchener, 1983; Short & Weissberg-Benchell, 1989); thus research is targeted towards delineating the factors that may help people to improve their problem solving performance. Metacognitive processing may be one set of factors that provides an avenue for understanding and developing such abilities.

## **METACOGNITION IN PROBLEM SOLVING**

The systematic study of metacognition is relatively recent (Bruning, Schraw & Ronning, 1999), although the philosophical roots of the concept date back much further (Yussen, 1985). Interest in metacognition within psychology harks back to the use of introspection by the early structuralist psychologists, in attempting to understand how a person's conscious awareness of his or her thinking affects those very thinking processes (Nelson & Narens, 1990). Contemporary interest arose as a reaction against the negative attitudes of the behaviourist and early information-processing schools towards consciousness (Tulving, 1994). Studying metacognition provides a first-person perspective of knowledge awareness, in contrast to the third-person perspective provided by earlier orientations. Flavell (1971, 1976) is credited with establishing metacognition as a research topic in its own right. He considered this to be "the central problem in learning and development" (1976: p231). Early literature indicates a primary concern with developmental aspects of self-reflective abilities in childhood (Metcalf & Shimamura, 1994; Yussen, 1985). Also, research focussed largely on memory, as opposed to other cognitive activities.

Flavell (1976) did briefly consider metacognitive aspects of problem solving. He suggested that children's problem solving is enhanced through the planful storage of information considered to be useful for future problem solving, the planful maintenance and revising of information for future retrieval, and the planful retrieval and systematic searching for relevant information when a problem requires solving. Flavell (1976) indicated that children must learn the 'how' (strategies), the 'where' (internal and external information sources), and the 'when' of problem-relevant information usage. He believed that people could become better problem solvers through learning how to improve their abilities to "assemble effective problem solving procedures from already available cognitive components" (p233).

It is apparent in the early literature that little theoretical construction or empirical research had been conducted to develop such ideas. Such theory and research has subsequently been developed, and important findings have appeared (Bruning et al, 1999). Indeed, the concept has proven of interest and worth in many research domains including memory (Bunnell, Baken, & Richards-Ward, 1999; Koriati, 1994, 1998; Leonesio & Nelson, 1990), problem solving (Berardi-Coletta, Dominowski, Buyer, & Rellinger, 1995; Betsinger, Cross, & DeFiore, 1994; Davidson et al, 1994; Davidson, 1995; Jausovec, 1994; Metcalfe, 1986a, 1986b; Metcalfe & Wiebe, 1987), perceptual processes (Bowers et al 1990; Carroll, 1993), language comprehension and production (Brown, Armbruster, & Baker, 1986; Greeno & Riley, 1987; Hacker, 1998; Pereira-Laird, 1996), social cognition (Gollwitzer & Schaal, 1998; Lories, Dardenne & Yzerbyt, 1998; Mischel, 1998), development (Butterfield, Nelson & Peck, 1988; Hertzog & Dixon, 1994; Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Schneider, 1998), neuropsychology (Shimamura, 1994; Shimamura & Squire, 1986), and motor activity (Simon & Bjork, 2001). Practical fields including education (Mayer, 1998), clinical practice (Dixon, Heppner, Burnett, Anderson, & Wood, 1993; Flett & Johnston, 1992; Mayo & Tanaka-Matsumi, 1996), and business/organizational practice (Smith, 1998; Williams & Yang, 1999) have also incorporated metacognitive perspectives.

At the most general level, Nelson and Narens (1990) provide a broad model for metacognition. They posit the existence of two levels of cognition: a lower

‘object-level’ at which cognitive activity takes place, and a higher ‘meta-level’ which contains a dynamic model of, and controls the activity of, the object-level. Two reciprocal types of information flow represent the relationship between these levels: ‘control’ from meta-level, which regulates and modifies the activity of the object-level; and ‘monitoring’ from the object-level, which informs the higher level of its activity, and modifies the meta-level model of the lower-level. Primarily, Nelson and Narens (1990) have applied this framework to memory processes, from acquisition to retrieval. However, this model may also be applicable to problem solving, albeit only as a descriptive tool.

Such definitions of ‘metacognition’ are criticized for their vagueness (Brown 1987; Paris and Winograd, 1990; Jausovec, 1994, 1999). For example, Brown (1987) argues that while the blanket term ‘metacognition’ encompasses an essential concept, it is rather nebulous and glosses over important distinctions. That is, metacognition is not one underlying process, but rather a set of processes that may differ across task and problem domain. Also, it is often difficult to distinguish ‘cognitive’ from ‘metacognitive’ processes (Weinert, 1987). The nature of metacognition also provides measurement difficulties (Paris & Winograd, 1990). Brown (1987) therefore advocates that, in the interests of “clarity and communicative efficiency” (p106), researchers should focus on the specific processes encompassed by the term, and the specific cognitive domains in which it is used (memory, communication, etc.). The term is still of value, however, as an orientation towards thinking of cognitive awareness and development, performance differences, and instruction (Paris & Winograd, 1990; Yussen, 1985). It is rendered more useful if efforts are made to delineate the specific processes under consideration, and to study these in detail as distinguishable but related processes.

In delineating more specific processes of metacognition, most theorists have distinguished between two major aspects: metacognitive knowledge or beliefs, and metacognitive strategies or executive processes (Brown, 1978; Brown et al, 1986; Flavell, 1987; Kluwe, 1982, 1987). Metacognitive experiences or feelings have also been identified as important (Flavell, 1987; Metcalfe, 1986a, 1986b; Metcalfe & Wiebe, 1987; Davidson, 1995). In addition, more recent theories have included

motivational factors, such as interest in task-engagement, desire to succeed, self-confidence, and performance attributions (see Ashman & Conway, 1997; Mayer, 1998; Short & Weissberg-Benchell, 1989), and epistemological assumptions (Kitchener, 1983; Schraw, Dunkle, & Bendixen, 1995) in a more comprehensive account of metacognitive activity. The following discussion outlines theoretical contributions to understandings of metacognitive knowledge, executive processes, and affective experiences.

### **1. Metacognitive Knowledge And Executive Control**

Most common in early models is the distinction between declarative and procedural components of metacognition, or between knowledge of cognition and regulation of cognition. For example, Brown (1978) distinguishes between “knowing what”, or knowledge of necessary process or strategy, and “knowing how and when” to use an applicable process or strategy. Kluwe (1982) states that the central aspects of metacognition are that a person has knowledge of one’s own and others’ thinking, and that a person has the ability to control or regulate his or her own thinking. Metacognitive self-appraisal has similarly been conceived as declarative (what you know), procedural (how you think), and conditional (knowing when and why certain knowledge and strategies should be used) (Paris & Winograd, 1990). The declarative-procedural distinction reflects a common differentiation throughout cognitive theory, most notably in theories of memory (Matlin, 1998), but its relevance to problem solving is apparent. Furthermore, acknowledging the metacognitive components of thinking emphasizes the active and self-directive features of cognition.

#### **a. Metacognitive Knowledge**

Metacognitive knowledge has been defined as knowledge of cognition (Brown, 1978), “one’s knowledge concerning one’s own cognitive processes and products, or anything related to them” (Flavell, 1976: p232), and as “the acquisition of knowledge, the amount of knowledge and the assumptions and opinions about the states and activities of the human mind” (Kluwe, 1987: p31). It is clearly a form of

declarative knowledge in the form of self-reflective thinking focused on the nature and on-going activity of cognitive processes. Flavell (1976, 1978, 1987), for example, distinguishes among three central forms of meta-level knowledge of cognitive phenomena: person-based knowledge, task-based knowledge, and strategy-based knowledge. Person-based knowledge includes understanding one's own intra-individual differences in ability across content domains, tasks, and time, understanding inter-individual differences in abilities between people within specific domains and tasks, as well as knowledge of universal factors in thinking common to all people, such as the fallibility of short-term memory or that more difficult tasks require greater effort. Task-based knowledge involves an understanding of how different activities or situations demand different types of strategies, processing, and effort. Strategy-based knowledge involves one's understandings not only of particular cognitive strategies that are applicable across different situations, but also of metacognitive strategies that monitor and control the use of lower-level cognitive strategies. Together with these three forms of knowledge, Flavell (1978) notes that sensitivity to knowing when particular forms of knowledge are necessary is an additional facet of metacognitive knowledge.

Kluwe's (1982) model of declarative knowledge explicates the nature of metacognitive knowledge in greater detail. According to Kluwe, at least six forms of metacognitive knowledge are distinguishable across the three dimensions of *domain specificity* versus *generality*, *cognitive activity* versus *transformation of activity*, and *generality* versus *diagnosticity*. He contrasts one's cognitive-level domain knowledge of specific content areas with metacognitive beliefs and assumptions that may be both *domain-specific*, such as believing that one is good at arithmetic but not so good at creative writing, and *domain-invariant* or constant across context. Domain-specific and domain-invariant forms of metacognitive knowledge incorporate understandings of cognitive *states*, *processes*, and *activities* as well as the *means to transform* those cognitive states and activities. These forms may be further divided into *general knowledge* about the organization of cognitive systems and *diagnostic knowledge* that guides beliefs of own and others' thinking in specific situations. For Kluwe, *general knowledge* represents a wide-based belief system

about the nature of thinking processes, while *diagnostic knowledge* is organized in the form of self-schemas that integrate beliefs about one's specific abilities.

### **b. Executive Control: Monitoring & Regulation**

Procedural aspects of metacognition have been recognised in terms of regulation or executive control, referring to the directed monitoring and guidance of ongoing cognitive activity. Kluwe (1982) discusses both cognitive and metacognitive aspects of procedural knowledge. At a cognitive level are *solution processes*, the strategies, processes and operations aimed at providing solutions to problems. At a metacognitive level are *executive processes* that monitor ongoing cognitive activity and regulate the selection, application, and effects of available cognitive strategies. The distinction between monitoring and regulation seems particularly important (Nelson & Narens, 1990). Whereas Nelson and Narens (1990) conceive monitoring as distinct from control, Kluwe's (1982) model subsumes monitoring and regulation together under the rubric of executive control. Monitoring of cognition allows for the gathering of knowledge about immediate thought processes, while regulation allows for efficient application of those processes towards perceived task demands in order to complete some task. Both monitoring and regulation are considered processes that provide executive control of thinking.

Brown (1978) states that essential executive skills in the self-regulation of problem solving include *prediction* of one's own capacity to solve a problem, *awareness of appropriate heuristic strategies* and how these should be applied, *identification* of the problem at hand, *planning* of potential strategies into a usable form, *monitoring* of the strategies as they are used, and ongoing *evaluation* of both the processes and products of problem solving to determine a suitable endpoint of one's efforts. Similarly, Kluwe (1982) distinguishes the *monitoring* activities of identification, prediction, checking, and evaluation, from the *executive regulation* of self-motivation and interest, one's resources and their allocation, the intensity of effort in the form of duration and persistence, and speed of processing.

Both metacognitive knowledge and executive control are assumed to be related, though distinct, forms of metacognition. Kluwe (1987) suggests that



declarative knowledge and executive control processes operate together when a person is confronted with a problem solving scenario. For example, one's declarative metacognitive knowledge allows one to recognise a problem situation and to encode relevant information about the problem's elements, to provide informed executive decisions about appropriate strategies and plans that may produce a solution. While it is the knowledge facets that provide problem-relevant information for the solver, it is the executive control and regulation functions that allow solution processes to proceed. However, knowledge and executive processes are logically and empirically distinct. Metacognitive knowledge appears to be reasonably stable, consciously statable, and late-developing, while executive processes may be more automatic, not consciously statable, context-dependent across specific tasks, and not age-dependent (Brown, 1978; Bruning et al, 1999; Pereira-Laird, 1996).

## **2. Metacognitive Experiences**

Metacognitive experiences have been identified as affective counterparts of metacognitive self-appraisal but have received less research attention than the knowledge-based or procedural control components (Flavell, 1987; Gick & Lockhart, 1995; Metcalfe, 1986a, 1986b; Metcalfe & Wiebe, 1987; Yussen, 1985). Such experiences, or feelings, are defined as "relatively spontaneous reactions or reflections that occur on line (during the cognitive process) while the cognitive enterprise is rolling along" (Yussen, 1985: p256). Whereas metacognitive knowledge refers to memory-based conceptions of one's knowledge, and metacognitive control is how people use their knowledge and strategy repertoires, metacognitive experiences represent immediate affective and cognitive responses to ongoing activity; for example, miscomprehending the nature of a problem, realizing that one is frustrated with progress on a problem, or having a sense of surprise at suddenly finding a workable solution.

These on-line feelings can be diagnostic, if interpreted correctly, in that they can direct the problem solver to aspects of their cognitive activity that require greater or lesser attention. There appears to be developmental differences in ability

to interpret such experiences, with younger children being less able than older children or adults to respond appropriately to their reflective feelings (Flavell, 1987). Metacognitive experiences may also be similar to ongoing attributions about the causes of ease and difficulty in a problem solving episode (Borkowski, Carr, Rellinger, & Pressley, 1990). Gick and Lockhart (1995) suggest that initial affective responses to a problem can motivate a person's decision to ignore or engage in problem solving.

The self-reflective and diagnostic nature of metacognitive feelings may be particularly useful in the continuous monitoring of cognitive activities, particularly problem solving attempts. For example, feelings of warmth or progress towards a goal should guide the direction of a person's subsequent strategies. Feeling that one is working in the right direction will allow narrowing of potential solution paths down to those deemed most productive; feeling that one is not working in the right direction encourages the solver to try a new solution path (Metcalf, 1986b; Simon, Newell, & Shaw, 1979). This obviously requires a measure of self-reflection involving explicit, or possibly implicit, appraisal of problem-relevant information. However, the affective experiences associated with problem solving may be negative (frustration at lack of progress, annoyance at not solving the problem earlier) as well as positive (pleasure at finding correct solution) (Gick & Lockhart, 1995). The affective quality of metacognitive appraisal may be most apparent in the solution of insight problems; solution to these problems is often accompanied by a sense of suddenness or surprise, resulting in the reputed 'Aha!' reaction (Metcalf, 1986b; Metcalf & Wiebe, 1987; Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995). The resolution and affective response to insight problems is similar to that experienced in 'getting' a joke (Gick & Lockhart, 1995).

Distinctions between elements of metacognition, particularly metacognitive knowledge and executive control, have been central to the confusion surrounding the construct of metacognition, and have led to doubts about the extent to which declarative knowledge and procedural control can be related. Some researchers (e.g. Kluwe, 1982, 1987; Nelson & Narens, 1990; Pereira-Laird, 1996) obviously see the two forms as interactive components, while others argue that either one or the other



form should alone be considered as metacognitive. The prevailing belief is that a full appreciation of metacognition and related behaviour requires consideration of knowledge, executive processes, and affective experiences together.

### **3. Metacognitive Information Processing In Problem Solving**

Information-processing approaches to metacognition have applied metacognitive knowledge and control to various higher-order processes across the course of a problem solving episode. Typical progressive metacognitive processes include identification, representation, planning, monitoring, and evaluation (Brown, 1978; Davidson et al, 1994; Flavell, 1978; Kluwe, 1982, 1987); presumably these processes are universally applicable across many domains of problem solving.

#### **a. Identification and problem finding**

Identification of a problem is a critical first step; recognizing that a problem exists, and having a desire to rectify the problem, encourages one to engage in problem solving activities. All problem solving reputedly requires the solver to identify, or encode, the relevant features of the problem, to store this information in working memory and long-term memory, and to relate the incoming information to existing relevant knowledge structures (Flavell, 1978; Newell & Simon, 1972). Identification of a problem requires a certain amount of self-reflection on the features of a situation to determine if a problem actually exists; that is, if there are obstacles to be overcome in achieving a goal. Many potential problematic or improvable situations may go unnoticed if a person cannot identify elements in a situation that can be changed. "Problem finding" has recently been identified as an important skill in post-formal adult thinking, and has been related to creative processes (Dominowski, 1995; Lubart & Sternberg, 1995; Perkins, 1981). People who can view existing situations in novel and creative ways can presumably focus on otherwise unnoticed but improvable conditions, or find better methods of organizing situations to facilitate some new goal (Arlin, 1989; de Bono, 1967).

Problem identification may be just as conceptually complex as subsequent metacognitive phases of problem solving. Sufficient identification of a problem and

its features allows for mental representation of the problem, prediction of impending success, and planning of solution strategies.

### **b. Problem representation and solution prediction**

Developing a useful mental representation of a problem's structure is essential to engaging in effective solution processes (Newell and Simon, 1972). Many problems may elicit representations automatically, without conscious control (Schooler & Melcher, 1995). However, metacognitive control over representation construction is possible, and is particularly useful where solution to a problem requires a change in representation (e.g. insight problems).

For example, Davidson and Sternberg's (1984; Davidson, 1995; Davidson et al, 1994; Sternberg & Davidson, 1982) three-process model of selective processing, a sub-theory of the triarchic theory of intelligence, outlines three metacognitive processes that influence the development of problem representations: selective encoding, selective combination, and selective comparison. These processes are arguably applicable to all problems, though Davidson and Sternberg emphasize the relevance to insight problems. Selective encoding involves focusing on that information which is deemed most relevant to a correct solution; if a solution is not possible, representational change may require selective encoding of problem features that were originally non-obvious. Selective combination involves integration of problem relevant information into patterns that facilitate solution; impasses in progress may be overcome by combining features in otherwise non-obvious ways. Selective comparison requires the solver to compare new problem-relevant information with existing knowledge, through analogies and metaphors for example, to develop a workable solution; again, non-obvious connections between new and old knowledge can facilitate changes in representation that facilitate solution. Davidson and Sternberg (1984) note that solving a problem may require any one, or a combination, of these processes. Research needs to consider under what conditions and with what problems each of these processes are valuable.

An understanding of the nature of a problem, acquired once a representation has been developed and one's relevant knowledge and competence has been

assessed, allows for predictions of imminent solution progress and anticipation of the likelihood of success. Such predictive judgements, or feelings-of-knowing (FOKs) the answer to some problem, are crucial to the forthcoming course of solving attempts: selecting problems that are considered solvable, indicating how much time, effort, and persistence should be allocated, and selecting appropriate strategies that could lead to solution (Kluwe, 1982; Metcalfe, 1998a; Paris & Winograd, 1990). These predictive functions in turn allow for planning of problem attempts.

Interestingly, research across problem solving and other cognitive activities (e.g. memory) has demonstrated a pervasive 'cognitive optimism' in people's predictive judgements; people generally believe that their performance will be better than it actually is (Metcalfe, 1998a). The relation of prediction to performance depends on how it is assessed. Schwartz and Metcalfe (1994) distinguish between micro-predictive and macro-predictive accuracy. With micro-prediction measurement, in absolute terms people tend to perform better on specific tasks that they are more confident about solving than on tasks they are less confident about; in this sense, people are generally accurate at predictive ranking of tasks in terms of relative difficulty. In contrast, macro-prediction refers to comparing the average predictions with respect to overall performance; on average, people overestimate the probability of imminent success. Over-prediction appears to be due to the nature of the information on which people base their estimates; namely, any relevant knowledge that is activated or accessible from memory, regardless of its accuracy (Koriat, 1994, 1998; Metcalfe, 1998a). That is, the more partial, even if inaccurate, information people can access upon cuing of the problem and its representation the higher their predictions tend to be. Unfortunately, the incomplete or inaccurate information upon which estimates are based does not actually help problem solution; hence failure is often the outcome.

The implication of overestimation is that it does not seem to support efficient problem solving; overestimations of success may lead to less efficient monitoring, prompting people to terminate solution attempts before the correct solution has actually been found.

### **c. Monitoring and evaluation**

Monitoring and evaluation are closely related higher-order activities, and may be difficult to distinguish. Monitoring is obviously a central aspect of metacognitive control in most cognitive activities, but has a particularly relevant role in the progress of solution activities as they occur, as discussed above. Monitoring itself represents one form of evaluation process, that of the ongoing solution process. Efficient on-line monitoring and regulation of solution processes may enable greater performance, through the generation of more accurate or useful solution products (Brown, 1978; Kluwe, 1982, 1987).

However, evaluation of the products themselves is also important. Once a potential or partial solution has been generated, the problem solver needs to evaluate the solution to determine if it indeed meets the requirements of the identified goal; if so, solution efforts may be terminated but, if not, the search for a new solution begins or is terminated because the solver does not wish to persevere with the problem. This latter case indicates why monitoring and evaluation are inseparable, because evaluation is ongoing throughout the solution episode until problem-related activity is terminated. Davies (2000) demonstrated the effectiveness of ongoing evaluation in solving a well-structured problem. Performance on the Tower of Hanoi task was enhanced for participants who were required to provide a verbalized or non-verbalized evaluation for each successive move, relative to participants who provided no evaluations. Additionally, participants providing evaluations were disrupted by undertaking a concurrent task while no-evaluation participants experienced no disruption from this task. This suggested that the act of ongoing progress evaluation enabled participants to develop explicit representations of their solution strategies; these representations were open to disruption by increased working-memory load.

Indeed, all of the metacognitive activities identified above may occur in a non-linear form as the problem solver reflects on their activity; all processes presumably occur in an interactive fashion together, and all are necessary if a goal is to be obtained.

Despite distinctions between knowledge, control, and affective experiences, metacognitive activity is assumed to be central to efficient thinking and performance across the course of problem solving. Indeed, Brown (1978) considers executive functioning to be “the crux of efficient problem solving” (p82). Kluwe (1987) suggests that it may be both intra- and inter-individual variations in executive control of thinking that account, to a reasonable degree, for performance differences and deficits. Research has helped establish the veracity of the hypothesized link between metacognition and cognitive performance, but not without controversy. Metacognitive processing has proven to be a challenging construct to investigate empirically, due to the subjective and higher-order nature of the processes suggested by theory.

Informative empirical findings have accrued through the use of ‘think-aloud’ verbalization techniques, and subjectively-based phenomenological techniques that reputedly tap into metacognitive experiences.

#### **4. Metacognitive Monitoring And Verbalization**

The impetus for verbalization procedures arises from the identified need to access a person’s flow of conscious thoughts as they engage in a problem, based on an assumption that a person’s immediate thoughts contain higher-level self-reflective ‘inner speech’ that can be characterized as metacognitive. Presumably, if the researcher can gain access to these higher-level thoughts then he or she can observe what metacognitive processes the problem solver is engaged in during the immediate moment of solution activity; this may also allow one to observe in what ways metacognitive thoughts may regulate concurrent cognitive activity and performance (Dominowski, 1998). Verbalization, or ‘think-aloud’, procedures require participants to speak whatever thoughts come to mind, presumably in working memory, as they work on a task. Verbalization methods may be retrospective or concurrent; concurrent methods may be either directed or non-directed (Ericsson & Simon, 1993). Retrospective methods require participants to describe their prior thoughts shortly after engaging in an activity; concurrent methods require on-line reporting of thoughts while engaged in a task.

Doubts have been cast over the accuracy of verbal procedures to provide a window into people's metacognitive reasoning (Jausovec, 1994; 1999). Nisbett and Wilson (1977) argue that relatively little of our cognitive processes are available to awareness; that these self-reports are subjective and unverifiable, and thus unreliable; and concurrent verbalization may in fact interfere with the processes deemed to be accessed. Jausovec (1994) adds that people can report their cognitions only sequentially, whereas many processes operate in parallel and at a rate too fast to report; also, the use of different coding protocols across studies encourages inconsistent interpretations of verbal data. Research demonstrates that verbalization can be an inaccurate record of cognitive and can adversely effect processing in some cases (Nisbett & Wilson, 1977; Schooler et al, 1993); however this depends on the type of verbalization instructions employed and the exact nature of what processes are under investigation (Brown, 1987; Ericsson & Simon, 1980, 1993). Admittedly, if people are required to provide on-line verbalizations about otherwise non-reportable information, the act of verbalization may hinder actual processing (Schooler & Engstler-Schooler, 1990; Wilson & Schooler, 1991; Brown, 1987).

Nevertheless, think-aloud procedures are one of the more common methods in metacognitive research. Furthermore, their use is supported both theoretically and empirically. For example, Ericsson and Simon (1980, 1993) contend that concurrent verbalization of working memory contents has a neutral, non-disruptive effect on cognitive processes. Where a person reports information that is not readily accessible in verbal form (e.g. some visual information), processing may be neutral but slowed-down as recoding into verbal form takes place. In either case, several commentators (e.g. Lieberman, 1979; Nelson & Narens, 1990) argue that introspective self-reports can be informative if considered as an imperfect means of self-awareness, and interpreted in this light. Verbalizations do not have to provide complete access to underlying processes to be informative and useful. It is the motivating and influential nature of these introspective reports with respect to cognitive performance that necessitates the need to study such processes. Conscious but incomplete thoughts and verbalizations may be particularly informative of metacognitive thinking.



Berry (1983) and Berardi-Coletta et al (1995) have demonstrated that verbalization provides access to metacognitive processes, and that this metacognition improves problem solving performance. Berardi-Coletta et al. (1995) found similar positive transfer results for process-oriented verbalization groups relative to other groups engaged in both the 'Tower of Hanoi' task and the Katona card problem. Two process-level groups (either focusing on process-level metacognitive (MC) monitoring of solutions or making "If...Then" (IT) statements) performed significantly better than other groups (problem-focused (PF), think-aloud (TA) control, silent control) on both practice and transfer trials, in terms of both the ratio of excess to minimum required moves, and time to solution. The researchers demonstrated that beneficial effects were due not to verbalization per se, but to metacognitive processes evoked by the requirement to explain one's thoughts. As expected, process-oriented statements were more common (60% of total statements) for both process groups, less common (5%) for the TA group, and absent for the PF group. These results demonstrated that the shift in processing to a more process-oriented level did indeed induce participants to engage in more metacognitive reasoning, and this improved their performance. Metacognitive statements were not made spontaneously in either the TA or PF groups. In a subsequent experiment, metacognitive (MC) group members, instructed to think about answers to process-level questions rather than to verbalise, performed better than a control group receiving no additional instructions beyond performing the task. This suggested that it is not overt or covert verbalizing per se, but the metacognitive processing induced in participants, that aided performance, a finding replicated by Davies (2000). The type of thinking encouraged by overt or covert thinking is crucial.

These studies demonstrate the usefulness of both verbalization and process-oriented thinking in problem-solving tasks, although the generality of such effects is not yet established. For example, the effects of verbalization may depend on the nature of the task studied (Dominowski, 1998). The studies cited above used well-defined problems; verbalization may not be so effective with non-incremental problems that require other than well-defined solution methods. Schooler et al (1993) have demonstrated that insight problem solving may be subject to the same

verbal overshadowing that has been shown with non-reportable processes such as facial recognition (Schooler & Engstler-Schooler, 1990) and aesthetic judgements of taste (Wilson & Schooler, 1991). Using both retrospective and concurrent verbalization methods Schooler et al. (1993) found that participants who were required to verbalise their thoughts while working on both visual and verbal insight problems solved fewer problems than participants not required to verbalize. This negative effect of verbalization did not occur for noninsight problems. Schooler et al. suggested that solution of the insight problems required processes that were not available for conscious inspection; the need to verbalize one's thoughts increases the salience of the verbalizable aspects of the stimulus, thus overshadowing the non-verbalizable aspects. For many insight problems, it may be the non-reportable aspects that allow solution of the problem; if these aspects are overshadowed by verbalization, solution is impeded. Given that language processes appear to hinder at least some insight problem solving, insight processes may operate independently of language and are distinguishable in this sense from more well-defined problems.

Such findings prompt interesting questions about how, or indeed whether, insight problem solving can be studied at a metacognitive level. Not only do insight processes appear inaccessible through verbalization, but also the act of focusing on verbalizable aspects may actually hinder performance. However, insight-related metacognitive processes may be meaningfully accessible through the investigation of another form of higher-order thinking; namely, metacognitive experiences and feelings.

## **METACOGNITIVE EXPERIENCE IN INSIGHT**

Given that verbalization techniques often have a negative effect on solution of insight-type problems, it is necessary to approach insight-based metacognitions from another perspective. Metacognitive experiences or feelings have been identified as emotional counterparts of metacognitive self-appraisals. Insight is an apt area in which to study metacognitive experience given that insight is often accompanied by strong affective responses (Gick & Lockhart, 1995). Research has found that using metacognitive experience measures is useful for studying insight-



related processes, and tends to support distinctions between insight and other problem solving processes. However, findings have raised issues about whether 'insight' can be defined by metacognitive experiences during solution, whether solutions to insight problems occur with discontinuous or incremental cognitive processes regardless of metacognitive feelings, and whether unconscious or conscious processes are responsible.

### **1. Discontinuous Versus Continuous Metacognition**

Metcalf's (1986a, 1986b; Metcalfe & Wiebe, 1987) research has provided some intriguing results linking metacognitive phenomenology to insight problem solving. Adapting a methodology more commonly employed in metamemory research (e.g. Nelson & Narens, 1980), Metcalfe has compared the metacognitive experiences of solving insight problems with those used in completing both memory-related tasks and well-defined problems. Specifically, Metcalfe's research has measured both 'feelings-of-knowing' (FOK) predictions and 'feelings-of-warmth' (FOW) or closeness to a solution as one monitors his or her progress on a task or problem.

#### **a. Feelings-of-knowing**

In Metcalfe's paradigm, participants are presented with either problems or trivia questions typed individually on flash cards. Items are shown sequentially, with the participant allowed about five seconds to produce a correct response; if no correct response is given, that item is put aside and the next item presented until a specified number of items (often five) remain unsolved. These unsolved items are then shuffled and re-presented together, and FOK predictions are made in either one or two ways: first, by ranking the items in terms of those the participant feels they are most likely and least likely to complete; second, by shuffling again and having the participant provide absolute probability judgments, rating (often on a 10-point scale) the likelihood of completing each item. The participant is then given a forced-choice recognition test for memory items, or has the opportunity to solve each problem, with a specified time limit for completing each item.

Comparing FOK predictions for memory trivia questions and insight-related puzzle problems, Metcalfe (1986a) found that macro-predictive accuracy was high for the memory task (gamma [G] correlation = .52 for ranked predictions, .54 for probability estimates), but non-existent for the insight problems ( $G = -.15$  ranked,  $-.18$  probabilities). While participants over-predicted their accuracy on both tasks, the over-prediction effect was much higher for insight problems than for memory performance. Metcalfe assumed that participants' predictions were reasonably accurate for the memory task because they could access incomplete but relevant partial information in memory on which to base their predictions. For insight problems, however, poor predictive accuracy resulted from problems encouraging an inaccurate sense of confidence, supported by the fact that correct solutions occurred suddenly, without warning; presumably without gradual accrual of relevant information prior to solution.

Metcalfe and Wiebe (1987) found a similar effect of FOK accuracy when comparing insight with non-insight problems. Non-insight or well-defined problems are presumably solved in an incremental, step-by-step fashion by the gradual accrual of relevant information and progress from problem state to goal state. Conducive with these expectations, non-insight problems, as with memory trivia questions, produced more accurate FOK predictions than the insight problems. Together, Metcalfe's results correspond with the 'cognitive optimism' effect manifested across cognitive tasks (Metcalfe, 1998a); namely, that metacognitive judgments tend to overestimate performance, but that overestimation is usually greater for more difficult tasks.

While Metcalfe's studies have demonstrated clear disparities between insight problem solving and other cognitive activities, it is unclear how accurately her procedure approximates real-world prediction behaviour. For example, people may seldom be aware of a whole set of problems in advance, such that they have an opportunity to assess the likelihood of solving each problem by ranking all known problems in terms of perceived difficulty. It seems more likely that problems are dealt with sequentially, and that one might not predict his or her chances of success until one has identified a given problem and is about to solve it. Issues related to the

measurement of problem prediction deserve more attention in metacognitive research.

#### b. Feelings-of-warmth

Metcalf (1986b) and Metcalf and Wiebe (1987) also examined the dynamic experience of metacognitive monitoring, by having participants provide feeling-of-warmth (FOW) ratings during solution episodes, indicating how close they believed they were to a solution. In this procedure, the participant is typically given a sheet of paper marked with a series of horizontal lines running sequentially down the page. Vertical lines are conceived as fixed-point (e.g. 7- or 10-point) scales. As the participant works on a problem, he or she uses a pencil to mark the scales with a vertical slash 'warmth' rating; slashes made to the left, 'cold', end of a scale indicate that he or she feels less close to a solution and slashes made further to the right, 'warm', end of a scale indicate he or she feels progressively closer to a solution, with a slash made to the extreme right of a scale when the participant is sure he or she has a solution. Only one rating mark is made on each horizontal scale, with the participant working down the page as he or she continues work on a problem. Prior to, or at the commencement of, a solution attempt, the participant marks the first line on the page with a slash at the extreme left of the scale, with subsequent ratings made every 10 or 15 seconds during the attempt until either the participant generates a solution or a time-limit is reached.

Metcalf (1986b) found that correct solutions to insight problems were typically, but not universally, preceded by a series of low warmth ratings until immediately prior to solution when ratings increased suddenly to high rating points. Incorrect solutions were instead accompanied by gradual increments in warmth, until a wrong solution was given. Similarly, Metcalf and Wiebe (1987) demonstrated that correctly solved non-insight problems were accompanied by more gradual increases in warmth, while correctly solved insight problems were again solved more suddenly. These researchers assumed that, in solving insight problems, people do not initially have available the correct representation or strategies required to produce the correct solution, with low warmth ratings reflecting a feeling of being

unable to generate any workable strategy; once a productive strategy or representation is generated, correct solution follows rapidly. However, gradual increases in warmth during insight problem solving typically indicate that an inaccurate representation or strategy has been developed and implemented. Metcalfe and Wiebe (1987) argued that metacognitive phenomenology could be used as a means of defining insight and incremental solution processes for a given individual.

Other researchers have supported Metcalfe's results. Jausovec (1994) assessed metacognitive factors in problem solving behaviours of high- and average-ability solvers. Good problem solvers, given their assumed greater metacognitive abilities, were predicted to differ in their FOW ratings when solving well-defined, ill-defined, and insight problems. No difference across problem type was expected for poor problem solvers lacking the requisite awareness to monitor solutions differently according to task demands. Results supported these hypotheses. In comparing the FOW ratings with think-aloud protocols, more able solvers demonstrated greater congruency between their ratings and the solution strategies reflected in their verbalisations, with poor solvers showing no congruency. Also, the able solvers provided more statements indicating metacognitive processing, with poor subjects providing no such statements. Altogether, in comparison to poor problem solvers, able solvers were better estimators of their abilities, had greater knowledge of cognitive processes, and engaged in more metacognitive activities.

Davidson (1995) extended Metcalfe's results to examine metacognitive experiences across problems pre-designated as involving processes identified from Davidson and Sternberg's (1984) three-process theory. Insight problems whose solutions were considered to require selective encoding, selective combination, or selective comparison were compared to non-insight problems. Forty-eight participants were divided for comparison into high- and average-ability groups on the basis of IQ-test scores. Performance on all problem types was greater for high-ability than for low-ability participants. When warmth ratings were analysed, the selective-encoding and selective comparison problems yielded the predicted insight effects in terms of lower ratings for correct than for incorrect solutions. Non-insight and selective-combination problems produced no difference in warmth for correct

and incorrect solutions. Further, lower ratings were provided for insight problems by high-ability than average-ability participants, presumably because average-ability participants gave more incorrect solutions (with accompanying high ratings). When warmth patterns were assessed, correctly solved selective-encoding and selective-comparison problems tended to yield typical insight patterns, while correctly solved noninsight and incorrectly solved selective-encoding and selective-comparison problems produced more incremental warmth patterns. Neither pattern tended to be produced for correctly solved selective-combination problems.

Apart from supporting Metcalfe's distinction between insight and well-defined problems, Davidson's (1995) results provided other interesting findings. First, highly intelligent solvers were more likely than less intelligent solvers to solve insight problems, and to solve these with subjectively sudden patterns of warmth, possibly because they engaged in more accurate information processing and monitoring of their progress (i.e. more likely to continue searching for an answer until the critical insight occurred; knowing they were unsure of a solution until it was produced suddenly). Second, results demonstrated that processes assumed to be necessary for representational change can promote insightful solutions, but are also distinguishable; selective-encoding and selective-comparison appear to facilitate sudden insight, while selective-combination may involve both insightful and incremental processes. Unfortunately, these problems were only pre-classified by independent raters; no verification was undertaken to assess whether participants actually used the hypothesized processes for each of the respective problem types.

The implications of the above studies have been challenged. For example, defining insight only in terms of accompanying phenomenology is circular without independent verification of the actual processes involved (Dominowski & Dallob, 1995; Weisberg, 1992). Weisberg and Alba (1981) examined performance on classic problems, such as the nine-dot problem, presumably solved with an insight experience. Presumably these problems are initially difficult because past experiences with similar problems leads people to become fixated on unwarranted assumptions that block correct solution, until assumptions are revised and rapid insight occurs. However, performance on the nine-dot problem was poor, even when

participants were given hints that negated the alleged fixation; very detailed instructions had to be given before any notable facilitation was observed. Attainment of the required insight did not appear to promote rapid solutions of the nine-dot problem. Weisberg and Alba (1981, 1982) interpreted these results to mean that solutions to 'insight' problems are not impeded by the inappropriate application of past experience that is overcome only through a sudden process of reorganization; rather, people rely on problem-cued memory retrieval wherein previous specific knowledge is used as the basis for subsequent solution attempts and revision of current understandings. However, subsequent findings that amnesiacs can produce near-normal solution performance on insight-like tasks suggest that implicit rather than explicit memory processes may be more crucial (Metcalf, 1998b). Also, Lung and Dominowski (1985) found that the nine-dot problem could be solved without specific instructions, depending on the quality of general instructions and practice. But again, the critical 'insight' for solution to this and similar problems does not seem easy to generate spontaneously, nor does it lead to immediate solution. However, these results do not negate the existence of insight. These problems may be impure examples of 'insight' problems, requiring both insight and incremental processes; the crucial insight may be necessary but not sufficient for correct solution; once the insight is achieved, further verification or development may be necessary before solution is achieved (Schooler & Melcher, 1995; Weisberg, 1995b). It seems imprudent to conclude from only a few 'insight' problems that insight as a phenomenon does not occur.

## **2. Conscious Versus Unconscious Processing**

Other researchers argue that, at a cognitive level, 'insight'-like tasks are actually solved incrementally rather than discontinuously, even when insight is subjectively experienced as sudden and discontinuous. Much gradual processing may occur unconsciously, prior to realisation of a solution. Using both visual and verbal tasks, Bowers et al (1990) and Bowers et al (1995) demonstrated that people could intuitively recognise a coherent solution pattern even when they could not provide a definite solution and lacked confidence in their choices. Also, people may



gradually approach a correct solution to a problem while lacking confidence in their closeness. Even when one achieves the correct solution it may be some time before they realise the significance of their discovery, in effect experiencing as discontinuous an otherwise continuous, if partly unconscious, process.

Similarly, Durso, Rea, and Dayton (1994) gathered participants' ratings of the relatedness of words and concepts that were either related or unrelated to the crucial concepts for solution to an insight problem. Graphs devised from these ratings indicated that for correct solvers, conceptual relations were cognitively restructured immediately upon solution. However, some intermediate restructuring was apparent prior to solution, indicating that gradual changes did occur before the perceived sudden realisation. The conscious experience of suddenness does not imply that cognitive processing of a problem occurs discontinuously rather than incrementally; hence, phenomenological ratings may not accurately reflect the underlying cognitive processes.

Bowers et al (1990, 1995) explain their results in terms of unconscious spreading activation: exposure to a problem provides cues that activate solution-relevant semantic networks, these cues lead to convergence on a coherent solution, but a solution may only appear in consciousness once coherent activation crosses a certain threshold. It is only when a solution enters consciousness that it may be verified as correct or incorrect. This interpretation concurs with other research that cites implicit processes in insight experiences (e.g. Patalano & Seifert, 1994; Seifert et al, 1995; Yaniv & Meyer, 1987); in each case, unconscious priming promotes subsequent recognition or generation of a correct response. However, Bowers et al stress more than others that unconscious priming may build gradually to a solution rather than 'rest' until an opportunistic event links an existing primed representation to new information. As Durso et al (1994) state:

"Like dynamite, the insightful solution explodes on the solver's cognitive landscape with breathtaking suddenness, but if one looks closely, a long fuse warns of the impending reorganization" (p98).

Insight processes can obviously be dissociated between discontinuous *metacognitive* experiences and continuous unconscious or conscious *cognitive*



change. However, some insight processes may occur discontinuously at a cognitive as well as a metacognitive level. Metcalfe (1998b) suggests that the occurrence of spontaneous restructuring in several physical, living, and perceptual systems indicates that spontaneous cognitive restructuring cannot be dismissed out-of-hand. Some evidence supports this claim. For example, Kounios and Smith (1995) and Smith and Kounios (1996) used speed-accuracy decomposition to measure the amount of partial information available to participants during anagram solution. Anagrams typically produce insight-like metacognitive patterns (Metcalfe, 1986b), but are not technically insight tasks (see Weisberg, 1995b). Kounios and Smith found that participants have little partial information prior to solution, suggesting that cognitive processing was indeed discontinuous rather than continuous. However, the implications of these findings beyond anagram solution have yet to be established. Nevertheless, Metcalfe (1998b) suggests that future definitions of insight should consider the phenomenology of metacognitive experiences in conjunction with the speed of structural changes, and the lack of stable intermediate cognitive structures. Exactly how these latter two constructs might be measured is unclear.

Controversy around insight problem solving obviously centers on whether restructuring occurs spontaneously or gradually, and whether restructuring is mainly due to conscious or unconscious processes. Perhaps all of these processes are possibilities at different times or for different tasks. Disagreements may be due partly to the inconsistent selection of problems in insight studies, and to the lack of a universally accepted definition of insight. For example, discrete categorization of problems as either insight or non-insight does not seem helpful. Insight is arguably a characteristic of particular cognitive or metacognitive processes rather than a static property of any given problem. As such, whether or not any person experiences insight depends on the interaction of that person's processing of a particular problem in a particular situation. Some 'insight' problems may be solved incrementally, and some 'incremental' problems may be solved suddenly (Metcalfe, 1986b; Metcalfe & Wiebe, 1987). Additionally, different problems may share common features, and may be solved by different people in different ways, possibly with both insight and

incremental processes (Schooler & Melcher, 1995; Schooler et al, 1993; Weisberg, 1995b); even problems identified as requiring insight appear to differ amongst each other (Davidson, 1995; Dominowski, 1995). As such, insight and incremental problems may only be defined as such in terms of a continuum of problems, based on the probability that any given problem encourages insightful versus incremental solution experiences (Smith, 1995). Alternatively, Weisberg's (1995b) taxonomy of problems classifies along dimensions of continuity versus discontinuity, restructured versus non-restructured discontinuity, and pure versus hybrid insight.

Clearly issues around the defining of insight and incremental problems and processes are central to continued progress in this research domain. The issue of metacognitive experiences of insight also requires development. Research discussed above demonstrates that, at least some of the time, solutions to insight problems occur gradually at a conscious or unconscious cognitive level. Nevertheless, studies testing phenomenological experiences consistently show that, at a metacognitive level, insight is often experienced with subjective suddenness, even without strictly cognitive-level restructuring. Are these metacognitive experiences enough to consider the accompanying solution processes as 'insightful'? If a person experiences self-reflective awareness of discontinuous and sudden changes in his or her understanding of a problem, have they in effect experienced insight regardless of the underlying processes? The answers to these questions depend on the researcher's theoretical perspective as well as on empirical findings. However, it seems important to consider metacognitive factors given the guiding role of metacognition in the problem solving process. Also, it seems important to clarify our understandings of the role of metacognitive experiences in problem solving before we can more clearly relate metacognitive to cognitive processes.

## THE PRESENT STUDY

The present study expands on the work of Metcalfe (1986a, 1986b), Metcalfe and Wiebe (1987), Jausovec (1994), and Davidson (1995) in order to further clarify the metacognitive processes in problem solving, and to add to the research base on which more informed conclusions can be drawn. However, it employs a somewhat different methodology than the above studies, as outlined below.

This study investigates processes in both well-defined and 'insight'-type problems, two types of problems around which controversy revolves. Both of these problem types are similarly presentable in verbal-puzzle form, such that problem presentation and variations in difficulty are reasonably equivalent across problem type. This study examines both specific metacognitive beliefs in the form of individuals' feeling-of-knowing (FOK) predictions of success on each problem prior to solution attempts, and dynamic metacognitive feeling-of-warmth (FOW) patterns in the form of FOW ratings during solution monitoring. Inclusion of these variables serves as a point of replication and comparison with previous research.

While primarily based on the research paradigm developed by Metcalfe (1986a, 1986b) and Metcalfe and Wiebe (1987), several differences are apparent in the present design. First, a computer program is used to administer the problems and to collect responses to each of the measures. This program allows self-paced engagement with the problems and rating tasks, with the experimenter available for additional support when necessary.

Second, problems are presented sequentially and predictions made intermittently with the corresponding problem, in contrast to Metcalfe's (1986a) and Metcalfe and Wiebe's (1987) procedure of having participants make predictions for all problems in a set before any problems are attempted. Also, predictions are made only in the form of probability estimates, rather than using the feeling-of-knowing ranking procedure. Metcalfe's procedure might not accurately reflect how predictions of problem solving performance are made in the real world; people may be more likely to make predictions sequentially on single problems as they are encountered.

Third, participants are free to provide any rating value for their first warmth rating on each problem, rather than providing a compulsory minimum (rating of 1) value. Having had a chance to read and predict performance on a problem, solvers may have a 'warmer'-than-minimum sense of closeness to the solution at the onset of solution activity. It is possible that the requirement to provide a minimum first rating may unduly influence subsequent ratings, without accurately representing the solver's perceived closeness to a solution. Also, to avoid potential confusion between the prediction given immediately prior to the first warmth rating, the initial warmth rating is given after the first 15-second solution interval rather than before solution activity has commenced.

In addition, this study takes an exploratory step in examining two other features of metacognition: metacognitive evaluation, and the relationship between general 'off-line' metacognitive beliefs and more specific 'on-line' beliefs represented by predictions and evaluations. Metacognitive evaluations are assessed by having participants evaluate their performance (number of problems solved) following completion of each problem set. The inclusion of this measure is to ascertain how evaluations are related to actual performance, prior predictions and monitoring. General metacognitive beliefs are assessed by having participants complete a short paper-and-pencil Problem Solving Inventory (PSI) (Heppner & Petersen, 1982; Heppner, 1988) (see Appendix B). This inventory has been designed to assess an individual's perceptions of his or her own problem solving behaviours and attitudes. The inclusion of the PSI is intended to examine the relationship between general and specific personal metacognitive beliefs, and between general beliefs and actual performance. Previous research has not commonly examined the influence that general beliefs about one's own abilities and performance may have on one's allocation of strategies and effort, and on consequent performance accuracy. Incorporating this variable in the present study may open the way for more detailed assessment in future studies.

The present study also differs from previous research in its reliance on a community sample of adult participants. Most studies, with the exception of Davidson (1995), have tended to use university undergraduate samples. There is a

need to examine whether the processes and relationships observed in student samples are similar or different in any way in the wider community. To this end, adults between the ages of 20 and 55 from the wider community of Palmerston North and surrounding areas were invited to take part in the present study.

### **Objectives**

In a broad sense, the objective of the proposed study is to evaluate the accuracy of individuals' metacognitive experiences and beliefs with respect to their actual problem-solving performance. This incorporates a number of more specific objectives and hypotheses, relating to each of the dependent variables.

The first specific objective is to assess how confident and accurate people are in predicting the outcome of their problem solving attempts, and evaluating whether predictive accuracy differs depending on type of problem, whether well-defined or 'insight'-type problems. The second objective is to assess how well people monitor their problem solving during solution of a problem, through the provision of progressive warmth ratings throughout solution episodes, and whether metacognitive experiences associated with monitoring differ depending on problem type. The third objective is to assess the accuracy of individuals' evaluations of their solution performance on the two types of problems. The fourth objective is to assess the relationship between participants' general 'off-line' beliefs and their more specific 'on-line' metacognitive predictions and evaluations.

### **Hypotheses**

#### **i. Performance**

**Hypothesis 1.** Performance (number of correct solutions) will be better on well-defined problems than on insight problems. Insight problems are expected to be more difficult because they presumably encourage development of inaccurate representations that impede identification of crucial information leading to solution.

#### **ii. Predictions**

**Hypothesis 2.** Predictive accuracy will be greater for well-defined than for insight problems, in terms of both micro- and macro-prediction. Participants should

be over-confident in their predictions for both problem types, but will be more over-confident for insight than for well-defined problems, as found by Metcalfe and Wiebe (1987). With respect to macro-predictive accuracy, mean predictions should overestimate mean performance to a greater degree for insight than for well-defined problems. With respect to micro-predictive accuracy, correlations between prediction and performance should be greater for well-defined than for insight problems. These tendencies fit with the pattern of cognitive optimism observed across cognitive tasks (Metcalfe, 1998a). Because insight problems usually encourage an initially inaccurate representation (whether subsequently solved correctly or incorrectly), performance should be less predictable than for well-defined problems wherein solution depends on following a correct solution path based on a clearly-prescribed representation.

### **iii. Monitoring: Feelings-of-warmth**

**Hypothesis 3a.** Within the insight problem set, correct solutions should manifest more ‘insight’ than ‘incremental’ warmth patterns.

**Hypothesis 3b.** Within the insight problem set, incorrect solutions should manifest more ‘incremental’ than ‘insight’ warmth patterns.

**Hypothesis 3c.** Between problem sets, more ‘insight’ warmth patterns should be manifested by correct insight solutions than by correct well-defined solutions.

**Hypothesis 3d.** Between problem sets, more ‘incremental’ warmth patterns should be manifested by correct well-defined solutions than by correct insight solutions.

A main effect of problem type on solution accuracy is expected, as participants’ feelings-of-warmth should be more incremental for well-defined problems than for insight problems, and less gradual but more sudden prior to solution for insight problems than for well-defined problems (Metcalfe, 1986b; Metcalfe & Wiebe, 1987; Jausovec, 1994; Davidson, 1995). An interaction between



problem type and solution correctness is also expected; warmth patterns should be incremental for both correctly and incorrectly solved well-defined problems, but should differ between correctly and incorrectly solved insight problems.

Specifically, correct insight problems should be solved with more often with 'insight' patterns than with 'incremental' patterns. In contrast, incorrectly solved insight problems should be accompanied by 'incremental' patterns more often than by 'insight' patterns as the solver follows an obvious solution path suggested by an inaccurate problem representation that ultimately does not bear the correct solution.

#### **iv. Evaluations**

**Hypothesis 4.** Evaluations (estimating number of correct solutions) should be more accurate for well-defined than for insight problems; for both absolute and relative evaluation measures. People will be more overconfident in their evaluations of insight than of well-defined problems.

Because well-defined solutions are typically based on a clearly specified problem representation, participants should be reasonably aware of whether or not they are using the correct strategy, whether or not they have achieved the correct solution, and therefore should be reasonably able to indicate how many problems were solved correctly. Because insight problems typically encourage people to follow an incorrect but obvious solution path, solvers may be convinced that they have solved a problem when they in fact have not; therefore, solvers should be less accurate in estimating how many problems they have actually solved.

#### **v. PSI Scores : Correlating general with specific beliefs**

**Hypothesis 5a.** Significant correlation, of low-to-moderate magnitude, between PSI scores and mean predictions for well-defined set, insight set, and both sets combined. Correlations will be negative in valence.

**Hypothesis 5b.** Significant correlation, of low-to-moderate magnitude, between PSI scores and mean evaluations for the well-defined and insight sets. Correlations will be negative in valence.



**Hypothesis 5c.** For both predictions and evaluations, correlations should be stronger with the problem solving confidence and personal control scales than with the approach-avoidance scale.

One's general beliefs about problem solving and his or her appraisals of personal capabilities should inform, and therefore correlate to some degree with, confidence in one's performance on specific problems. Arguably, however, predictions and evaluations are also informed by other sources of information, including immediate information acquired from the presentation of a problem or type of problem, and appraisals of one's immediate performance on a specific problem. Also, the PSI has been developed from a personal/social problem solving domain rather than a hypothetical verbal puzzle domain, so may not be completely appropriate for the current context, limiting association with the specific measures.

Predictions and evaluations in both problem sets will be more related to the Problem Solving Confidence and Personal Control scales of the PSI than to the Approach/Avoidance scale, given the content of these scales. General appraisals of confidence and control should influence one's appraisals of future and past performance on any specific problem. In the present study, tendencies to approach or avoid problems should be less relevant. The puzzles should provoke little anxiety, so should not provoke avoidance feelings.

It is unclear whether general beliefs should be more related to predictions than to evaluations, or more related to well-defined than to insight related beliefs, so no concrete predictions are made in this regard.

## METHOD

### Participants

Participants aged between 20 and 55 years were recruited through advertisements (see Appendix A) placed in a local community newspaper, and through word-of-mouth within the social networks of the researcher. In total, thirty-six participants (15 male, 21 female) took part in the first session of the study; however, four people were unable to return for the second session and their data were excluded from further consideration. On completion of both sessions participants were offered \$10 compensation in remuneration for their time and costs incurred in taking part; twelve people elected to receive this compensation.

Of the thirty-two participants (13 male, 19 female) to complete the study, the average age was 33.66 years ( $SD = 9.80$ ) for the total sample, 32.69 years ( $SD = 10.48$ ) for males, and 34.32 years ( $SD = 9.55$ ) for females. The majority of participants (13 male, 16 female) were from New Zealand European ethnic backgrounds, with others classifying themselves as both New Zealand European and Maori (2 females), and as German European (1 female).

Nine participants (4 male, 5 female) were in full-time paid employment, ten (5 male, 5 female) were in part-time paid employment, three (all female) were in no paid employment, and ten participants (4 male, 6 female) were full-time students. Across the sample seventeen people (6 male, 11 female) were currently studying either full- or part-time. In terms of the highest level of education attained by participants, seven had completed a postgraduate qualification, thirteen had completed an undergraduate qualification, three had completed a trade or vocational qualification, six had undertaken some tertiary education, two had gained Sixth Form Certificate, and one had completed School Certificate.

### Materials

#### 1. Problems

Two sets of problems were used in the present study: well-defined and insight problem sets (see Appendix B for complete list). The problems were selected

on the basis that they were of similar length, could be read within fifteen seconds, and that they were all verbal in nature. Verbal-sentence problems were chosen in both cases so as to standardize the surface features of the problems across the two problem sets. This style of presentation involves problems in the form of short paragraphs of only two or three sentences in length. Some particular information is provided to the solver followed by a question that requires the solver to use the provided information (together, in many cases, with additional knowledge or assumptions from long-term memory) to find a correct answer to the question.

The decision to include only verbal problems necessitated the exclusion of other problem forms that have been previously studied, including pictorial and hands-on problems. Commonly studied well-defined problems in these alternative presentation forms include, for example, the missionaries and cannibals problem and the Tower of Hanoi. Common insight problems presented in pictorial form include the nine-dot problem and the triangle of circles problem. Exclusion of these forms allowed participants the freedom of working on problems mentally while recording warmth ratings by hand, without requiring the use of hands for tasks related to the solution process. Other problem forms were deliberately excluded, such as anagrams and algebra problems, and problems requiring particular domain knowledge in any area. Excluding such problems from the present study has the disadvantage of eliminating the investigation of metacognitive processes in other important problem forms. However, the advantages of using verbal problems alone include ease of presentation via computer screen, ensuring that participants could represent and operate on all problems mentally without recourse to some physical representation or manipulation, and ensuring that presentation form across both problem sets was as similar, and as comparable, as possible. It was decided that these advantages outweighed the costs of restricting the materials to a single presentation form.

Most of the present problems were drawn from previous studies employing similar tasks. In a number of cases, additional problems were selected from additional sources on the basis that they shared similar properties with those problems from previous research. Each set consisted of two practice problems and eight main trial problems.

## **2. Problem Solving Inventory (PSI)**

The Problem Solving Inventory (PSI) was employed in the present study to examine the relationship of 'off-line' general metacognitive beliefs with participants' 'on-line' specific beliefs of performance in the form of predictions and evaluations for problem solving episodes. The PSI (Appendix B) is a 35 item self-report, paper-and-pencil test that assesses "an individual's perceptions of his or her own problem-solving behaviours and attitudes" (Heppner, 1988: p1). It is designed to measure an individual's appraisals of his or her capabilities, rather than actual problem-solving skills, therefore representing a measure of metacognitive-level beliefs.

Respondents rate their level of agreement/disagreement with each statement on a 6-point scale (1= strongly agree; 6= strongly disagree). Low scores indicate positive self-appraisals of problem solving ability; high scores represent negative self-appraisals of ability. Administration time is 10 to 15 minutes. Four scores are derived, including a Total score and three scale scores: Approach-Avoidance Style (AA), Problem-Solving Confidence (PSC), and Personal Control (PC). The AA scale (16 items) measures "a general tendency to approach or avoid problem-solving activities" (p2). The PSC scale (11 items) measures "self-assurance while engaged in problem-solving activities" (p1). The PC scale (5 items) measures "being in control of one's emotions and behaviours while solving problems" (p2). Scales were established through factor analysis. Items relate to each of five common phases identified in models of problem solving: "general orientation, problem definition, generation of alternatives, decision making, and evaluation" (p7). Items from each phase are randomly distributed amongst each of the three scales.

Internal consistency of the PSI is high, with Cronbach's alphas for the scales ranging from .72 to .90. Test-retest reliability is also high, with coefficients ranging from .83 to .89 over a period of two weeks (Heppner & Petersen, 1982) and from .44 to .65 over a period of two years (Reeder, 1986, cited in Heppner, 1988). The PSI has also exhibited reasonable levels of construct and discriminant validity (Heppner, 1988).

The PSI has been used primarily in research investigating the cognitive, behavioural, and emotional correlates of coping with real life problems. Thus, it relates more closely to social and personal problem solving than to abstract, hypothetical well-defined and insight-type problems. However, because it is anticipated that there will be some relationship in a person's problem solving processes between more abstract problems and more social-personal problems, a low-to-moderate relationship is expected between participants' PSI scores and their predictions and evaluation for the present tasks. Even if there is no relationship between these measures it would still be important to know there is little relationship, at least with the PSI as the measure of general beliefs.

### **Apparatus**

A Toshiba T2110CS laptop computer with a monochrome LCD screen was used. A DOS-run program was designed to present the participant with summarized procedural instructions and the to-be-solved problems, and to record the participant's predictions, feeling of warmth ratings, and evaluations. The program was written by Harvey Jones, School of Psychology, Massey University.

### **Procedure**

Participants were tested individually, across two separate sessions spaced approximately one week apart. In each session either the well-defined or insight problem sets were presented, with the order of set presentation counterbalanced alternately across participants. The nature of the two problem sets was not disclosed to participants prior to their completing all aspects of the study; however, they were informed that the two sets did differ in some way. Testing was conducted free from distractions within a research room at the School of Psychology, Palmerston North, Massey University.

#### **1. Preliminary Procedures**

In the first session, formal consent procedures preceded the experiment administration. Participants had previously had an information sheet (Appendix A)

mailed to them, but were given the opportunity to re-read this in the presence of the experimenter and to ask any questions for clarification. The experimenter emphasized the participant's ethical rights and the participant signed a consent form (Appendix A) when satisfied with the information provided. An answer booklet was provided within which participants were to record their solutions to the presented problems. The front page of this booklet contained a short section of instructions regarding the booklet's use, followed by a number of biographical items for the participant to indicate their sex, age, ethnicity, employment status, occupation, and highest level of education.

The participant was seated at a desk, upon which was placed the lap-top computer. They read a uniform instruction sheet (Appendix B) that described in detail the procedure of the study and the nature of responses to be made upon presentation of the problem set. A chance was given to ask questions about the procedure before the experimenter started the computer program to commence presentation of the problem set. An initial information screen summarized the instructions that the participant had read on the instruction sheet. On-screen instructions guided the participant through all aspects of the program, allowing self-paced progress, and indicating what responses to make and when to make these responses. The participant had the chance to work on two practice problems before attempting the eight main trials. The practice problems were presented in the same order across participants; the main problems were randomly ordered across participants.

## **2. Problem Presentation**

The procedure was replicated for each problem. The participant initially had 15 seconds to read the problem before it was removed from the screen; the participant had been instructed that they could think about how well they believed they could solve the problem, but without actually attempting to solve the problem at this point. He or she then provided a prediction of the likelihood of solving the problem correctly, by pressing a number between 1 and 10 (represented by a '10' label affixed to the '0' key) on the top-row of the keyboard; a rating of 1 indicated



that he or she was not at all confident of solving the problem, while a rating of 10 meant that he or she was completely confident of solving the problem. As soon as this prediction had been made, the problem was returned to the screen for solution.

From this point, the participant had four minutes to work on the problem. Every 15 seconds a beep sound from the lap-top signalled that the participant should enter a feeling-of-warmth (FOW) rating, indicating how close the participant believed he or she was to solving the problem at that point. A number on the keyboard, from 1 to 10, was to be pressed; a rating of 1 was 'very cold' (the participant had no idea of the answer), a rating of 10 was 'very hot' (the participant was certain of knowing the answer). The participant had been instructed to make these ratings as quickly as possible, based on their immediate impression of how close they were to the answer at that moment. FOW ratings were made every 15 seconds until either the participant indicated that they had solved the problem or the four minute time limit had expired.

The participant could indicate completion of the problem at any time by pressing the 'F' key to finish; the problem was removed from the screen and the participant was instructed to enter a final warmth rating of 10, indicating that the participant believed the problem to be solved. If the time limit had expired, the problem was removed and a final FOW rating was to be made, in this case indicating how close to the solution the participant believed him- or her-self to be at this endpoint (a 10 rating was not required). The participant was then instructed to write their solution in the space provided in the answer booklet or, if no solution had been reached, to indicate their best-possible answer or to write that they did not know the answer. When ready, pressing the space-bar allowed the participant to proceed to the next problem.

Once all eight problems in the set had been attempted, the participant was directed to record two final evaluations. First, an absolute evaluation indicated how many problems in the set he or she believed had been solved correctly, represented by pressing a number key between 1 and 8. Second, a relative evaluation indicated how well they believed they had performed on the set compared to 'most other people', by pressing a number from 1 to 5; a rating of 1 indicated that their



performance was “well below average”, a rating of 2 “a little below average”, a rating of 3 “average”, a rating of 4 “a little above average”, and a rating of 5 “well above average”. With both evaluation ratings made, the computer program terminated itself and the participant was thanked for their time. The experimenter informed the participant that at this point they would not be told how well they had performed on the problems, nor would they be told the actual answers to the problems. Instead, this information would be disclosed at the completion of the study, with a results summary to be sent to interested participants.

### **3. Second Session**

The second session involved a similar procedure to the first, with several exceptions. First, some participants were asked to complete the Problem Solving Inventory (PSI) prior to engaging in the computer task. Because the PSI is sold in packs of 25 copies, for financial reasons only one pack could be purchased for the present research. Twenty-five participants were randomly chosen from the total sample to complete the inventory. Instructions were provided on the PSI form. Participants completed the form at their own pace, with most finishing within ten minutes. All participants were given the computer task to complete, and were presented with the problem set alternate to that presented in the first session. Following completion of the computer task, participants were again informed that they would not yet be given access to their success rate or to the actual answers. However, they were given a ‘Thank You Sheet’ (Appendix A) that explained the study in greater detail, including the difference between the two problem sets and the issues that the experimenter was examining. An opportunity was given to sign up for a results summary to be sent to participants at the completion of the study. Also, the participant was invited to take \$10 as compensation for the time and any costs incurred in his or her participation. The participant signed a payment form (Appendix A) indicating whether or not they accepted this compensation.

## RESULTS

The 32 participants who completed both well-defined and insight problem sets provided a potential pool of 512 (256 well-defined, 256 insight) problems for analysis. First, preliminary analyses provide an overall picture of performance through solution accuracy, individual performance, and time to solution. Second, participants' predictions are analysed in terms of absolute confidence levels and predictive accuracy in relation to actual performance. Third, feelings-of-warmth, representing on-line metacognitive monitoring of problem solving, are assessed in relation to mean warmth ratings across solution intervals, distributions of insight and incremental patterns, and differential and angular warmth measures. Fourth, absolute and relative evaluations are analysed in terms of post-performance confidence on the well-defined and insight sets. Finally, the relationships between general and specific problem solving beliefs are explored by correlating PSI scores with predictions and evaluations. All statistical tests are two-tailed, unless otherwise noted.

### 1. Performance Analyses

#### a. Solution accuracy

Solutions were graded as correct, incorrect (a solution was given but this was not the correct solution), unsolved (no solution was given, or the participant admitted to not having an answer to give), or excluded (the participant already knew the solution prior to the experiment). From the pool of 512 problems, 222 problems (43.4%) were solved correctly, 240 (46.9%) were solved incorrectly, 32 (6.3%) were unsolved, and 18 (3.5%) were already known to participants and were excluded from further analyses (Figure 1). The well-defined problems were solved correctly more often than the insight problems,  $\chi^2(1) = 7.93, p = .005$ . As seen in Figure 1, 50% of the well-defined problems were solved correctly. Within the insight set, 36% of problems were solved correctly, with half solved incorrectly. This reinforces findings from earlier studies (e.g. Metcalfe, 1986b; Metcalfe & Wiebe, 1987) demonstrating that these problems tend to be quite difficult. Only in a small number of cases were no solutions provided by participants, while all the problems excluded at this point due to previous knowledge came from the insight set.

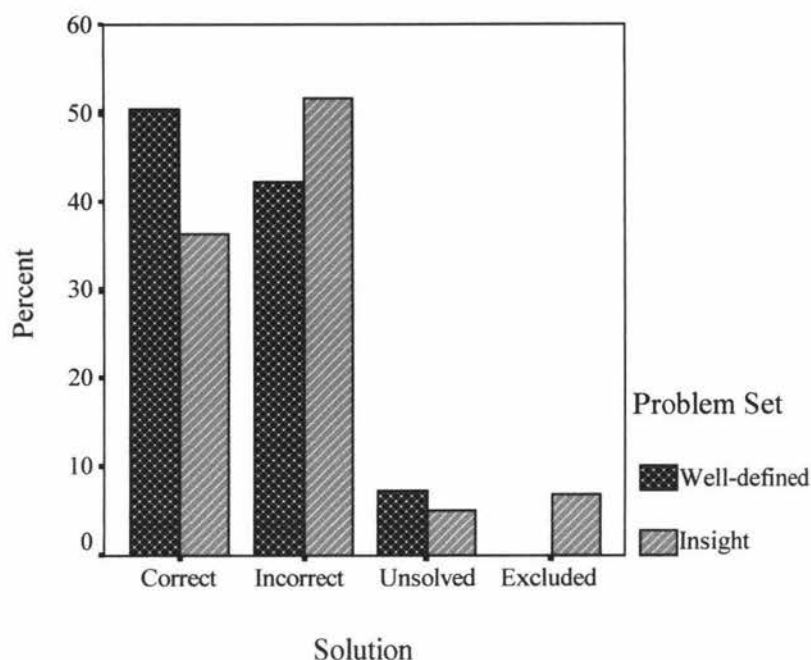


Figure 1. Percentage of solutions for well-defined and insight problem sets.

Solution accuracy was also assessed across problems, providing an indication of relative problem difficulty (see Appendix C, Table 1). No single problem was solved by all participants. The easiest problems came from the well-defined set, with the most frequent correctly solved problems being the ‘Race Results’ (WD problem 8) and the ‘Water Jugs’ (WD 1) problems. Within the insight set, the ‘Tower’ problem (INS 1) and the ‘Water Lilies’ problem (INS 2) were solved most frequently. The least solved problem was the ‘Frog’ problem (WD 2), with the ‘Chain’ (INS 5) and ‘Triplets’ (INS 8) problems each solved by only four participants. Almost half of the excluded problems were from one source, the ‘B.C.’ problem (INS 7); the ‘Water Lilies’ problem (INS 2) was already known to five people.

In all of the following analyses, only the correctly and incorrectly solved problems in each problem set were considered.

#### **b. Time to solution**

Time to solution was measured in terms of the number of feeling-of-warmth (FOW) intervals to solution (Figure 2). The maximum possible number of fifteen-second intervals over the allowable solution time of four minutes was 15. A 2 x 2 (problem set x

solution) analysis of variance (ANOVA) was conducted on the mean numbers of warmth ratings. There was a main effect of solution,  $F(1, 461) = 14.624, p < .001$ , with, on average, fewer intervals needed to complete correct ( $M = 5.09, SD = 3.653$ ), compared to incorrect ( $M = 6.53, SD = 4.794$ ), solutions. There was no effect of problem set,  $F(1, 461) = 2.253, p = .134$ , with well-defined problems ( $M = 6.02, SD = 4.286$ ) and insight problems ( $M = 5.65, SD = 4.396$ ) solved within similar interval periods. No interaction was significant between problem set and solution,  $F(1, 461) = 2.948, p = .087$ .

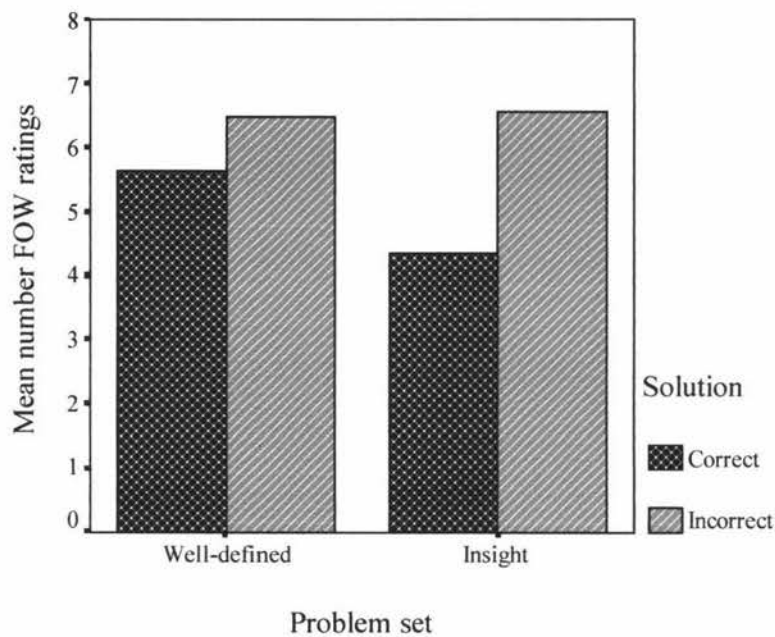


Figure 2. Mean number of FOW ratings made across problem set and solution accuracy.

Most problems were solved within only a few rating intervals, such that distributions of FOW intervals skewed towards the lower interval periods (Appendix C, Figures 1(a), (b), (c), (d)). Distributions of FOW intervals were more positively skewed for correct insight solutions than for correct well-defined solutions.

### c. Individual performance

Individual performance was assessed in terms of the total number of correct solutions both within each problem set and across the two sets combined. Overall, the

mean number of correct solutions for participants was 6.9 out of a possible 16. Most participants correctly solved either 6 or 8 problems (28% and 25% of participants respectively). On average, the number of correctly solved problems was greater in the well-defined ( $M = 4.03$ ,  $SD = 1.60$ ) than in the insight set ( $M = 2.91$ ,  $SD = 1.45$ ),  $t(31) = 4.03$ ,  $MSe = .28$ ,  $p < .001$ . Within the well-defined set, modal performance was four correct solutions from the total of eight problems, with 75% of the sample attaining between three and five correct solutions. Modal performance on insight problems was three correct solutions, with 65% of the sample achieving between two and four correct solutions.

## **2. Predictions: Confidence and Predictive Accuracy**

Only correctly and incorrectly solved problems with at least two FOW ratings were included in analyses of the predictions. Protocols with only one FOW rating (solved within 15 seconds) were excluded to avoid the possibility that participants may have actually known the answers to these problems although not admitting so to the experimenter. This was particularly likely for the insight set, where the modal solution time was within one rating interval. This left a pool of 395 problems, including 208 well-defined (116 correct, 92 incorrect) problems and 187 insight (68 correct, 119 incorrect) problems. Predictions were assessed in terms of absolute confidence levels, and predictive accuracy in relation to actual performance.

### **a. Confidence**

The mean predictions for both problem sets are shown in Table 1. A 2 x 2 (problem set x accuracy) ANOVA was conducted to test the differences between these means. There was a main effect of accuracy,  $F(1, 391) = 10.53$ ,  $p < .01$ , indicating that predictions were higher prior to a correct than an incorrect solution. There was no effect of problem type,  $F(1, 391) = .32$ ,  $p = .57$ , indicating that predictions were similar for both well-defined and insight problems, and no interaction between the factors,  $F(1, 391) = .001$ ,  $p = .97$ . In absolute terms, predictions indicated that participants were, on average, about 50 percent certain of achieving the correct solution to the problems, regardless of problem type.

To examine the pattern of predictions more closely, the frequency of predictions made at each rating level (from 1 to 10) was calculated (Appendix C, Table 2). Participants tended to be slightly more conservative over the range of predictions for the insight problems compared to the well-defined problems. For the well-defined problems, 55% of ratings were of low confidence (ratings of 1 to 5) compared to 45% of high confidence (ratings of 6 to 10). For the insight problems, 62% of ratings were of low confidence compared to 38% of high confidence. Differences between sets were not significant. When a low confidence rating was given for well-defined problems, a correct solution followed in 50% of cases; for insight problems, a correct solution followed in only 30% of cases. When a high confidence rating was given, a correct solution followed in 63% of well-defined cases and in 47% of insight cases.

Table 1. *Mean (M) predictions for incorrectly and correctly solved problems in well-defined and insight sets*

Problem Set	Solution	N	M	SD	MSe
Well-defined	Correct	116	5.7	2.6	.24
	Incorrect	92	4.9	2.7	.28
	Combined	208	5.3	2.7	.18
Insight	Correct	68	5.6	2.4	.29
	Incorrect	119	4.7	2.6	.24
	Combined	187	5.0	2.5	.19

**b. Predictive accuracy**

The accuracy of participant’s predictions was assessed in two ways: first, by calibrating the degree of overestimation between predictions and actual performance; second, by calculating the correlation between predictions and performance.

Predictions and performance were calibrated by first calculating, in each problem set for each participant, the proportion of correct solutions and the mean probability estimate based on predictions. Prediction ratings (on a 10-point scale) were converted into probability estimates by dividing the participant’s mean prediction by 10, providing estimate scores between .1 and 1.0. The proportion of correct solutions was calculated by

dividing the number of correct solutions in each set by 8 (the total number of problems presented in each set). Thus, solution proportions could range from .0 (no correct solutions from eight problems) to 1.0 (eight correct solutions from eight problems). An ANOVA was used to compare prediction probability estimates and correct solution probabilities across both problem sets. There was an effect of estimate such that, across both problem sets, mean prediction estimates were greater than solution probabilities,  $F(1, 31) = 16.20$ ,  $MSe = .986$ ,  $p < .001$ . Essentially, predictions of success were greater than the rate of actual correct solutions. However, this effect of overestimation was greater in the insight set than in the well-defined set, demonstrated by a significant interaction between estimation and problem set,  $F(1, 31) = 12.94$ ,  $MSe = .215$ ,  $p = .001$  (see Figure 3).

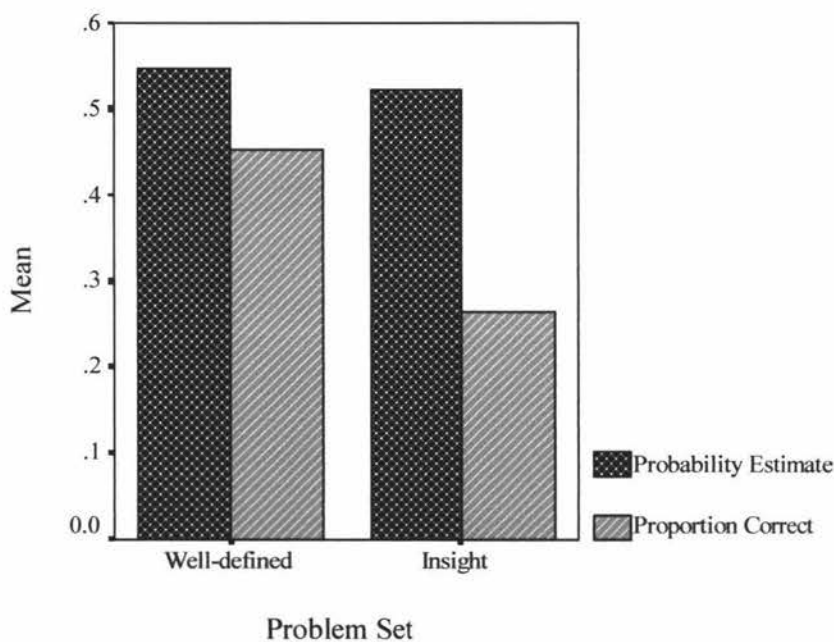


Figure 3. Mean prediction probability estimates and proportions of correct solutions for well-defined and insight sets.

Predictive accuracy was also assessed by correlating predictions with solution accuracy. Goodman-Kruskal gamma correlations were calculated between predictions and solution accuracy, following the procedure established by Metcalfe (1986b). The gamma ( $G$ ) statistic is a non-parametric rank-order correlation suitable for ordinal data



that contains many ties (Nelson, 1984; Siegel & Castellan, 1988), and is well established for research involving feeling-of-knowing (FOK) judgements (Butterfield, Nelson, & Peck, 1988; Metcalfe, 1986a). For each participant, contingency tables enabled calculation of a gamma correlation between predictions (rank-ordered from lowest to highest) and frequencies of incorrect and correct solutions. This calculation was repeated for both problem-sets. Gamma can only be calculated where a participant has at least one correct and at least one incorrect response in each set, and where equal predictions are not given for all problems in the set. Due to these restrictions, only 23 participants had summary gammas in both the well-defined and insight sets. The mean gamma across participants was  $G = .33$  for well-defined problems, and  $G = .19$  for insight problems, indicating that predictions were more accurate for well-defined than for insight problems. However, the difference between the two gammas was not significant,  $F(1, 44) = .19, p = .67$ . Also, mean gammas were not significantly different from 0 by one-tailed t-tests, either for the well-defined gammas [ $t(22) = 1.10, MSe = .30, p = .14$ ] or the insight gammas [ $t(22) = 1.37, MSe = .14, p = .09$ ].

In summary, greater predictive overestimation was observed for insight than for well-defined problems, although correlations between predictions and performance were not significant for either problem set. Altogether, analysis of predictions produced little conclusive differences between problem sets.

### 3. Metacognitive Monitoring: Feeling of Warmth (FOW) Analysis

Differences between problem sets were next analysed in relation to dynamic 'feeling of warmth' monitoring of solution progress. For analyses of the FOW ratings, only protocols with at least four warmth ratings were included; that is, those protocols with at least three ratings from the first to the penultimate rating. It was considered that meaningful patterns of warmth ratings would not be discernible for protocols with any fewer than four ratings. The remaining pool of 281 problems included 150 well defined (86 correct, 64 incorrect) and 131 insight (45 correct, 86 incorrect) problems. The following analyses of the warmth ratings follow the procedures established by Metcalfe (1986a, 1986b) and Metcalfe and Wiebe (1987). Warmth ratings were examined in terms of differences in mean ratings across intervals, differences in proportions of

insight and incremental patterns, and correlations of differential and angular warmth measures.

#### **a. Mean FOW ratings across solution intervals**

The mean FOW rating made in each of the 15-second intervals was calculated for each participant for correct and incorrect solutions in both problem sets. Evaluating the mean ratings across solution intervals has been identified as an imprecise, if not misleading, measure of changes in warmth across the solution process (Weisberg, 1992), for several reasons. Because problems were solved over differing numbers of intervals, the mean value for any one interval contains ratings that occur at different points within different rating protocols; for example, the mean value of the fourth interval prior to solution contains the first rating in a protocol of five ratings, the second rating in a protocol of six ratings, and the sixth rating in a protocol of ten ratings. Thus, the mean for any interval will contain a potential mix of lower (earlier) and higher (later) values, obscuring the actual overall pattern of ratings made. For the present data set, this issue is compounded because participants were free to make their first rating as any value they pleased; hence, first ratings were not universally at the lowest extreme (1) for all protocols. Nevertheless, in line with previous research, an attempt was made to gain an overall perspective of the successive intervals.

However, the present analysis differs in one respect from conventional analyses. Whereas previous research has assessed mean interval ratings for all problems with four or more ratings the present study examines a narrower band of protocols, being the subset of protocols with four, five, and six ratings only. These protocols were the most common of all protocols with four or more ratings; also, as these fall within a narrower band, it was hoped that this subset would be less affected by variance in interval ratings across the range of protocol sizes. This subset contained 124 problems, including 47 with four ratings, 44 with five ratings, and 33 with six ratings, and comprising 67 well-defined (44 correct, 23 incorrect) protocols and 57 insight (23 correct, 34 incorrect) protocols (Appendix C, Table 3 contains the complete table of overall means for all protocols with four or more ratings).

i. Warmth as a function of correctness in insight problems.

First, warmth patterns for correct and incorrect solutions were assessed in the insight set alone. Ten participants had at least one correct and one incorrect solution within this subset. The mean FOW ratings in the three intervals prior to solution are displayed in Table 2. A repeated-measures ANOVA indicated no main effect of solution accuracy,  $F(1, 9) = .14$ ,  $MSe = .63$ ,  $p = .72$ ; means in each interval were similar regardless of solution, with the mean warmth rating across the three pre-solution intervals being 4.85 and 5.06 for correct and incorrect solutions, respectively. A main effect of interval indicated that ratings increased as solution approached,  $F(2, 18) = 9.29$ ,  $MSe = 9.99$ ,  $p < .01$ . However, there was no significant interaction between solution and interval,  $F(2, 18) = .20$ ,  $MSe = .16$ ,  $p = .82$ ; the increase in warmth from fourth-last to second-last ratings was similar for correct (1.50) and incorrect (1.15) solutions. Finally, mean warmth in the interval preceding solution was 5.8 prior to both correct and incorrect solutions.

Table 2. Mean warmth ratings in last four intervals for correct and incorrect solutions in insight and well-defined sets, for problems solved within 4, 5, or 6 intervals only

Problem Set	Solution	N	Interval			Final (solution) rating
			Solution - 3	Solution - 2	Solution - 1	
Insight	Correct	10	4.25	4.55	5.75	10.00
	Incorrect	10	4.62	4.78	5.77	10.00
Well-defined	Correct	16	5.44	6.22	7.29	10.00
	Incorrect	16	5.04	5.35	6.88	10.00

ii. Warmth as a function of correctness in well-defined problems.

Within the well-defined set, sixteen participants had at least one correct and one incorrect solution. A repeated-measures ANOVA found no main effect of solution accuracy,  $F(1, 15) = 1.36$ ,  $MSe = 7.59$ ,  $p = .26$ ; mean warmth ratings across the three pre-solution intervals were 6.32 and 5.76 for correct and incorrect solutions, respectively

(Table 2). There was a main effect of interval, such that warmth increased across intervals towards solution,  $F(2, 30) = 17.40$ ,  $MSe = 28.56$ ,  $p < .001$ , with an increase in warmth from fourth-last to second-last intervals of 1.84 for correct solutions and 1.83 for incorrect solutions. No interaction was apparent between solution and interval,  $F(2, 30) = .40$ ,  $MSe = .55$ ,  $p = .67$ . Immediately prior to solution, participants gave mean warmth ratings of 7.3 and 6.9 for impending correct and incorrect solutions, respectively.

iii. *Warmth as a function of problem type.*

Warmth ratings were analysed as a function of problem type for correct solutions only, to determine whether well-defined problems were solved more incrementally than insight problems. In this analysis, mean FOW ratings for all protocols solved over four or more intervals were compared; mean ratings for these protocols over the three pre-solution intervals are shown in Table 3. Ratings were, on average, greater for well-defined ( $M = 5.85$ ) than for insight ( $M = 5.13$ ) problems, but this effect was short of significance,  $F(1, 22) = 3.18$ ,  $MSe = 18.06$ ,  $p = .09$ . There was a main effect of interval, with warmth increasing towards solution,  $F(2, 44) = 30.93$ ,  $MSe = 32.33$ ,  $p < .001$ , but there was no interaction between problem type and interval,  $F(2, 44) = .71$ ,  $MSe = .52$ ,  $p = .50$ . The analyses were repeated for protocols solved over four, five, and six interval periods only, but produced similar results.

Table 3. *Mean warmth ratings in last four intervals for correctly solved insight and well-defined problems*

Problem type	N	Interval			Final (solution) rating
		Solution – 3	Solution – 2	Solution – 1	
Insight	23	4.38	4.91	6.10	9.71
Well-defined	23	5.04	5.87	6.65	9.62

Note: Means of final ratings are less than ten because a minority of protocols reached the four-minute time limit and were given final ratings less than 10 (participants were still unsure of their answer once time had expired).

### **b. Warmth pattern analyses**

The use of differential warmth measures corrects for the effect on the FOW interval means of problems with varying numbers of FOW ratings. Differential warmth indicates the increase in warmth across the solution period until immediately (within 15 seconds) prior to solution. It was calculated, for each solution protocol with at least four FOW ratings, as the difference between the first and penultimate ratings. Essentially, it allows analysis of warmth patterns *forward from the first to penultimate ratings* rather than *backward from the penultimate to first ratings* as in the previous analysis, the latter being an earlier procedure of Metcalfe's that was justifiably criticized by Weisberg (1992). Hence, differential warmth gives a more accurate indication of progress on each problem than does the averaged interval ratings above.

First, the differential warmth measure was used to assess the proportions of problems that were solved either with an 'insight' pattern or an 'incremental' pattern of responding. Metcalfe (1986b) defined an insight pattern as a sequence of warmth ratings that increased by no more than one rating point from first to penultimate ratings. The rationale here is that if participants were metacognitively aware of any solution progress across the solution period, such that their warmth ratings increased by more than one point, then one could argue that the problem was actually solved in a gradual, incremental fashion, rather than in a sudden fashion as would be suggested by the construct of insight. An incremental pattern was defined as a sequence of warmth ratings that increased by more than five points from first to penultimate ratings. Metcalfe's analysis examined these patterns for correct and incorrect insight-type problems only. Davidson's (1995) and the present study's analyses consider differential warmth patterns for both well-defined and insight problems.

All protocols with at least four FOW ratings were again considered. However, the present analysis incorporated further differences. As participants were free to choose their first ratings, and these varied widely across the scale from one to ten, protocols were anchored at different rating points (Appendix C, Table 4 shows the frequency of first ratings for each problem set). It was believed that these anchor points might have different influences on subsequent rating patterns, so protocols were divided into those with either low (first rating 1-3) or high (first rating 4-6) anchor points. Because

meaningful patterns of insight or increment cannot be identified for protocols with first ratings of seven or higher, such protocols (15% of the problem pool) were excluded from the following analysis. This left a pool of 238 problems. Furthermore, to accommodate the number of protocols with high anchored first ratings, the incremental warmth pattern was re-defined in all cases as an increase of three or more rating points from first to penultimate rating. Tables 4 and 5 show, for each problem set, the proportions of each solution type whose protocols fall into the insight and incremental patterns, for both low and high anchoring of the first rating. (Appendix C, Table 5 (well-defined) and Table 6 (insight) show the frequency of differential warmth values for the respective problem sets).

Table 4. *Proportion of problems in insight set showing an insight or incremental pattern of warmth ratings, for low- and high-anchored protocols*

First Rating Anchor	Solution	N	Pattern		
			Insight	Incremental	Other
Low (FOW = 1-3)	Correct	17	.29	.59	.12
	Incorrect	32	.44	.50	.06
High (FOW = 4-6)	Correct	22	.41	.27	.32
	Incorrect	43	.51	.33	.16

Note: Within each anchor level, proportions sum across patterns for each solution type. 'Other' solution pattern represents increase of two rating points across a protocol.

Table 5. *Proportion of problems in well-defined set showing an insight or incremental pattern of warmth ratings, for low- and high-anchored protocols*

First Rating Anchor	Solution	N	Pattern		
			Insight	Incremental	Other
Low (FOW = 1-3)	Correct	27	.26	.67	.07
	Incorrect	35	.46	.40	.14
High (FOW = 4-6)	Correct	42	.38	.43	.19
	Incorrect	20	.45	.30	.25

Note: Within each anchor level, proportions sum across patterns for each solution type. 'Other' solution pattern represents increase of two rating points across a protocol.



Difference in proportion tests revealed significant differences for only a few comparisons. Amongst correct insight solutions, a greater proportion of protocols with insight rather than incremental patterns were expected. For low-anchored protocols there were in fact more problems solved with an incremental than an insight pattern,  $Z = 1.75$ ,  $p = .04$  one-tailed. For high-anchored protocols the expected difference transpired but was not significant. Comparing between patterns for incorrect insight solutions, greater proportions of incremental rather than insight patterns were expected. Again, the opposite difference transpired for high-anchored protocols exhibiting significantly more insight patterns than incremental patterns,  $Z = 1.70$ ,  $p = .04$  one-tailed. In contrast, the expected difference appeared for low-anchored protocols but was not significant. Amongst well-defined problems, correct solutions manifested greater proportions of incremental rather than insight patterns, and this difference was significant for low-anchored protocols,  $Z = 3.02$ ,  $p = .001$  one-tailed, but not for high-anchored protocols,  $Z = .47$ ,  $p = .32$ .

Other comparisons did not yield significant differences. Correct insight protocols were predicted to exhibit more insight patterns than incorrect insight protocols; the reverse was in fact true, but not significantly so either for low-anchored protocols,  $Z = -1.03$ ,  $p = .15$ , or for high-anchored protocols,  $Z = -.76$ ,  $p = .22$ . Incorrect insight protocols were expected to manifest more incremental patterns than correct insight protocols; the expected difference was true for high-anchored protocols but was not significant,  $Z = -.50$ ,  $p = .31$ ; the reverse was true for low-anchored protocols but the difference again was not significant,  $Z = .60$ ,  $p = .27$ .

Comparing well-defined and insight problems, correct solutions in the latter set were predicted to produce more insight patterns than correct well-defined solutions. Proportions of insight patterns were similar between the two sets, as confirmed by non-significant difference tests for both low- and high-anchored protocols. Finally, more incremental patterns were expected for correct well-defined protocols compared to correct insight protocols; the predicted difference was observed but was not significant at either low- or high-anchor levels.



### c. Differential and angular warmth correlations

Finally, warmth ratings were assessed by correlating warmth measures with problem type. Differential warmth was again employed, along with an angular warmth measure devised by Metcalfe and Wiebe (1987). Differential warmth represents increases in warmth irrespective of the number of intervals to solution, whereas angular warmth represents increases in warmth as a function of the number of intervals to solution. Hence, two protocols with the same differential warmth but solved over different interval periods will differ in angular warmth, with the protocol solved over fewer intervals having lower angular warmth and being less incremental than the protocol solved over more intervals. An important distinction between the present calculation of angular warmth and Metcalfe and Wiebe's (1987) calculation is noted in Appendix C, Note 1.

The relationship between warmth pattern and problem type for correctly solved problems was tested, using the pool of 132 correctly solved problems for all protocols with four or more ratings. For both warmth measures, a Goodman-Kruskal gamma correlation was calculated for each participant by cross tabulating the correctly solved insight and well-defined problems with the ascending rank-ordered warmth scores. A gamma score could only be calculated where a participant had at least one correct problem in each problem set and more than one value of the warmth measure (and where the warmth value was not constant); twenty-three participants had valid gamma scores. A positive correlation was expected, indicating that incremental patterns were more common for well-defined than for insight protocols. Across the sample, the mean gamma for differential warmth was .13, and was not significantly different from zero,  $t(22) = .73$ ,  $MSe = .17$ ,  $p = .47$ . Likewise the mean gamma for angular warmth,  $G = .07$ , was not significantly different from zero,  $t(22) = .47$ ,  $MSe = .15$ ,  $p = .62$  one-tailed. In both cases, there seemed to be little difference in warmth patterns between the two problem sets.

To examine whether the gammas were limited by the number of protocols with high-anchored initial warmth ratings, the correlations were re-calculated by excluding from the problem pool the 23 protocols with initial ratings of seven or higher, leaving 109 protocols. However, reanalysis from this subset of protocols produced mean gammas that were of no greater magnitude or significance from those reported above.

In summary, few differences were found between the two problem sets across the analyses of interval means, solution patterns, and differential and angular warmth measures. Solutions of well-defined and insight problems tended to be accompanied by similar monitoring of FOW ratings. Within each problem set, correct solutions produced monitoring responses similar to those produced for incorrect solutions.

Differences between the problem sets were next analysed in the form of post-solution evaluations, in terms of both overall confidence and accuracy of evaluations, for both absolute and relative evaluation measures.

#### **4. Evaluations: Confidence and Evaluative Accuracy**

The evaluations made on completion of each problem set were analyzed by again using the pool of 395 problems that excluded protocols with only one FOW rating. In contrast to the predictions, participants made evaluations based on perceived performance for complete problem sets rather than for individual problems within the sets. Analyses for each evaluation type (absolute and relative) are reported separately.

##### **a. Absolute evaluations**

Participants indicated how many problems in each set he or she believed had been solved correctly. Mean absolute evaluations (see Table 6) indicate that, on average, participations believed they had correctly solved approximately five problems in each set. These evaluations can be compared with the number of actual correct solutions in Table 6.

There was no significant difference in the mean absolute evaluations between the two problem sets,  $t(31) = .75$ ,  $MSe = .33$ ,  $p = .46$ . However, the two evaluations were significantly correlated with each other ( $r = .36$ ,  $p < .05$ ), indicating that participants who evaluated themselves highly on one set tended also to rate themselves highly on the other set.

Within the well-defined set, both the evaluations and the number of correct solutions were significantly correlated to a moderate degree ( $r = .57$ ,  $p = .001$ ) such that participants who solved more problems correctly also provided higher evaluations of their success. However, participants were, on average, over-confident in their evaluations of solution success; they evaluated themselves as having correctly solved 1.5

more problems than had actually been solved. This difference between mean evaluation and mean solution rate was significant,  $t(31) = 5.65$ ,  $MSe = .27$ ,  $p < .001$ .

Table 6. *Mean (M) absolute evaluations and correct solutions for well-defined and insight sets*

Problem Set		M	SD	MSe
Well-defined (n = 32)	Absolute Evaluation	5.13	1.68	.30
	No. Correct Solution	3.63	1.54	.27
Insight (n = 32)	Absolute Evaluation	4.88	1.64	.29
	No. Correct Solution	2.13	1.45	.26

Within the insight set, a somewhat higher degree of over-confidence was manifested; on average, participants believed they had correctly solved 2.75 more problems than had actually been solved. Again, this difference was significant,  $t(31) = 6.68$ ,  $MSe = .41$ ,  $p < .001$ . Evaluations and success rates were not significantly correlated for the insight problem set ( $r = -.13$ ,  $p = .48$ ).

The above-mentioned pattern of over-confidence was examined in more detail by calculating an absolute evaluation accuracy score for each participant in each problem set. Scores were determined by subtracting the participant's success rate from their absolute evaluation. Hence, an accuracy score of 0 represented an accurate evaluation (success rate and evaluation were identical), a negative accuracy score represented an under-confident evaluation (e.g., score of -2 indicated participant solved two problems more than they had believed), and a positive accuracy score represented an over-confident evaluation (e.g., score of 2 indicated participant solved two less problems than they had believed). The accuracy data was recast into three bands of confidence scores: over-confidence, accurate confidence, and under-confidence. Table 7 shows, for well-defined and insight sets respectively, the frequency of participants who made over-confident, accurately confident, and under-confident evaluations.

Within the well-defined set, only 19% of the sample made accurate evaluations, the majority (78%) being over-confident to some degree; most participants were over-confident only by a matter of one or two problems. Within the insight set, only 12.5% participants made accurate evaluations, with 81% demonstrating some degree of over-

confidence. In this case, participants tended to be over-confident either to a small degree (scores of 1 and 2) or to a large degree (scores of 5 or 6). However, a  $\chi^2$  test conducted on these scores revealed no significant differences between the distributions of the two problem sets; levels of over-confidence and evaluative accuracy were similar for both sets.

Table 7. *Frequency of over-confident, accurately confident, and under confident absolute evaluations within well-defined and insight sets*

Problem Set		Absolute Evaluation Confidence			Total
		Under-Confident	Accurately Confident	Over-Confident	
Well-defined	Frequency	1	6	25	32
	%	3.1	18.8	78.1	100
Insight	Frequency	2	4	26	32
	%	6.3	12.5	81.3	100
Combined	Frequency	3	10	51	64
	%	4.7	15.6	79.7	100

### b. Relative Evaluations

Participants indicated how well they believed they had performed on both problem sets compared to ‘most other people’, without knowing how well they or others had performed. Mean relative evaluations reveal that, on average, participants ( $n = 32$ ) believed they had performed ‘about average’ on both problem sets compared to their inferences of others’ performances. Participants appeared to make similar evaluations between the two sets; there was no significant difference between well-defined ( $M = 2.81$ ,  $SD = .82$ ) and insight ( $M = 3.03$ ,  $SD = .82$ ) evaluations,  $t(31) = 1.32$ ,  $MSe = .17$ ,  $p = .198$ . This is also reflected in a near-significant correlation ( $r = .34$ ,  $p = .054$ ) between the two evaluations.

The distribution of relative evaluations within the well-defined set (Figure 4) indicates that most participants believed they were either ‘average’ (41%) or ‘below average’ (34%) compared to inferred performances of others. A similar tendency emerges from the insight evaluations (53% ‘average’; 22% ‘below average’). Data

restrictions (40% cells frequency < 5; average cell count = 6.4; minimum expected count = .5) limited the use of a  $\chi^2$  calculation to test differences in the distributions between the sets. The  $\chi^2$  that was calculated indicated no significant differences in these distributions,  $\chi^2(4) = 3.00, p = .69$ .

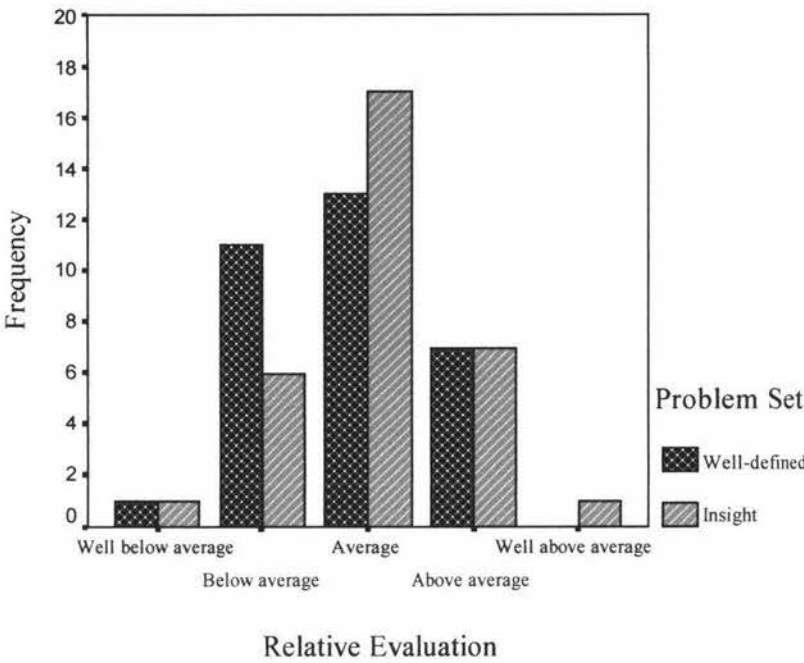


Figure 4. Distributions of relative evaluation ratings in well-defined and insight sets.

To assess the accuracy of the ratings for each set, participants were divided into five performance bands based on the number of problems they had solved correctly. These bands were to be analogous to the five levels of the relative evaluation measure. The modal solution rate for each set was appointed as the ‘average’ band, depending on the frequencies of solution rates within the problem set. Evaluation accuracy was then determined, for each participant, by subtracting his or her accuracy band level (numbered 1 to 5 from very low to very high) from his or her relative evaluation. The resulting evaluation accuracy scores indicated whether the participant had made an under-confident evaluation (negative score), an accurate evaluation (score of 0), or an over-confident evaluation (positive score) of their solution success relative to other participants. The distribution of confidence scores for both problem sets (Table 8) were not significantly different,  $\chi^2(2) = 1.28, p = .53$ .

Table 8. *Frequency of over-confident, accurately confident, and under confident relative evaluations within well-defined and insight sets*

Problem Set		Relative Evaluation Confidence			
		Under-Confident	Accurately Confident	Over-Confident	Total
Well-defined	Frequency	11	9	12	32
	Percent. (%)	34.4	28.1	37.5	100.0
Insight	Frequency	7	10	15	32
	Percent. (%)	21.9	31.3	46.9	100.0
Combined	Frequency	18	19	27	64
	Percent. (%)	28.1	29.7	42.2	100.0

Overall, there were few differences between the problem sets on either the absolute or relative evaluation measures, in terms of confidence or accuracy. However, participants tended to be over-confident in their evaluations, with overconfidence greater for insight than for well-defined problem sets due to lower performance on insight problems; this between-set difference in overconfidence was more pronounced for absolute evaluations than for relative evaluations.

## 5. PSI Analysis: General Versus Specific Metacognitions

Finally, the relationships between general and specific metacognitive beliefs was assessed by comparing participants' responses to the PSI with their predictions and evaluations. In each case, differences between well-defined and insight problem sets were again examined. The PSI was completed by a sub-sample of 25 participants. Lower scores on all scales indicated a more positive appraisal of one's problem solving abilities, with the scales representing a Total PSI score (possible score range 32-192), an Approach-Avoidance Style (AA) score (16-96), a Problem-Solving Confidence (PSC) score (11-66), and a Personal Control (PC) score (5-30).

### a. General problem solving confidence and inter-scale correlations

Prior to the main analysis, PSI performance and inter-scale correlations were examined. Overall means (*M*) and standard deviations (*SD*) for the PSI Total score and

the three subscales are shown in Table 9. Male participants tended to exhibit less confident appraisals than females, overall and on each subscale. T-tests ( $p < .05$ , two-tailed) revealed that males had significantly higher scores than females for PSI total scores,  $t(23) = 2.42$ , AA scores,  $t(23) = 2.08$ , and PC scores,  $t(23) = 2.10$ . Although mean scores for the PSC scale were higher for males than for females, this difference was not significant,  $t(23) = 1.72$ ,  $p = .10$ .

Table 9. *PSI scale means (M) and standard deviations (SD)*

PSI Scale	Total Sample (n = 25)		Male (n = 11)		Female (n = 14)	
	M	SD	M	SD	M	SD
PSI TOT	78.1	18.2	87.6	17.9	70.2	16.7
PSI AA	38.4	11.5	43.3	12.2	34.3	9.5
PSI PSC	23.6	6.8	26.1	6.0	21.6	6.9
PSI PC	16.4	4.5	18.4	3.7	14.9	4.5

Note: PSI TOT = PSI Total score; PSI PSC = PSI Problem Solving Confidence score; PSI PC = PSI Personal Control score; PSI AA = PSI Approach-Avoidance score

All PSI Total and scale scores were significantly correlated with each other (Table 10). These results match, to a reasonable extent, Heppner's (1988) original inter-scale correlations.

Table 10. *Correlations between PSI scale scores and prediction measures*

	PSI PSC	PSI PC	PSI AA	Com. Pred.	WD Pred.	INS Pred.
PSI TOT	.77***	.75***	.91***	-.40*	-.34*	-.36*
PSI PSC	—	.51**	.49**	-.37*	-.36*	-.28
PSI PC	—	—	.54**	-.43*	-.32	-.46*
PSI AA	—	—	—	-.26	-.23	-.25

Note: PSI TOT = PSI Total score; PSI PSC = PSI Problem Solving Confidence score; PSI PC = PSI Personal Control score; PSI AA = PSI Approach-Avoidance score; Com. Pred. = Combined predictions (both problem sets); WD Pred. = Well-defined set predictions; INS Pred. = Insight set predictions.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ; all one-tailed.



### **b. General versus specific metacognitions: Predictions**

The relationship between general and specific metacognitions was examined by correlating the PSI scores (representing general meta-level appraisals) with participants' mean predictions prior to solving problems (representing meta-level appraisals of ability to solve specific problems). Correlations were calculated for all problems combined, for well-defined problems separately, and for insight problems separately (Table 10). Negative correlations meant that *lower* PSI scores, representing greater confidence in one's capabilities, were related to *higher* predictions or greater confidence in solving specific problems.

Mean predictions for problems combined, regardless of actual performance, were significantly correlated in the predicted direction for all but the AA scale ( $p = .09$ ), which did fall in the predicted direction. For well-defined problems alone, correlations were slightly lower in magnitude but mean predictions were significantly related to the PSI Total and PSC scores, but not to the PC ( $p = .06$ ) or AA ( $p = .13$ ) scale scores. For insight problems alone, mean predictions were significantly correlated with the PSI Total and PC scale scores, but not with the PSC ( $p = .09$ ) or AA ( $p = .12$ ) scale scores, though these latter correlations did fall in the predicted direction. Overall results indicated that, at least at a low-to-moderate level, people who tend to be positive in appraising their general problem solving abilities are also more positive in predicting their performance on the types of problems studied in the present research. Also, specific predictions appear to be more related to problem solving confidence and feelings of personal control than to tendencies to approach or avoid problems.

### **c. General versus specific metacognitions: Evaluations**

Relationships between general and specific metacognitive beliefs were also examined by correlating the PSI scores with evaluations. Correlations with the PSI were again hypothesized to be negative in direction, with higher (more confident) evaluations related to lower (more confident) PSI scores, and to be low-to-moderate in magnitude. Note, however, that evaluations were made regardless of actual performance; participants were not notified of their actual success rate prior to making their

evaluations. Table 11 shows the correlations between PSI scores and the absolute and relative evaluations for each problem set.

Overall, general beliefs tended to be more related to the absolute evaluations than to the relative evaluations. Also, evaluations appeared to be related more to feelings of problem solving confidence and personal control than to tendencies to approach or avoid problems. Within the well-defined set, significant correlations of moderate strength were found between absolute evaluations and PSC and PSI Total scores, but not for PC ( $p = .05$ ) or AA ( $p = .12$ ) scales. For the relative evaluations, significant correlations were found with the PSC, PSI Total, and PC scales, but not with the AA scale ( $p = .13$ ). Within the insight set, significant correlations were found between absolute evaluations and all four PSI scores. For relative evaluations, significant correlations of slightly lower magnitude were found with the PSC, PSI Total, and PC scale scores, but not with the AA scale ( $p = .12$ ).

Table 11. *Correlations between PSI scale scores and absolute and relative evaluations*

PSI Scale	Well-defined		Insight	
	Absolute	Relative	Absolute	Relative
PSI TOT	-.41*	-.43*	-.59*	-.39*
PSI PSC	-.52*	-.59*	-.42*	-.43*
PSI PC	-.33	-.39*	-.61*	-.39*
PSI AA	-.24	-.23	-.50*	-.25

Note: PSI TOT = PSI Total score; PSI PSC = PSI Problem Solving Confidence score; PSI PC = PSI Personal Control score; PSI AA = PSI Approach-Avoidance score.  
\*  $p < .05$ .

In summary, moderate correlations were observed between the PSI, as a measure of general off-line metacognitive beliefs, and both predictions and evaluations, as measures of specific on-line metacognitive beliefs. Both predictions and evaluations were related more to appraisals of problem solving confidence and personal control than to approach-avoidance tendencies. Significant relationships were found with both well-defined and insight set measures. Trends between problem sets suggest that on-line appraisals for well-defined problems are slightly more related to problem solving confidence than are appraisals for insight problems, with the latter being somewhat more related to personal control beliefs than are appraisals for well-defined problems.

## **DISCUSSION**

The central purpose of the present study was to assess the nature and accuracy of participants' metacognitive beliefs and experiences during problem solving activities, and to determine the nature of any differences in metacognitive beliefs and experiences between two types of problem: well-defined and insight. These two broad classes of problem have been the focus of debate regarding the nature of human problem solving processes; namely, whether cognitive processing is continuous or discontinuous during the solution of particular problems. Previous research suggests that if there are differences in the nature of solution processes between insight-related and more well-defined problems these differences should become manifest in a person's metacognitive experiences and judgements of performance. Therefore, the present study examined differences of metacognitive beliefs, between well-defined and insight problems, in the form of predictions, monitoring (feelings-of-warmth), evaluations, and the relationships between general, off-line, beliefs and specific, on-line, beliefs.

Several hypotheses were supported and interesting results were found. These results are discussed for each metacognitive activity and hypothesis in turn, in each case noting the interpretive restrictions necessary due to the methodologies employed. First, some initial comments are made regarding the actual problem solving performance of the present sample. Second, results relating to the predictions are discussed. Third, the monitoring and warmth data are interpreted. Fourth, both absolute and relative evaluations are discussed. Fifth, the relationship between general and specific beliefs is discussed. Limitations of the present study are then noted, and suggestions made for future research. In conclusion, implications of the present study are considered in relation to our understandings of insight and well-defined problem solving processes.

### **PROBLEM SOLVING PERFORMANCE**

Overall performance on the problem solving task was quite poor, in terms of both the total problem pool and across individual performance. For all problems combined, less than half the problems in the total pool were solved correctly; the

average individual success rate was about seven correct solutions from a possible sixteen solutions. While one participant achieved a success rate of fourteen correct solutions, most participants could solve only between six and eight problems. Performance was also low within each problem set, for both well-defined and insight problems.

The first hypothesis, that performance (number of correct solutions) would be greater for the well-defined than for the insight set, was confirmed. This supports previous research demonstrating that problems characterized as insight problems tend to be quite difficult (Metcalf, 1986b; Sternberg & Davidson, 1982; Weisberg & Alba, 1981), particularly in comparison to more well-defined problems (Davidson, 1995; Metcalf & Wiebe, 1987). Davidson (1995) found that performance was higher on well-defined than on insight problems for average-ability (IQ) subjects but not for high-ability (IQ) subjects; mean success rates for the latter group were high across all problem types.

Presumably, typical insight problems tend to be more difficult because they encourage an immediately obvious, but inaccurate, representation that leads to the use of unhelpful reasoning or strategies, that in turn leads to the incorrect response (Dominowski, 1995; Schooler & Melcher, 1995). Extended activation of the inaccurate representation may promote fixation on misleading features of the problem, causing the solver to arrive at an impasse; further progress towards a solution ceases (Davidson, Deuser, & Sternberg, 1994; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Knoblich, Ohlsson, & Raney, 2001; Schooler & Melcher, 1995). The correct solution may only be achieved once the solver restructures his or her understanding of the problem to generate a more facilitative representation; however, it may be difficult to break out of a fixated representation in order to achieve a new understanding of the problem.

In contrast, well-defined problems may be less difficult because the elements of the problem- particularly the problem state, the goal state, and the required operations- are clearly specified; difficulty arises not from the generation of an inaccurate representation but from executing the required operations in the correct fashion (Newell & Simon, 1972; Weisberg & Alba, 1981). Arguably, generating a

new and more accurate representation, as required for solving insight problems, may be more difficult than working through the necessary steps of an initially accurate representation.

Alternatively, insight and well-defined problems might instead be solved in similar ways, but with the former systematically more difficult to solve than the latter. That is, insight-like problems may be solved in an incremental fashion, through the gradual use of memory-retrieval and knowledge application based on past experience, just as it is assumed more well-defined problems are solved (Weisberg, 1995a, 1999; Weisberg & Alba, 1981; Weisberg & Alba, 1982). From this perspective, the difficulty of insight-like problems would arise not from fixation on an inaccurate representation of the problem. Rather, an accurate representation may be generated, but may consist of a sufficiently large problem space such that finding the correct solution path of possible operations is not always immediately possible. Finding and executing the correct solution path might be more difficult for traditional insight problems than for traditional well-defined problems because the necessary operations may be less clearly specified or derivable for insight problems. Insight problems might be ‘trickier’, even if solved incrementally, because solvers may follow an incorrect path while gradually accruing some solution-relevant information that only later leads to the correct solution path. Alternatively, solvers may make incremental progress along a correct solution path without acquiring critical knowledge of a crucial step that produces the solution.

Further analysis, discussed below, of the underlying cognitive and metacognitive processes participants engaged in, is needed to clarify which of the above perspectives best explains the relative difficulty of the two problem sets. The former explanation, assuming often sudden experiences of representational change, is the typical interpretation offered by most insight researchers (e.g. Davidson, 1995; Davidson et al, 1994; Dominowski, 1995; Knoblich et al, 1999; Metcalfe, 1986a, 1986b; Metcalfe & Wiebe, 1987). The latter perspective challenges these assumptions, suggesting that people do solve insight problems with more incremental processes (e.g., Weisberg, 1995a; Weisberg & Alba, 1981). Examination of problem solvers’ metacognitions, including their predictions,

monitoring, and evaluation of performance during problem solution, may provide useful information about the processes engaged, whether incremental or sudden, to solve both types of problems.

## **PREDICTIONS**

### **1. Predictive Confidence**

It was hypothesized that there would be no difference in magnitude between predictions for well-defined and insight problems. This hypothesis was confirmed. There was no main effect of problem type on the mean absolute levels of predictions; participant's mean predictions were similar for both insight and well-defined problems. On average, participants were 53 percent and 50 percent confident of correctly solving well-defined and insight problems, respectively. Essentially, this finding demonstrates that participants were making middle-of-the-road predictions for most of the problems, regardless of problem type. It seems that, on average, participants could perceive no information from the problem statements that could inspire more confidence for one type of problem than for the other. It is unclear whether participants were making mid-range 'not sure' predictions because they genuinely believed both problem types were of similar difficulty, or because this was a default response in the face of uncertainty regardless of problem type.

No hypothesis was made regarding a main effect of solution accuracy on the magnitude of predictions. However, a main effect was evident wherein mean predictions were higher prior to correct solutions than prior to incorrect solutions, regardless of problem type. In absolute terms, the difference in mean predictions between correct and incorrect solutions is small: less than one rating point for each problem set. Nevertheless, the main effect suggests that participants may have had some indication of which problems they were more likely to solve correctly or incorrectly. However, little can be confirmed about the veracity of participants' predictions without assessing the accuracy of these predictions.



## **2. Predictive Accuracy**

It was hypothesized that predictions would be more accurate for well-defined than for insight problems. This was tested in two ways: first, by comparing mean probability predictions with the proportion of correct responses in each problem set, to determine the degree of any performance overestimation; and second, by calculating the rank-order correlation between predictions and performance.

### **a. Macro-predictive accuracy: Comparing estimate and performance proportions**

First, it was hypothesized that participants' predictions would overestimate performance in both problem sets, but that overestimation would be greater for insight problems than for well-defined problems. These hypotheses were confirmed. For both well-defined and insight sets the mean prediction estimates were greater than the proportions of correctly solved problems. However, overestimation for insight problem performance was more than twice the overestimation for the well-defined problems. Predictions in the well-defined set overestimated performance by approximately ten percent; predictions in the insight set overestimated performance by twenty-five percent.

The present results support findings from the few studies that have examined metacognitive predictions in problem solving. For example, Metcalfe (1986a) found that predictions of success overestimated actual performance on both recognition and recall memory tasks and on insight problem solving tasks; however, overestimation was greater for the insight than for the memory tasks. Metcalfe and Wiebe (1987) compared predictive accuracy for well-defined (algebra) and insight problems. Again, predictions overestimated actual performance on both types of problems, and the effect of overestimation was greater for insight than for well-defined problems. In the present study, the difference between prediction and performance for insight problems (.25) is similar to the difference found by Metcalfe and Wiebe (1987) for insight problems (.25). However, the difference between prediction and performance for well-defined problems is lower in the present study (.10) than in Metcalfe and Wiebe's results (.18).



The results of the present study differ from Metcalfe and Wiebe's (1987) results also in terms of the basis of the overestimation effect. Metcalfe and Wiebe (1987) found that both predictions and performance were higher for the well-defined than for the insight problems. In the present study, mean predictions were equivalent across problem type while performance was greater for well-defined than for insight problems. The magnitude of predictions for well-defined problems may differ between the present and Metcalfe and Wiebe's studies partly because of the types of problems used. That is, people may be more confident in their abilities to solve algebra problems (as in Metcalfe and Wiebe's study) compared to verbal puzzles (as in the present study), at least with the participant samples employed. This greater confidence for algebra problems may be justified, as performance on well-defined problems was higher for Metcalfe and Wiebe's participants (algebra problems) than for participants in the present study (verbal puzzles). Future research should examine in more detail both predictions and performance for different types of well-defined problems, as the types of well-defined problem selected may have differing implications for comparisons with insight problems.

The results of the present and above-cited research concur with a persistent finding in metacognitive research across a range of cognitive tasks; namely, that metacognitive judgments tend to overestimate performance, but that overestimation is usually greater for more difficult tasks. This finding, that Metcalfe (1998a) terms 'cognitive optimism', has consistently been demonstrated in studies of metacognitive judgements including reading comprehension (e.g., Glenberg, Wilkinson, & Epstein, 1982), memory-based feelings-of-knowing, tip-of-the-tongue experiences, and hindsight bias. Metcalfe (1998a) contrasts the optimism exhibited in these tasks with the type of optimism exhibited by people in stressful or life-threatening situations. In the latter case, people often engage in motivated optimistic self-denial to protect themselves from the negative emotional and psychological consequences of their distressing circumstances. The true nature of these circumstances is usually understood but, to enable coping and to maintain self-esteem, the people affected are motivated to persuade themselves that more optimistic beliefs are true.

In contrast, the optimism exhibited by people engaged in cognitive tasks is more likely a result of accessibility heuristics involved in making performance judgements (Koriat, 1994, 1998; Metcalfe, 1998a). That is, over-prediction appears to result from the nature of the accessible information on which people base their estimates. This includes any apparent problem-relevant information that a person can access from memory or from assessments of task difficulty and personal ability (e.g., Flavell, 1976, 1978, 1987), regardless of the accuracy of this information. Obviously, the problem solver does not realize the inaccuracy of any information he or she is considering; if the information were known to be inaccurate, surely it would not be considered. Also, because predictions are made prospectively and often quickly, prior to detailed analysis or solving of the problem, the solver would have little opportunity to assess the accuracy of accessible information (Koriat, 1998). The more partial, or even inaccurate, information that a person can access upon cuing of a problem or other cognitive task, the higher his or her estimation of success will tend to be. Ironically, however, the same inaccurate information that heightens predictions of performance may also hinder performance by facilitating an incorrect response. Hence, over-prediction will be common because estimates, based partially on information unrecognized as inaccurate, will often be greater than the ensuing performance.

Such over-confident predictions may have a greater impact on insight than on well-defined problem performance because, arguably, insight problems may be more difficult and may promote impressions that an obvious but inaccurate solution is actually correct.

#### **b. Micro-predictive accuracy: Correlating predictions and performance**

It was also considered useful to calculate a micro-predictive index of prediction accuracy; that is, the direct correlation between participants' predictions and performance. It was hypothesized that the correlation between predictions and performance would be greater for well-defined than for insight problem sets; also, the well-defined set correlation, but not the insight set correlation, was expected to be significantly greater than zero. The central hypothesis was not supported by the

present results. While the relationship between predictions and performance was greater for well-defined (mean  $G = .33$ ) than for insight (mean  $G = .19$ ) sets, the difference was not significant. Concurrently, as expected, the insight set correlation was not significantly different from zero but, contrary to expectations, the well-defined set correlation was likewise not significantly different from zero. Therefore, neither correlation was reliably free from chance effects. Despite a (non-significant) tendency for well-defined predictions to be more predictive of actual performance than insight predictions, participants' metacognitive predictions could not reliably predict their performance, either for well-defined or insight problems.

The present results are surprising in the light of previous findings. Metcalfe (1986a) and Metcalfe and Wiebe (1987) used both ranking and probability estimate measures to assess correspondence between predictions and performance. In both studies, and for both measures, the predicted relationships were found. Metcalfe and Wiebe (1987) found that the relationship between prediction and performance was greater for well-defined algebra problems (mean  $G = .40$  on both ranking and probability predictions) than for insight problems (mean  $G = .08$  on ranks; mean  $G = .18$  on probabilities). For both measures, the well-defined set correlations were significantly different from zero while the insight set correlations were not. Metcalfe (1986a), comparing memory tasks and insight problems, found that prediction-performance correlations were greater for the memory tasks than for the insight problems, for both ranked and probability estimate measures; again, correlations for the memory task, but not for the insight problem task, were significantly different from zero.

In terms of insight theory, the lack of a significant prediction-performance relationship for insight problems was expected. Metcalfe (1986a) and Metcalfe and Wiebe (1987) assumed that participants could not reliably predict the likelihood of success on insight problems because of the very nature of these problems. That is, insight problems are assumed to encourage the generation of an initially inaccurate representation or understanding of the problem. If predictions are based on the solver's initial understanding of the problem, and if this initial representation is misleading, then a prediction based on an inaccurate understanding of the problem

can not be expected to be diagnostic of eventual solution accuracy. Confidence does not vary systematically with solution because one's confidence tends to be based on an incomplete understanding of the problem. The solver cannot predict that his or her representation of the problem will need to be restructured to generate a correct solution.

In contrast, the present study's null result for well-defined problem predictions is surprising and difficult to explain. Well-defined problems are assumed to encourage an accurate representation of the problem because, in being well-defined, the elements of the problem situation (problem state, goal state, allowable operations, limiting constraints, etc.) should be clearly specified. There should be little doubt about what the solver is required to do in order to generate the correct solution. Hence, feelings of high or low confidence in one's chances of solving the problem should be based on an adequate understanding of the problem, and hence should be diagnostic of success or failure. One should know when he or she is capable or not capable of solving the problem.

Several possible explanations may account for the lack of predictive accuracy for well-defined problems. First, the problems selected in the present study may not be good examples of well-defined problems; that is, the present problems may not have clearly specified elements at all. However, several of the problems have been identified in previous research as well-defined problems, and have exhibited the qualities expected from well-defined problems. Problems not used in previous research were selected for the present study on the basis that they appeared to manifest the qualities defined as well-defined; that is, the problem elements were considered to be clearly specified, and the problem solvable in an incremental, step-by-step, fashion. These assumptions were supported by pilot testing and confirmed by the feeling-of-warmth data discussed below. Nevertheless, limitations in problem selection are discussed in greater detail later.

A second explanation is that the problems chosen were indeed adequately well-defined but that participants' predictions were not accurate because they responded more randomly to the prediction request. As noted above, predictions

tended toward the mid-range of the prediction scale, as if participants were giving default 'unsure' predictions in the face of uncertain performance.

A third possibility relates to the difference in the methodology used for obtaining predictions between the present and previous studies. The procedure employed by Metcalfe (1986a) and Metcalfe and Wiebe (1987) may have enabled participants to make more accurate predictions compared to those made in the present study, at least for well-defined problems. Metcalfe (1986a) transplanted to the problem solving context a common methodology used in feeling-of-knowing memory research: a dual ranking and probability estimate procedure. Specifically, participants in Metcalfe's (1986a) and Metcalfe and Wiebe's (1987) studies initially ranked problems (from least to most likely to be solved) in terms of appraised likelihood of solution, and then made probability estimates for individual problems. In contrast, the present study used only the probability estimate procedure. Also, Metcalfe had participants make predictions for all the problems prior to any solution activity. In contrast, the present study had participants make predictions sequentially as they solved each problem.

Metcalfe's procedure may have inflated predictive accuracy for several reasons. First, her participants may have had more time to study the problems prior to solution; they had chances during both rating procedures to read the problems. Therefore, participants may have been able to gain a greater understanding of the problems prior to solving them, or may have inadvertently started solving some of the problems prior to the prescribed solution period. Secondly, the act of initially ranking the problems in terms of relative difficulty may have facilitated more accurate probability estimates than would otherwise have been the case. Both Metcalfe (1998a) and Schwartz and Metcalfe (1994) have noted that people are more accurate at rank ordering test items relative to other problems in terms of difficulty, than at providing specific predictions of performance for each item. Therefore, if a person has the opportunity to rank-order a set of problems prior to making absolute predictions, they may subsequently be better able to allocate appropriate absolute estimates for individual problems. If the solver's rank ordering is reasonably accurate, his or her predictive estimates may also be more accurate than if the

estimates were made without the advantage of the initial ranking task. Future research should assess the effect of the ranking procedure on probability estimates by counterbalancing the two procedures.

The rank-ordering procedure was not employed for the present study because it was felt that this procedure might not accurately reflect how predictions of problem solving performance are made in the real world. For a similar reason, predictions were not obtained collectively for all problems prior to any problems being attempted; rather, a prediction was made prior to each problem being solved. It seems unlikely that in most real-life problem solving situations a person will know in advance the complete set of related or unrelated problems, at least enough that he or she might have the opportunity to assess the likelihood of solving each problem by ranking all the known problems in terms of perceived relative difficulty. Rather, it was assumed for the present study that problems are dealt with sequentially, and that one might not predict his or her chances of success until one has identified a problem and is about to solve it. Incorporating these assumptions into the present study may have altered the accuracy of predictions relative to those made in the earlier research. Specifically, predictions for latter problems may have been unduly influenced by perceived performance on the earlier problems rather than by an appreciation of the perceived difficulty of the immediate problem. Under the present conditions, one's predictions, even for well-defined problems, may not be very accurate.

In summary, the accuracy of problem solving predictions may depend to some degree on the method used to observe predictions. If first given the chance to rank-order a set of well-defined problems in terms of relative difficulty, a person's subsequent probability estimates may be reasonably accurate. If, however, this opportunity to rank problems is not available, probability estimates may be less reliably predictive of forthcoming success.

Despite these explanations, the lack of any predictive accuracy in the present study is still surprising and somewhat unsettling. According to most theories of metacognition, predictions serve a crucial purpose in the course of solution attempts; namely, guiding allocations of time, effort, and perseverance and selection of



appropriate strategies (Kluwe, 1982; Metcalfe, 1998a; Paris & Winograd, 1990). At least for well-defined problems, predictions *should* serve these functions reliably and therefore should be diagnostic of success. If predictions are unable to efficiently serve these functions, it is unclear how adequate solution planning can proceed. The comparison between proportions of estimates and performance success, however, illustrates that predictions for well-defined problems are more closely calibrated with performance than are predictions for insight problems. This may indicate that cognitive processing is less continuous in solution of insight than well-defined problems. Consideration of metacognitive monitoring patterns should provide greater appreciation of continuous and discontinuous processes during both insight and well-defined problem solution.

### **MONITORING: FEELINGS OF WARMTH**

Feeling-of-warmth (FOW) ratings are metacognitive appraisals of how close a problem solver is to a solution at various intervals during solution attempts. The warmth procedure was employed in the present study as the primary test that problems had been solved in either an incremental or a subjectively sudden (insightful) fashion. Several hypotheses were proposed regarding the expected patterns for both problem sets and both correct and incorrect solutions. The hypotheses were tested using three statistical assessments of the warmth protocols: comparison of *mean ratings* across pre-solution intervals, comparison of *differential warmth patterns* from first to final pre-solution ratings, and *correlation* of warmth patterns with problem set, for both differential and angular warmth measures. Across all three statistical procedures, few of the hypotheses were supported.

It was hypothesized for insight problems that, if encouraging insightful processing, correct solutions should exhibit more insight than incremental response patterns. This hypothesis was not supported. When differential warmth patterns were compared for solutions with low initial ratings, significantly more correct solutions were accompanied by incremental rather than insight patterns. For protocols with higher initial ratings, no difference in warmth patterns was found. The hypothesis that incorrect insight solutions would exhibit more incremental than insight patterns



was not supported: for protocols with low initial warmth, no difference in warmth patterns was found; for protocols with high initial warmth, more insight than incremental patterns were observed. For correct solutions across problem type, the hypothesis that insight problem protocols were more likely than well-defined problem protocols to exhibit insight warmth patterns was not supported. Likewise, the hypothesis that well-defined protocols were more likely than insight problem protocols to exhibit incremental response patterns was not supported. Neither problem set had predominantly more insight or incremental patterns than the other set. There were no differences in the mean pre-solution warmth ratings for either problem set, no differences in the proportion of insight or incremental patterns of differential warmth, and no significant correlation between warmth pattern and problem set for either differential or angular measures. Overall, it appeared that participants' metacognitive monitoring of solution progress could not discriminate between insight and well-defined problems, or between correct and incorrect solution progress.

The present results suggest that insight problems may be solved incrementally just as often as well-defined problems. At least at a metacognitive level, solution of insight problems may involve more continuous monitoring than has been assumed. This finding would concur with research suggesting that some insight-like problems are not solved following sudden realization of a crucial step or hint; rather, additional processing is required before a correct solution may be produced (Weisberg & Alba, 1981). The present finding also relates to research demonstrating that, at a non-conscious level, cognitive processing of insight-like tasks may gradually progress toward a coherent and correct solution (albeit without the solver's awareness of progress) (Bowers, Farvolden, & Mermigis, 1995; Bowers, Regehr, Balthazard, & Parker, 1990; Durso, Rea, & Dayton, 1994).

The present results starkly contradict the findings from previous studies. Research by Metcalfe (1986b), Metcalfe and Wiebe (1987), Jausovec (1994), and Davidson (1995) have consistently found differences in people's monitoring of solution progress for correctly solved insight and well-defined problems. Specifically, correct solutions to insight problems are accompanied by low pre-

solution warmth until a solution is accompanied by a sudden increase in the final warmth rating. Presumably, solvers are initially unsure of any progress because they have generated an inaccurate representation of the problem that does not facilitate any constructive progress towards a solution. However, restructuring of one's understanding of the problem may allow the solver to overcome uncertainty and impasse, and experience an insight into the problem that rapidly leads to the correct solution (Davidson et al, 1994; Knoblich et al, 1999; Knoblich et al, 2001; Schooler & Melcher, 1995). The typical pattern for correct insight solutions differs from the incremental pattern consistently found for correctly solved well-defined problems (Davidson, 1995; Jausovec, 1994; Metcalfe & Wiebe, 1987). Because the elements of a well-defined problem are reasonably clearly structured, the solver presumably makes gradual progress through a step-by-step procedure towards solution.

Methodological differences between the present and past studies limit the potential conclusions from the present data. First, the present study allowed participants the freedom to choose their first warmth ratings as any point on the 10-point warmth scale. Previous studies have required participants to make their first warmth rating, often prior to initial problem presentation, at the extreme 'cold' end of the scale (Metcalfe, 1986b; Metcalfe & Wiebe, 1987). However, the present researcher felt that this requirement might unduly constrain both the solver's actual initial warmth and the course of warmth progress, where participants had already had a chance to read the problem and predict performance. The finding that initial warmth ratings varied widely across the rating scale supports the assumption that initial confidence is indeed variable across individuals for different problems, and may be higher than the above-cited studies suggest. This finding necessitated the division of warmth protocols into those with low or high initial anchor ratings; consequently, comparisons with previous studies can only be made for protocols with low anchors (first FOW rating of 1 to 3).

A second methodological complication is that, in the present study, the first warmth rating was made 15 seconds into the solution period rather than before, or at the zero-second point of, the solution period. This adjustment was made so that participants would not confuse the first warmth rating with the immediately

preceding prediction rating. Also, Metcalfe and Wiebe (1987; experiment 2) had found that when solvers were not required to provide a minimum initial rating, they still tended to make minimum ratings. So it was assumed for the present study that initial ratings (at 15-second point) would still be low enough to allow a zero-second rating to be extrapolated as the scale minimum (rating of 1) by default. With hindsight, this assumption was not warranted.

This finding does suggest the interesting possibility that if, for the insight problems, participants had been as confident at the 15 second point in Metcalfe's and others' research as were the participants in the present study, then earlier studies may not have found the insight warmth pattern to be as prevalent as had previously been demonstrated. This further implies that, at least in some cases, the obligation to record initial warmth as the minimum scale value may unduly constrain participants' subsequent ratings to the low end of the scale. However, the lack of base-line, zero-second, ratings in the present study makes it difficult to infer accurate patterns of warmth from the present protocols, and to draw reliable comparisons with previous research.

A third methodological difference in the present study was the method for collecting warmth ratings. Keyboard number-key presses were used rather than the previously employed visual line-scale method (with the solver marking slashes on scales printed successively down a page). Use of computer keyboard responses may have inadvertently encouraged less consistent progress along the FOW scale than does the visual scale procedure. A visual scale provides the solver with a permanently accessible record of previous ratings. When unsure of progress, the solver may refer back to previous ratings to make subsequent ratings in line with previous ratings. Given that insight problems may often promote uncertainty in progress, participants in earlier studies may have been more likely to make default progress ratings by repeating previous (usually low) ratings, until the solution was suddenly achieved and higher confidence assured. Such a default response would produce the typically observed insight pattern of warmth. The present keyboard response method, however, does not provide an accessible record of previous responses; if the solver were unsure of progress, and wishes to be guided by

previous ratings, he or she would need to rely on memory of past ratings. Where memory of previous ratings was not readily accessible or was inaccurate, less consistent ratings might inadvertently be recorded. This may be more a function of accessibility to earlier ratings rather than the mode of providing ratings; potentially, access to one's earlier keyboard responses may have the same effect as access to previous line-scale responses. Such methodological effects need to be more clearly assessed before conclusions from the computer-response methodology can be adequately evaluated.

Beyond methodological concerns, other factors may have influenced the warmth patterns observed in the present study. Intelligence and problem solving ability have been demonstrated to influence the monitoring of problem solving. Jausovec (1994) found that warmth patterns were influenced by problem solving ability, as defined by performance on a set of problems. High performers were able to discriminate between problem types in their progress monitoring; the typical differences between well-defined and insight problems were exhibited. Less distinctive differences were found for average performers, and no differences in warmth patterns were observed for low performers. Davidson (1995), comparing performance on the basis of independently-measured IQ scores, found that highly intelligent participants performed better than average intelligent participants on both insight and incremental problems. When metacognitive monitoring was considered, the highly intelligent participants exhibited more distinctively the typical problem type differences than did the average participants; the former participants provided lower overall ratings for the insight problems and higher overall ratings for the well-defined problems than did the averagely intelligent solvers.

In contrast, Metcalfe (1986b) and Metcalfe and Wiebe (1987) did not control for intelligence or ability but still obtained the expected pattern of problem set differences. Arguably, the participants in these latter two studies, being undergraduate university students, were of higher-than-average intelligence within the communities from which they were drawn. Differences between high and low performers within these samples were not discussed. Likewise, the present study did not control for intelligence, and the small sample size and restricted variation in

individual performance limited post-hoc comparisons of high and low performers. Arguably, the present sample also constituted an above-average group compared to the population from which it was selected; most participants had completed some undergraduate university education or attained an undergraduate or postgraduate qualification. Nevertheless, results did not match those observed in the previous studies. If level of intelligence were to account for the present results, similar patterns would be expected for the present as for the previous research. It seems more likely that the rating procedure limited warmth patterns rather than levels of intelligence or ability; alternatively, perhaps the rating procedure masked the potential effects of intelligence.

Problem selection in the present study was not rigorous, and may have limited the expression of insight. That is, the present problems may not have been adequate for encouraging the traditional processes of impasse and sudden solution realization. However, most of the insight problems selected for the present study have been used in previous research obtaining the expected patterns and between-set differences. Nevertheless, for some reason, these problems may have been solved in the current context without participants experiencing sudden insightful metacognitions. It is unclear why this might have been the case to a greater degree in the present study compared to past studies.

A related possibility is that participants did indeed experience insight while solving many of the present problems, but that the rating scale did not adequately capture this experience. For example, problems may have been solved in a hybrid fashion, involving both incremental and insightful processes (e.g., Weisberg, 1995b). Participants may have been able to make some perceived gradual progress toward a solution, but achieving solution only once a critical insight into the problem had been achieved. Such a process would be accompanied by protocols that appeared incremental in nature, even if a subjective insight or ‘Aha!’ experience had occurred. Alternatively, a person may be reasonably sure of the correct solution (with awareness coming relatively suddenly), leading to sub-maximum increases in pre-solution warmth, but persist in their progress for some time in order to verify the answer before terminating solution activity with the final rating and solution

response. Hence, persistence for the sake of solution verification may mask insight warmth patterns even when the problem has been solved with a subjectively sudden process. Interestingly, such a verification process has been assumed in models of insight (e.g., Wallas, 1926, cited in Robertson, 1999; Smith, 1995). Possible limitations in the ability of the warmth-rating procedure to capture insight experiences may mask the observation of this phenomenon. Combined with the present methodological limitations, these latter interpretations, to varying degrees, may be the best (albeit unverified) explanations for the observed results.

## **EVALUATIONS**

As with the predictions, there was no significant difference between problem sets in the average ratings of performance evaluations, for either absolute or relative evaluation measures. For absolute evaluations, participants on average tended to believe they had correctly solved five out of eight problems in each problem set. However, within the total sample, only twelve and four participants actually performed at or above this level, for well-defined and insight sets respectively. With relative evaluations, most participants believed they had performed ‘about average’ compared to other participants, regardless of problem type.

### **1. Evaluation Accuracy**

It was hypothesized that evaluations would be more accurate for the well-defined problem set than for the insight problem set. Support for this hypothesis depended on how accuracy was assessed. Comparison of mean evaluation ratings with the mean number of correct solutions supported the hypothesis; evaluations overestimated performance for both problem types, but overestimation was greater within the insight set than within the well-defined set. However, when distributions of evaluation accuracy were considered, the hypothesis was not supported; there was no significant difference between the problem sets in the distributions of accurate, under-confident, and over-confident evaluations.

Problem solving researchers have not previously studied performance evaluation in the same way as the present study, making theoretical interpretation of



the above results difficult. Evaluations, or post-dictions, have been studied in metamemory research, with results similar to those found in metacognitive research for predictions; judgments tend to exhibit cognitive optimism, with overconfidence manifested for most tasks but greater overconfidence for more difficult tasks (Bunnell, Baken, & Richards-Ward, 1999; Metcalfe, 1998a). In the present study, the comparison between mean evaluations and mean performance supported the general observation of cognitive optimism. In this respect, the evaluation results may be interpreted in the same manner as for the prediction results above. Even after completing the problem sets, participants' metacognitive evaluations of success were higher than was warranted by their actual performance; the opportunity of actually working on the problems, experiencing ease or difficulty in solving the problems, and experiencing perceived success and/or failure did not inhibit the tendency to be overly optimistic about one's overall performance.

Although mean evaluation ratings were similar across problem sets, greater overconfidence was observed for insight than for well-defined problems because performance was considerably lower on the insight problems. It seems that participants could not accurately judge that their insight problem performance was so much lower than their well-defined problem performance. The differing nature of the problems may account for this effect. Because insight problems presumably encourage initially obvious but inaccurate representations, these may facilitate incorrect answers that may otherwise seem correct, or at least feasible. Thus, when incorrect solutions are produced, it may be difficult to know that one's solution is indeed incorrect; the solver assumes that the answer is actually correct because an answer has been produced.

Evaluation is presumably more straightforward for well-defined problems; because the goal state and permissible operations are usually clearly specified, the solver should have a clearer indication that his or her generated solution either does or does not match the stipulated goal state. However, even for well-defined problems, evaluations do seem to exhibit some over-confidence, so there is some lack of correspondence between the actual and appraised matching of generated and stipulated goal states.



The distributions of absolute evaluation confidence support the conclusion that participants tended to be over-optimistic in their evaluations. Less than 20 percent of participants made accurate evaluations in either problem set, with approximately 80 percent being overconfident to some degree. In contrast to the averaged comparisons, however, the distributions of confidence did not differ significantly between problem sets. This lack of a problem type effect may be due to the collapsing of accuracy into the three general bands of accurate-, under-, and over-confidence. Specifically, the over-confident band masks differences between problem sets in the degree of over-confident ratings. If evaluations are compared within the over-confident band alone, differences are still evident between the problem sets. Although most participants were over-confident with their well-defined set evaluations, they tended to be over-confident only by a matter of one or two problems. In contrast, the majority of over-confident evaluations for the insight set were over-confident by five or six problems. When these differences are taken into account, it appears that overconfidence was indeed greater for insight than for well-defined problems.

The relative evaluation measure provides an interesting counterpoint to the absolute evaluation findings. Although absolute evaluations indicated a tendency to overestimate how many problems had been solved, participants tended to believe they had performed either 'average' or 'below average' when comparing own performance with the inferred performance of others. It seems that the need to judge oneself against others, without knowing how well either oneself or others have actually performed, tends to make evaluations more conservative than when one judges one's absolute performance alone. Such social comparison in the face of uncertain performance may encourage people to hedge their bets by evaluating oneself as average, even when they believe their performance is better than average.

However, relative evaluations were not necessarily accurate. When participants were divided into five performance bands in line with the five levels of the evaluation measure, only one-third of participants, in each of the problem sets, made accurately confident evaluations. Frequencies of participants were generally equivalent across confidence bands so that, compared to absolute evaluation

confidence, more people were under-confident and fewer were over-confident in their relative appraisals. These trends were similar for both problem types. The equivalence across confidence bands is probably an artifact of the rating scale and accuracy calculation, given the tendency for participants to rate themselves as ‘average’ performers; obviously, not all participants can actually perform at average levels so, depending on how ‘average’ performance is defined, many will be inaccurate in their assessments even where the range of responses is narrow.

## **2. Utility of Evaluation Measures**

The study of metacognitive evaluations seems an important extension of problem solving research, given the importance of evaluation in theories of metacognition. Previous research examining metacognitive experiences in insight and other problem solving has not included similar evaluation measures. In feeling-of-warmth studies, evaluation has been conflated with monitoring, and particularly with the final warmth rating; a maximum warmth rating is taken as the participant’s evaluation of solution certainty. However, the FOW methodology usually requires participants to give the maximum rating as the final rating, whenever they think they may have a correct solution. Maximum ratings are required even where solvers may still have some doubt about their final solution including, for example, if a satisficing strategy has been used; the final rating may not always be indicative of maximum satisfaction with a solution. Use of a separate evaluation measure, therefore, may provide a better indication of a solver’s confidence in the accuracy of his or her solution.

Nevertheless, the evaluation measures used in the present study may not provide wholly appropriate or accurate indicators of performance appraisal. It is unclear how the present absolute evaluation measure, appraising performance on a complete set of problems, is relevant to real-life appraisal procedures. Few real-life problems come in such neatly-demarcated ‘sets’; people may be more likely to make evaluations on a problem-by-problem basis. Thus, the present absolute evaluation measure may not be an accurate gauge of personal appraisal of post-solution performance.

The relative evaluation measure may have been less reliable or meaningful than the absolute measure. The absolute evaluation required participants to infer only one piece of information: their own perceived performance. The relative evaluation required participants to infer two pieces of information: their own perceived performance *and* the perceived performances of some general ‘others’, to make a comparison between the two performance levels. Without knowledge of one’s own success, relative comparisons may have been unduly impaired. It is probably difficult to infer one’s performance in relation to others when feedback regarding one’s own performance is not available for comparison. The present difficulties with both the absolute and relative evaluation measures limits the conclusions that can be drawn from the present results, and confidence in the above assessments of evaluation accuracy must be reserved.

Despite the noted limitations, use of evaluation measures may enhance future research of metacognitive experiences in problem solving. Future studies could have participants make absolute evaluations for individual problems, in the form of probability estimates that the generated solution is actually correct. This would represent a closer post-solution equivalent to the pre-solution prediction, allowing comparison of both predicted and evaluated performance. Such evaluations would also be useful in conjunction with warmth data, to aid detection of potential satisficing strategies; where a solver is not completely confident of a solution, a less-than-maximum evaluation confidence estimate might be given even where a maximum final warmth rating is required.

More accurate, or at least meaningful, relative evaluations may be exhibited if participants are informed of their own success rate prior to making the relative evaluations. Future research could consider relative evaluation effects in greater detail. Even where one’s actual performance is unknown, the degree of one’s appraised absolute performance may have differing consequences for comparisons with other solvers; self-appraised ‘low’ performers may be more or less confident than self-appraised ‘high’ performers when comparing self with others. Also, the nature of the reference group might influence relative evaluation confidence, perhaps in conjunction with the degree of self-appraisal. In the present study, no

reference was made to a specific group beyond general 'others', but reference to specific groups might alter relative appraisals, perhaps depending on the perceived similarity of oneself to the reference group. Finally, it might be interesting to assess the influence of social group performance information on self-appraised absolute performance, by first giving participants information on how other people tend to perform and then observing how this information influences confidence in self-performance. While the relative evaluation as presently employed is flawed, the measure itself may be useful for incorporating social cognitive understandings into metacognitive research in the domain of problem solving.

## **GENERAL VERSUS SPECIFIC METACOGNITIONS**

The Problem Solving Inventory (PSI) was used to measure general metacognitive beliefs about personal problem solving abilities; that is, beliefs not connected to any immediate problem solving episode. Specific metacognitions were represented by the predictions and evaluations that participants had made during the problem solving tasks. As the comparison between general and specific beliefs was being examined in an exploratory fashion, few predictions were made regarding performance on the PSI and the relationships between these scores and the specific belief measures. However, some interesting results were found.

### **1. PSI-Prediction and PSI-Evaluation Correlations**

The primary purpose of the PSI was to measure correlations between general problem solving appraisal and the specific metacognitive beliefs represented by predictions and evaluations of performance on specific sets of problems. Few hypotheses were proposed regarding possible differences between problem sets or solution performance, due to the lack of any existing theory or research to support any assumptions regarding potential differences. Generally, it was hypothesized that significant correlations, of negative valence and low-to-moderate magnitude, would be observed both between PSI scores and mean predictions and between PSI scores and evaluations. These hypotheses were supported. For both problem sets, several significant correlations, ranging from  $-.34$  to  $-.46$ , were found between predictions

and PSI total and scale scores. Likewise, for both problem sets, significant correlations, ranging from  $-.39$  to  $-.61$ , were found between both absolute and relative evaluations and the PSI scores.

Overall, these results indicate that general metacognitive appraisals of problem solving abilities are related, to some degree, to specific metacognitive beliefs in terms of predictions and evaluations of problem solving performance. Particularly, the more confident one is of one's general problem solving abilities, the more confident predictions and evaluations one will make of performance on the verbal problems studied in the present research. This makes intuitive sense. When a solver is presented with a problem and needs to make a prediction regarding the likelihood of success, he or she has little information on which to base such a prediction. Perhaps the only information available is the nature of the specific problem as presented, specific knowledge one may have about the problem domain, and one's knowledge of how well he or she has dealt with similar problems, or with problems generally, in the past. Context- and problem-specific information will be helpful to some degree, but one cannot be certain of future success. In the face of such uncertainty, the solver may rely on his or her beliefs about problem solving ability based on past experience. If the solver is generally confident of his or her abilities and his or her available repertoire of relevant knowledge and strategies, one can have a general sense of confidence in performance, and thus can make a more confident prediction.

The same may be true of evaluations. Once a person has completed a problem, or a set of problems, a new important piece of diagnostic information is available: actual experience with the problems at hand. However, certainty in success or failure is not assured until one is informed of his or her success. Until performance feedback is secured, the solver may need to base evaluations, at least to some degree, on beliefs of general abilities.

While the above results may make intuitive sense, little research has examined relationships between off-line and on-line metacognitions in problem solving. This is surprising given the value that models of metacognition place on people's global beliefs about abilities and performance. For example, Flavell (1976,

1978, 1987) emphasizes that knowledge of cognition, with regard to person-, task-, and strategy-based knowledge, should guide problem solving activities. Kluwe (1982) distinguishes between general and diagnostic metacognitive knowledge that incorporates both domain-specific and domain-invariant beliefs, as well as knowledge of both cognitive states and strategies to transform cognitive states. Kluwe's (1982) model indicates that there should be some relationship between general metacognitions (for example, broad knowledge of the nature of human thinking, memory, and problem solving processes), and diagnostic knowledge regarding one's beliefs about his or her own particular cognitive processes, strengths, limitations and temporal and situational variations. Also, both Kluwe (1987) and Brown (1978) assume that metacognitive knowledge should interact with metacognitive control and regulation of ongoing problem solving activity, through the processes of identification, prediction, monitoring, and evaluation. Researchers have been slow to empirically demonstrate these proposed links between context-independent and context-specific beliefs. The present study offers some initial verification of such links, but leaves the way open for research to explicate more detailed connections.

## **2. Differences Between PSI Scales**

A second, related, hypothesis in the present study was that correlations for the specific belief measures would be of greater magnitude with the problem solving confidence (PSC) and personal control (PC) scales than with the approach-avoidance style (AA) scale. This hypothesis was confirmed for all measures. Predictions, within each set and in both sets combined, were significantly related to the PSC and PC scales, but not significantly related to the AA scale. Likewise, evaluations in both sets were significantly related to the PSC and PC scales, but only the insight absolute evaluation was significantly related to the AA scale.

These results are interpretable given the content of the respective scales with respect to the present problems. Feelings of confidence and personal control of solution activity are probably relevant to any problem situation or task, including the verbal puzzles used here. However, the tendency to approach or avoid problems is



probably of greater relevance to more affectively-laden problem situations than to the current verbal problems. The present problems are rather innocuous and unlikely to cause anxiety to the extent that participants should feel a need to avoid the problems.

An additional, and unexpected, finding was that different PSI scales were more related to either well-defined or insight problem beliefs. Specifically, well-defined problem measures were more strongly and significantly related to the PSC scale than were insight problem measures. In contrast, the PC scale correlations were stronger and more significant with the insight set measures than with the well-defined set measures. In both cases, these trends were true for predictions and evaluations, and for both absolute and relative evaluations. It is unclear why feelings of problem solving confidence would relate more to specific metacognitions for well-defined problems than for insight problems; likewise, there is no further evidence to explain why personal control beliefs would relate more to insight problem metacognitions than to well-defined problem metacognitions. This differentiation is particularly curious because there was little apparent difference between problem types in any of the specific belief measures. However, these trends suggest that there was some difference in participants' metacognitive ratings between problem sets, and that this difference may somehow be related to different aspects of the participants' general appraisals of performance.

The observed correlational results are also interesting in that they represent relationships between general and specific beliefs without regard to actual performance. That is, regardless of performance, participants tended to demonstrate similar levels of confidence in both their off-line and on-line appraisals. The findings that the on-line appraisals may have been relatively inaccurate in the present study may suggest that general metacognitions are also limited in their diagnostic utility. General metacognitive measures need to be related to actual performance as well as to performance appraisals in order to clarify relations between general beliefs, specific beliefs, and actual performance.

Interpretation of the present results and observed relationships need to be qualified. First, there is only correlational evidence of the suggested relationships

between general and specific metacognitions; no evidence is available to suggest any causal relationships between levels of beliefs. Consequently, the observed correlations may be due to some other, as-yet-unidentified, variable. Second, the present sample completing the PSI is quite small and may be idiosyncratic in nature. Larger samples would be needed to test the reliability of the suggested relationships. Third, the PSI may not be the most accurate tool for measuring general beliefs in the present context. The PSI was selected for its convenience and availability rather than for complete appropriateness to the present research, due to a lack of other tools that measure metacognitive beliefs. This inventory was developed from within a personal and social problem/coping domain (Heppner, 1988) rather than from a hypothetical/non-social puzzle problem-solving domain; there may be important differences in global beliefs between these two domains. However, face validity of the PSI items indicates that most items target metacognitive appraisals that are broad enough to be applicable to any type of problem solving scenario. That respectable correlations were found between the PSI scores and the current specific measures supports the contention that the PSI seems an adequate tool for the present research, at least until a more adequate inventory is designed.

If general beliefs are to have an effect on performance, they must be translated to specific episodes of activity. Context-specific beliefs may represent the mediating interface of such a translation process; the influence of general beliefs on specific beliefs may mediate metacognitive control of problem solving activity. The present research provides observations that warrant closer investigation of these ideas in future research.

## **LIMITATIONS OF THE PRESENT STUDY**

Given the inability of the present study to replicate some consistent findings from previous research, the limitations of the present study should be considered. Methodological weaknesses are noted, followed by relevant conceptual concerns. Non-rigorous selection of appropriate insight and well-defined problems may have hindered the current results, together with lack of control for intelligence and problem solving ability. Finally, the lack of independent verification of participants'

subjective experiences associated with monitoring patterns is considered a limitation of both the present and previous research.

### **1. Methodological limitations**

Methodological weaknesses associated with the prediction, monitoring, evaluation, and general belief measures, as well as the computer-administered mode of task presentation, have already been considered. In summary, the methods used to collect both prediction and monitoring (warmth) responses differed in the present study compared to earlier studies. It is likely that these changes affected participant responses, thus limiting reliable interpretation of the present results. Having participants provide predictions sequentially for individual problems as each is solved, rather than for all problems prior to any solution activity, may have reduced predictive accuracy. For warmth ratings, the lack of a zero-second rating point removed the opportunity for a base-line indication of warmth before, or at onset of, solution activity. This made the ensuing warmth patterns difficult to interpret, particularly in relation to previous research. The present evaluation measures were not ideal, and may have given a misleading picture of evaluation accuracy. The use of complete-set evaluations may not reflect real-world evaluation processes, and lack of self-performance feedback provided participants with little basis for inferring performance evaluations relative to other solvers. While the PSI provided some interesting results as a measure of global problem solving appraisals, it may not have been a wholly appropriate measure for the present research context; a more suitable test may have to be developed for use in related future research.

The use of a computer-administered procedure rather than a paper-and-pencil procedure may have affected the present results in subtle ways, complicating comparisons with previous research. The recording of warmth ratings on the keyboard number-keys, rather than by marking of visual scales printed sequentially down a page, may have subtly affected the course of rating progress. However, access to earlier ratings, regardless of response mode, may be the more pertinent issue.

Undetected effects of intelligence or ability may have limited the capacity of the present study to support previous findings. Ideally, sample selection should be made on the basis of pre-determined intelligence scores or performance on an independent sample of problems; comparisons between participants should be made on the basis of these variables, or at least on the basis of post hoc problem solving performance. Also, problem selection in the present study was not rigorous. Future research should incorporate more rigorous problem selection, including, if necessary, extensive pilot testing to test the assumptions upon which problems are selected.

Additional methodological concerns must be considered. Time to solution may have affected the course of warmth ratings. In the present study, solution time (in the form of number of rating intervals) ranged from four to the maximum fifteen intervals. Warmth patterns may differ for shorter or longer solution periods; if solution periods differ markedly between problem types, comparisons between types may be confounded. This was controlled for some analyses in the present study by limiting analysis to rating protocols within a narrower band of solution intervals. However, this resulted in many protocols (for example, those solved within seven or more intervals) being exempted from analysis. Previous studies have not discussed the effect of solution-interval variance on observed warmth patterns; however, future research should provide more systematic control over solution period comparisons, in ways that allow the maximum possible number of protocols to be considered.

The present feeling-of-warmth procedure required that protocols with less than four ratings not be considered for warmth analyses; meaningful warmth patterns cannot be observed for such protocols. However, the majority of problems for both problem sets in the present study were solved in four or less intervals (within one minute); thus, a majority of protocols were unable to be analyzed for monitoring patterns. Analysis of these protocols would be instructive of monitoring engaged in during more rapid solution activities. The use of shorter (e.g., 10 second) rating intervals might be useful (e.g., Metcalfe, 1986b, Experiment 1), at the risk of greater distraction for participants' concentration on solution efforts. Future research

might consider more appropriate methods of assessing monitoring processes in cases of more, as well as less, rapid solution activity.

Additionally, screening for problems that participants already know or have previously solved is crucial. Metcalfe (1986a, 1986b) and Metcalfe and Wiebe (1987) presented all problems to participants and excluded from further use those with which participants were familiar or had solved very quickly. Although participants in the present study were asked at completion of each set if they had already known any problems, with some problems excluded on this basis, some participants may have denied familiarity when it was actually present. Also, problems solved within one warmth interval were excluded from further analysis. However, the lack of an initial screening process indicates that measures may have been biased because solutions were already known.

## **ii. Conceptual Limitations**

The present discussion highlights two general issues within the insight problem solving research; first, what the term 'insight' refers to and, second, the lack of independent verification of insight occurrences.

### **a. Defining insight and insight problems**

While researchers freely refer to 'insight' problems, this terminology is somewhat misleading. Insight is a particular type of cognitive, or metacognitive, process experienced by problem solvers when solving some problems; insight is *not* an innate property of a particular class of problem. Insight-classified problems may not always encourage an insight experience; any given 'insight' problem may be solved incrementally by an individual just as a more 'well-defined' problem may be solved incrementally. Thus, the mere correct solution of an insight problem does not automatically verify that sudden restructuring or an insight experience has occurred.

Basically, any problem traditionally designated as an insight problem may still be solved correctly in an incremental fashion. The findings of Metcalfe (1986b), Metcalfe and Wiebe (1987), and Davidson (1995) support this contention; in each study, a substantial proportion of warmth protocols for insight problems have

exhibited an incremental pattern (although typically lower than proportions of insight patterns). Similarly, a problem designated as well-defined may not necessarily be solved in a gradual, incremental manner; such a problem could be solved with a more typical insight process, because the insight is a manifestation of a particular type of thinking about a problem rather than a property of the problem itself. Again, the present study, and each of the above-cited studies, found a proportion of well-defined problems had been accompanied by insight patterns of warmth. What determines whether an insight pattern of monitoring is manifested during any solution episode is the interaction of the problem solver's abilities, strategies, and beliefs *with* the nature of the problem, *within* the particular physical, social, and temporal context in which problem solving occurs.

The above reasoning implies that the meaning or focus of the term 'insight' needs to be made more explicit. Most researchers would realize that the term refers to a cognitive process rather than an innate quality of a given problem. However, reference to 'insight' problems, without qualification of what this really means, promotes the misleading implication that such problems should universally be solved with insight experiences. In terms of defining insight and well-defined problems, the above reasoning leads the present researcher to believe that problems should be classified less as an insight/non-insight dichotomy and more as a continuum of problems. At one extreme, problems *tend to promote* insightful cognitive processing for most people most of the time, and at the opposing extreme, other problems *tend to promote* incremental cognitive processing for most people most of the time (cf. Weisberg, 1992). Problems between these two extremes would tend to be solved, to varying degrees or some of the time, with either insightful or incremental processes. Problems towards the mid-range of the continuum would tend to be 'hybrid' problems, solved with both insightful and incremental processes during the one solution episode, or by different processes by different individuals (see Weisberg, 1995b).

Weisberg (1995b) has highlighted a similar need for a more systematic conceptualization of individual problems. His proposed taxonomy categorizes problems on the basis of the types of processes that particular problems encourage;



specifically, whether discontinuity of solution processes and restructuring of problem representations are experienced, and whether restructuring involves pure insight or more hybrid processes. Weisberg (1995b) has used this classification scheme to categorize some common research problems. However, his classifications are based on his individual appraisal of problem content, and include his assumptions about the nature of specific problems. Verification of classification would require other researchers to analyze the same problems using the proposed scheme. Weisberg's (1995b) scheme has the advantage of being theoretically-driven; his categorizations are organized in accordance with traditional Gestalt conceptions of restructuring and insight. However, any systematic classification of problems would require verification through observation of problem solvers' actual processing of selected problems. Specification of how individual problems might be categorized in Weisberg's taxonomy, or where in a continuum of solution processes an individual problem might be placed, would require norming studies that determine frequency prevalence of insight, incremental, and hybrid processes for each problem as experienced by as many problem solvers as feasible. Ideally, norming studies would include all of the problems currently designated in existing problem solving research as insight or well-defined problems. Obviously, the proposed norming research would be a considerable undertaking. However, until such research is conducted, our understanding is limited regarding what research problems are better defined as promoting insight, incremental, or some combination of monitoring processes.

### **b. Verifying insight experiences**

The above classification schemes would improve selection of problems in research. However, independent verification of solution processes for individual solvers in each solution episode is also necessary (Weisberg, 1995b). The present and past research examining metacognition in insight is limited by a lack of independent verification of monitoring protocols. Researchers implicitly assume that particular patterns of warmth ratings correspond with particular subjective experiences of participants. For example, typical 'insight' warmth patterns are

assumed to correspond with the experience of impasses and subjectively sudden insights; incremental patterns are assumed to correspond with gradual increases in one's appraisal of solution proximity.

However, warmth rating tasks may incorporate unidentified response bias, or may lack validity by failing to measure the monitoring behaviour that is intended to be measured, so that warmth patterns do not match with actual subjective experiences. In the present study, the predominance of incremental warmth patterns amongst insight problem solutions does not confirm that these problems were solved without insight. Conversely, the correct solution of insight problems, even for problems identified by future norming studies as usually involving pure insight processes, does not automatically confirm that the problem, in any single case, was actually solved with an insight experience. Without independent verification, these assumptions may be unfounded.

A retrospective verbal report, immediately following solution, would confirm the participant's experience in solving a problem; for example, whether an 'Aha!' reaction, indicative of a sudden insight experience, had occurred. While concurrent verbalization has been shown to inhibit insight problem solving (Schooler & Engstler-Schooler, 1990; Wilson & Schooler, 1991), post-solution retrospective accounts can provide valuable information about the solver's subjective experience of insight problem solving. Although verbalization procedures have been criticized for lack of objective information about cognitive processes (e.g. Nisbett & Wilson, 1977), a subjective report is sufficient if the solver's subjective experience is the object of study (Ericcson & Simon, 1980, 1993; Lieberman, 1979; Nelson & Narens, 1990). Additionally, future research may also incorporate more objective indicators of insight and restructuring experiences; for example, tracking of eye movements across problem materials (Knoblich et al, 2001), speed-accuracy decomposition (Kounios & Smith, 1995; Smith & Kounios, 1996), or neurological measures (e.g., Bowden & Beeman, 1998; Jausovec, 1999). Without independent verification, whether subjective or objective, of the actual processes experienced in an individual solution episode, interpretations of warmth patterns may be misleading regarding the

course of metacognitive monitoring, making comparisons between proposed problem types unreliable.

## **FUTURE RESEARCH**

Most directly, future research related to metacognitive experiences in problem solving should incorporate the recommendations suggested above regarding the methodological weaknesses of the present study. Particular attention should be given to the collection of feeling-of-warmth ratings, and verification of warmth patterns with self-reports of subjective experience.

While design limitations are apparent in the present study, the observed results provide some interesting comparisons between the present and previous research. It is suggested that rating responses may be more dependent on particular procedures than has previously been acknowledged. In the present study, for example, predictions made sequentially for each problem were less accurate, at least for well-defined problems, than were predictions made together following a related ranking procedure, as in Metcalfe's (1986a) and Metcalfe and Wiebe's (1987) studies. The ranking procedure, together with the ability to compare all problems when providing probability estimates, may have inflated predictive accuracy in the earlier studies. Research directly contrasting the relative merits and outcomes of alternative prediction procedures may clarify our understandings of prediction tendencies and accuracy for different problems. Furthermore, there is some indication in the present research that feeling-of-warmth ratings may be 'warmer' earlier in the monitoring process than previously expected. If so, this would have implications particularly for interpretation of patterns, especially with insight problems. While confounded by a lack of zero-second initial ratings in present protocols, the tendency for early warmth ratings to be higher may require further investigation.

The present study used only a select number of problems in both the well-defined and insight sets. Selection was not rigorous, and the present problems may not be representative of other well-defined or insight problems. There are different types of insight and noninsight problems that deserve attention. For example, the

present sample included only relatively short verbal-puzzle problems. There are more problems of this nature that could be studied. Metacognitive experiences may differ depending on the surface modality of problems, whether verbal or visual. Other problems are more visual in nature, within both insight (e.g., nine-dots problem, triangle of coins problem, matchstick problems) and well-defined (e.g., Tower of Hanoi, Missionaries and Cannibals problem) problem types.

Problems and experiences within each type might also differ in terms of underlying features or processes. For example, well-defined problems are distinguishable in terms of transformation, induction, and arrangement processes (Greeno, 1980). Insight problems may be distinguishable in terms of the selective encoding, comparison, and combination processes that are encouraged (Davidson, 1995), or whether pure insight or hybrid insight-incremental processes are evoked (Weisberg, 1995b). Past research has also made use of tasks, such as perceptual identification (e.g., Bowers et al 1995; Bowers et al, 1990; Carroll, 1993) and anagram tasks (e.g., Kounios and Smith, 1995; Metcalfe, 1986b; Smith and Kounios, 1996), that have accompanying phenomenology resembling insight but that may differ in important respects to traditional insight problems (Weisberg, 1995b). Further work in this area should consider more fully the range of problems to be solved.

Beyond well-defined and insight problems, metacognitive experiences of other problem types might be considered. For example, ill-defined problems are recognised as important in the problem solving literature. Perhaps most problems faced in real-life contexts, including personal and social problems, are ill-defined in nature; that is, lacking clearly defined goals or solution paths (Kahney, 1993; Kitchener, 1983; Kitchener & King, 1990). It would seem important to expand the horizons of the current paradigm to include more ecologically relevant problem situations. Jausovec (1994) examined ill-defined problems in conjunction with well-defined and insight problems; however, only a few problems of each type were considered, and his results deserve replication and extension.

Beyond problem solving, problem finding has been identified as equally, or perhaps more, important to adapting to and moving beyond one's current

circumstances (Arlin, 1989; Lubart & Sternberg, 1995; Perkins, 1981). Also, problem finding is often characterized as an insightful process, with a person able to envisage existing circumstances in novel and creative ways (Dominowski, 1995). Problem finding is difficult to examine in the present research paradigm, as there is little for participants to find when they are provided with the problems required for solution. It would be interesting to shift the focus of metacognitive research from problem solving to problem finding, particularly in terms of the experiences involved in finding new modes of conceiving one's environment.

Epistemic beliefs and cognitions are generally considered relevant to both problem finding (Merriam & Caffarella, 1991) and ill-defined problem solving (Jausovec, 1994; Kitchener & King, 1990; Koplowitz, 1987; Schraw, Dunkle & Bendixen, 1995). Kitchener (1983) affirms that, at least in the above contexts, epistemic reasoning must be considered in conjunction with both cognitive and metacognitive processes. However, the empirical links between these three levels of thought are less clear. Consideration should be given to the relationships between epistemic and metacognitive processes; for example, in identifying the characteristics of subjective metacognitive experiences during epistemologically-relevant problem reflection.

Developmental aspects of metacognitive experiences during problem solving have received little attention. This is unfortunate given the array of developmental research in relation to metacognitive knowledge and executive control in problem solving (Carr & Jessup, 1997; Kuhn, Garcia-Mila, Zohar, & Andersen, 1995), metamemory (Bunnell, Baken, & Richards-Ward, 1999; Butterfield, Nelson, & Peck, 1988; Hertzog & Dixon, 1994; Schneider, 1998), and comprehension (Brown, Armbruster, & Baker, 1986; Greeno & Riley, 1987; Hacker, 1998). Research in these contexts has identified important age-related considerations, both between older and younger children and between younger and older adults. The present study did not include a developmental dimension. However, a community-based sample of young-to-middle adulthood participants was selected to encourage future researchers to look beyond university undergraduate samples in examining the issues of interest. Adopting a more diverse selection of age-based samples will allow examination of

developmental aspects of metacognitive experiences. It is unclear at present if age-related differences are important to metacognitive experiences in solving well-defined and insight problems. However, both epistemological reasoning and problem finding are related to cognitive development; to the extent that these processes are related to metacognition and problem solving, subjective metacognitive experiences may also undergo age-related changes.

Finally, the present study examined metacognitive beliefs and experiences, such as feelings-of-knowing and feelings-of-warmth, without consideration of other aspects of metacognition, such as knowledge and regulation of problem solving activity. Future research should more extensively attempt to link together these aspects of metacognition, perhaps also with epistemic reasoning. Also, the present investigation assessed metacognitive processes apart from the underlying cognitive processes that are presumably controlled and monitored by higher-order thinking. For example, subjective experiences were considered without relation to the cognitive strategies that participants were using to solve problems. Jausovec (1994) found that cognitive strategy use and flexibility varied with problem type and ability. He found specific links between strategic flexibility and metacognitive factors. It would seem necessary for future work to more deeply appreciate how metacognitive appraisals are related to the concurrent cognitive-level processes and strategies that are being dynamically appraised.

## **SUMMARY AND CONCLUSIONS**

The present study had several related purposes. First, to assess the accuracy of metacognitive experiences during solution of insight and well-defined problems. Second, to compare patterns of metacognitive experiences between the two problem types, in terms of prediction, monitoring, and evaluation. Third, to explore the relationships between global off-line beliefs of personal problem solving ability and specific on-line beliefs during an immediate solution episode. Previous research of subjective problem solving experience indicates that metacognitive processing may differ depending on the type of problem attempted; specifically, insight problems encourage more discontinuous appraisals and less predictable performance whereas



well-defined problems encourage more continuous appraisals and more predictable performance.

Altogether, the present findings were unable to provide resolution to issues regarding the continuous or discontinuous nature of problem solving processes. However, important results were uncovered in relation to previous research, and potentially fruitful steps were made in extending research into further important areas. Methodological differences between the present and earlier studies limited the interpretation of the present results, while also pointing to potentially important differences in the observation of metacognitive experience depending on how those experiences are measured. This study also highlighted more fundamental conceptual issues relating to the definition of insight and insight problems, and to the independent verification of insight experiences.

In attempting to replicate previous research, the present study was equivocal in its findings relating to metacognitive differences between problem types. In line with past research, predictive overconfidence was exhibited for both problem types, with greater overconfidence for insight than for well-defined problems. This finding concurs with the general observation of optimism exhibited across a wide range of cognitive tasks; people tend to overestimate how well they will perform a cognitive task, but demonstrate greater overconfidence on difficult compared to easier tasks.

Observations of participants' metacognitive monitoring, through feeling-of-warmth ratings, have been used as the central test in determining whether a problem has been subjectively experienced as a continuous or discontinuous process. The present results contradict earlier findings that correct insight solutions are accompanied by subjectively sudden 'insight' monitoring patterns compared to the incremental patterns typically observed with incorrect insight solutions and correct well-defined problem solutions. This study found no differences between problem types in the metacognitive warmth ratings of participants. Rather than undermining the theory of insight supported by previous studies, the present results are probably due to methodological differences. However, the present findings do suggest that solvers may be more confident of progress at onset of problem activity, and that the

warmth rating procedure may not as adequately capture sudden insight experiences, than has previously been assumed.

In extending previous research, the present study produced interesting findings regarding solution evaluation and relationships between general and specific performance beliefs. Average absolute evaluations demonstrated a similar pattern of cognitive optimism to the predictions. While evaluations overestimated performance on both problem sets, overestimation was greater for insight than for well-defined problems, because performance was lower on insight problems. However, when distributions of confidence were examined, no effect of problem set was evident. Relative evaluations revealed most participants believed personal performance to be average compared to inferred performances of others, regardless of problem set; again no effect of problem set on distributions of confidence was evident. As presently employed, the evaluation measures lack sophistication but point the way to more useful means of observing metacognitive appraisals of people's solution performance.

Correlations between the PSI and both predictions and evaluations provided intriguing findings concerning relationships between general and specific problem solving beliefs. Low-to-moderate correlations were found between PSI scores and the specific measures, suggesting a definite link between global beliefs of ability and more context-dependent metacognitions in immediate problem situations. While the present sample may be too small for strong conclusions, and the PSI may not be the most appropriate tool for identifying global beliefs, the present findings suggest that the links between general and context-specific beliefs merit more detailed explication.

In conclusion, the present research was unable to clarify many issues regarding the continuous and discontinuous nature of metacognitive experiences in solving insight and well-defined problems. However, it provides several paths, both conceptual and methodological, by which these issues may be further addressed. Typically, solution of insight problems has been observed with sudden, insightful, experiences of subjective monitoring (Davidson, 1995; Jausovec, 1994; Metcalfe, 1986b; Metcalfe & Wiebe, 1987). This is particularly interesting given an alternative

line of research indicating that at a non-conscious cognitive level, insight-like problems may indeed be solved more incrementally (e.g., Bowers et al, 1995; Bowers et al, 1990; Durso et al, 1994). The links between cognitive and metacognitive processes need to be untangled if our understandings of insight are to be advanced.

Where metacognitive research is concerned, progress will only be made if researchers are more rigorous in heeding several concerns: defining insight as a property of psychological processes rather than an innate property of certain problems; specifying which problems are more likely than others to encourage insightful, incremental, or hybrid processes; and by improving the ability of measurement procedures to capture insight experiences by incorporating independent verification of problem solver's experiences. Given the pervasive presence of problems in our lives, and the often mysterious nature of insight, the development of metacognition research in the above ways could provide intriguing answers to fundamental questions about the way people solve their problems.

## REFERENCES

- Arlin, P. K. (1989). Problem solving and problem finding in young artists and young scientists. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.), *Adult development: Vol. 1. Comparisons and applications of developmental models* (pp. 197-216). New York: Praeger.
- Ashman, A. F. & Conway, R. N. F. (1997). *An introduction to cognitive education: Theory and applications*. London: Routledge.
- Berardi-Coletta, B., Dominowski, R. L., Buyer, L. D. & Rellinger, E. R. (1995). Metacognition and problem solving: a process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 205-223.
- Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *Quarterly Journal of Experimental Psychology*, 35A, 39-49.
- Betsinger, A. M., Cross, J. F. & DeFiore, R. M. (1994). Problem solving and metacognition. *Perceptual and Motor Skills*, 78, 1072-1074.
- Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 53-92). Hillsdale, NJ: Lawrence Erlbaum.
- Bowden, E.M. & Beeman, M.J. (1998). Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science*, 9, 435-440.
- Bowers, K. S., Farvolden, P., & Mermigis, L. (1995). Intuitive antecedents of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 27-51). Cambridge, MA: MIT Press.

Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, 22, 72-110.

Brandreth, G. (1987). *The puzzle omnibus*. London: Guild Publishing.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1) (pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum.

Brown, A. L., Armbruster, B. B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Orsanu (Ed), *Reading comprehension: From research to practice* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum.

Bruning, R. H., Schraw, G. J. & Ronning, R. R. (1999). *Cognitive psychology and instruction* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Bunnell, J. K., Baken, D. M., & Richards-Ward, L. A. (1999). The effect of age on metamemory for working memory. *New Zealand Journal of Psychology*, 28, 23-29.

Butterfield, E. C., Nelson, T. O. & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, 24, 654-663.

Carr, M. & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, 89, 318-328.

- Carroll, M. (1993). Feelings of warmth during picture identification. *Australian Journal of Psychology*, 45, 151-154.
- Cassidy, T. (1999). *Stress, cognition and health*. London: Routledge.
- Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.). *The nature of insight* (pp. 125-155). Cambridge, MA: MIT Press.
- Davidson, J. E., Deuser, R. & Sternberg, R. J. (1994). The role of metacognition in problem solving. In Metcalfe, J. & Shimamura, A. P. (Eds.), *Metacognition: Knowing about knowing* (pp. 207-226). Cambridge, MA: MIT Press.
- Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly*, 28, 58-64.
- Davidson, J. E. & Sternberg, R. J. (1998). Smart problem solving: How metacognition helps. In Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.), *Metacognition in educational theory and practice* (pp. 47-68). Mahwah, NJ: Lawrence Erlbaum.
- Davies, S. P. (2000). Move evaluation as a predictor and moderator of success in solutions to well-structured problems. *The Quarterly Journal of Experimental Psychology*, 53A, 1186-1201.
- de Bono, E. (1967). *The use of lateral thinking*. Middlesex, UK: Penguin.
- Dixon, W. A., Heppner, P. P., Burnett, J. W., Anderson, W. P., & Wood, P. K. (1993). Distinguishing among antecedents, concomitants, and consequences of problem-solving appraisal and depressive symptoms. *Journal of Counseling Psychology*, 40, 357-364.



Dominowski, R. L. (1995). Productive problem solving. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 73-95). Cambridge, MA: MIT Press.

Dominowski, R. L. (1998). Verbalization and problem solving. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 25-45). Mahwah, NJ: Lawrence Erlbaum.

Dominowski, R. L. & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 33-62). Cambridge, MA: MIT Press.

Durso, F. T., Rea, C. B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science*, 5, 94-98.

Ellen, P. (1982). Direction, past experience, and hints in creative problem solving: Reply to Weisberg and Alba. *Journal of Experimental Psychology: General*, 111, 316-325.

Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.

Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis* (Rev. ed.). Cambridge, MA: MIT Press.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Hillsdale, NJ: Lawrence Erlbaum.

Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural/process models of complex human behaviour* (pp. 231-245). Alphen aan den Rijn, The Netherlands: Sijthoff & Noordhoff.

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21-29). Hillsdale, NJ: Lawrence Erlbaum.

Flett, G. L. & Johnston, M. E. (1992). Depression, problem-solving ability, and problem-solving appraisals. *Journal of Clinical Psychology*, 48, 749-759.

Gick, M. L. & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 197-228). Cambridge, MA: MIT Press.

Gilhooly, K. J. (1988). *Thinking: Directed, undirected and creative* (2nd ed.). London: Academic Press.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, 10, 597-602.

Gollwitzer, P. M. & Schall, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and Social Psychology Review*, 2, 124-136.

Greeno, J. G. (1978). Natures of problem solving abilities. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 5. Human information processing* (pp. 239-270). Hillsdale, NJ: Lawrence Erlbaum.

Greeno, J. G. (1980). Trends in the theory of knowledge for problem solving. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research* (pp. 9-23). Hillsdale, NJ: Lawrence Erlbaum.

Greeno, J. G. & Riley, M. S. (1987). Processes and development of understanding. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 289-313). Hillsdale, NJ: Lawrence Erlbaum.

Hacker, D. J. (1998). Self-regulated comprehension during normal reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165-191). Mahwah, NJ: Lawrence Erlbaum.

Hanley, G. L. (1995). Teaching critical thinking: Focusing on metacognitive skills and problem solving. *Teaching of Psychology*, 22, 68-72.

Hayes, J. R. (1980). Teaching problem-solving mechanisms. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Lawrence Erlbaum.

Heppner, P. P. (1988). *The Problem Solving Inventory: Manual*. Palo Alto, CA: Consulting Psychologists Press.

Heppner, P. P. & Petersen, C. H. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counseling Psychology*, 29, 66-75.

Hertzog, C. & Dixon, R. A. (1994). Metacognitive development in adulthood and old age. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 227-251). Cambridge, MA: MIT Press.

Jausovec, N. (1994). *Flexible thinking: An explanation for individual differences in ability*. Cresskill, NJ: Hampton Press.

Jausovec, N. (1999). Metacognition. In M. A. Runco & S. R. Pritzer (Eds.), *Encyclopedia of creativity* (Vol. 2) (pp. 201-207). San Diego, CA: Academic Press.

Kahney, H. (1993). *Problem solving: Current Issues* (2nd ed.). Buckingham, UK: Open University Press.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development*, 26, 222-232.

Kitchener, K. S. & King, P. M. (1990). The reflective-judgment model: Transforming assumptions about knowing. In J. Mezirow and Associates (Eds.), *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning* (pp. 159-176). San Francisco: Jossey-Bass.

Kluwe, R. H. (1982). Cognitive knowledge and executive control: Metacognition. In D. R. Griffin (Ed.), *Animal mind – human mind* (pp. 201-224). Berlin: Springer-Verlag.

Kluwe, R. H. (1987). Executive decisions and regulation of problem solving behaviour. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31-64). Hillsdale, NJ: Lawrence Erlbaum.

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, 1534-1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, 29, 1000-1009.

Koplowitz, H. (1987). Post-logical thinking. In D. N. Perkins, J. Lochhead, & J. C. Bishop (Eds.), *Thinking: The second international conference* (pp. 213-232). Hillsdale, NJ: Lawrence Erlbaum.

Koriat, A. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. . In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115-135). Cambridge, MA: MIT Press.      ♡

Koriat, A. (1998). Illusions of knowing: The link between knowledge and metaknowledge. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Social and cognitive dimensions* (pp. 16-34). London: Sage.

Kounios, J. & Smith, R. W. (1995). Speed-accuracy decomposition yields a sudden insight into all-or-none information processing. *Acta Psychologica*, 90, 229-241.

Kuhn, D., Garcia-Mila, M., Zohar, A. & Andersen, C. (1995). *Strategies of knowledge acquisition*. Monographs of the Society for Research in Child Development, 60(4, Serial No. 245).

Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Leonesio, R. J. & Nelson, T. O. (1990). Do different metamemory judgements tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 464-470.

Lieberman, D. A. (1979). Behaviorism and the mind: A (limited) call for a return to introspection. *American Psychologist*, 34, 319-333.

Lories, G., Dardenne, B., & Yzerbet, V. Y. (1998). From social cognition to metacognition. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Social and cognitive dimensions* (pp. 1-15). London: Sage.

Lubart, T. I. & Sternberg, R. J. (1995). An investment approach to creativity. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 271-302). Cambridge, MA: MIT Press.

Lung, C-T. & Dominowski, R. L. (1985). Effects of strategy instruction and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 804-811.

Matlin, M. W. (1998). *Cognition* (4th ed.). Fort Worth: Harcourt Brace.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W.H. Freeman.

Mayer, R. E. (1995). The search for insight: Grappling with Gestalt Psychology's unanswered questions. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight*. Cambridge, MA: MIT Press.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, 26, 49-63.

Mayer, R. E. (1999). Problem Solving. In M. A. Runco & S. R. Pritzer (Eds.), *Encyclopedia of creativity* (Vol. 2). San Diego, CA: Academic Press.

Mayo, V. D. & Tanaka-Matsumi, J. (1996). Think aloud statements and solutions of dysphoric persons on a social problem-solving task. *Cognitive Therapy and Research*, 20, 97-113.

Merriam, S. B. & Caffarella, R. S. (1991). *Learning in adulthood*. San Francisco: Jossey-Bass.

Metcalf, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 288-294.



Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 623-634.

Metcalfe, J. (1998a). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review*, 2, 100-110.

Metcalfe, J. (1998b). Insight and metacognition. In G. Mazzoni & T. O. Nelson (Eds.), *Metacognition and cognitive neuropsychology: Monitoring and control processes* (pp. 181-197). Mahwah, NJ: Lawrence Erlbaum.

Metcalfe, J. & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Metcalfe, J. & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, 15, 238-246.

Mischel, W. (1998). Metacognition at the hyphen of social-cognitive psychology. *Personality and Social Psychology Review*, 2, 84-86.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95, 109-133.

Nelson, T.O. & Narens, L. (1980). A new technique for investigating the feeling of knowing. *Acta Psychologica*, 46, 69-80.

Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, 26, 125-173.

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

- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231-279.
- Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1). London, UK: Harvester-Wheatsheaf.
- Paris, S. G. & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-51). Hillsdale, NJ: Lawrence Erlbaum.
- Patalano, A. L. & Seifert, C. M. (1994). Memory for impasses during problem solving. *Memory & Cognition*, 22, 234-242.
- Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.
- Pereira-Laird, J. A. (1996). *Causal metacognitive-motivational models of reading comprehension in reading disabled and normal achieving readers*. Unpublished doctoral dissertation, Massey University, Palmerston North, New Zealand.
- Robertson, S. I. (1999). *Types of thinking*. London: Routledge.
- Schneider, W. (1998). The development of procedural metamemory in childhood and adolescence. In G. Mazzoni & T. O. Nelson (Eds.), *Metacognition and cognitive neuropsychology: Monitoring and control processes* (pp. 1-21). Mahwah, NJ: Lawrence Erlbaum.
- Schooler, J. W. & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 17, 36-71.

Schooler, J. W. & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97-133). Cambridge, MA: MIT Press.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122, 166-183.

Schraw, G., Dunkle, M. E. & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*, 9, 523-538.

Schwartz, B. L. & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 93 –113). Cambridge, MA: MIT Press.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 65-124). Cambridge, MA: MIT Press.

Shimamura, A. P. (1994). The neuropsychology of metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 253-276). Cambridge, MA: MIT Press.

Shimamura, A. P. & Squire, L. R. (1986). Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesiac patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 452-460.

Short, E. J. & Weissberg-Benchell, J. A. (1989). The triple alliance for learning: Cognition, metacognition, and motivation. In C. B. McCormick, G. E. Miller, & M. Pressley (Eds.), *Cognitive strategy instruction: From basic research to educational applications* (pp. 33-63). New York: Springer-Verlag.

Siegel, S. & Castellan, N. J. (1988). *Nonparametric statistics for the behavioural sciences* (2nd ed.). New York: McGraw-Hill.

Simon, D. A. & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 907-912.

Simon, H. A. (1980). Problem solving and education. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research* (pp. 81-96). Hillsdale, NJ: Lawrence Erlbaum.

Simon, H. A., Newell, A., & Shaw, J. C. (1979). The process of creative thinking. In H. A. Simon (Ed.), *Models of thought*. New Haven, CT: Yale University Press.

Smith, G. F. (1998). *Quality problem solving*. Milwaukee, WI: ASQ Quality Press.

Smith, S. M. (1995). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 229-251). Cambridge, MA: MIT Press.

Smith, R. W. & Kounios, J. (1996). Sudden insight: All-or-none processing revealed by speed-accuracy decomposition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1443-1462.

Sternberg, R. J., & Davidson, J. E. (1982). The mind of the puzzler. *Psychology Today*, 16, 37-44.

Sternberg, R. J., & Davidson, J. E. (1999). Insight. In M. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (Vol. 2) (pp. 57-69). San Diego, CA: Academic Press.

Swartz, R. J. & Perkins, D. N. (1990). *Teaching thinking: Issues and approaches*. Pacific Grove, CA: Midwest Publications.

Tulving, E. (1994). Foreword. In J. Metcalfe & A. P. Shimamura, (Eds.), *Metacognition: Knowing about knowing* (pp. vii-x). Cambridge, MA: MIT Press.

Weinert, F. E. (1987). Introduction and overview: Metacognition and motivation as determinants of effective learning and understanding. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 1-16). Hillsdale, NJ: Lawrence Erlbaum.

Weisberg, R. W. (1986). *Creativity: Genius and other myths*. New York: Freeman.

Weisberg, R. W. (1992). Metacognition and insight during problem solving: Comment on Metcalfe. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 426-431.

Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. New York: Freeman.

Weisberg, R. W. (1995a). Case studies of creative thinking: Reproduction versus restructuring in the real world. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 53-72). Cambridge, MA: MIT Press.

Weisberg, R. W. (1995b). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157-196). Cambridge, MA: MIT Press.

Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 226-250). Cambridge, UK: Cambridge University 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, 169-192.

Weisberg, R. W. & Alba, J. W. (1982). Problem solving is not like perception: More on Gestalt theory. *Journal of Experimental Psychology: General*, 111, 326-330.

Williams, W. M. & Yang, L. T. (1999). Organizational creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 373-391). Cambridge, UK: Cambridge University Press.

Wilson, T. D. & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality & Social Psychology*, 60, 181-192.

Yaniv, I. & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential biases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 187-205.

Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. Gary Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 253-283). Orlando, FL: Academic Press.



**APPENDIX A**  
**PARTICIPATION FORMS**

Recruitment Advertisement

Information Sheet

Consent Form

Payment Form

Thank You Sheet

## **STUDY ON PROBLEM SOLVING**

We are researchers at Massey University studying people's beliefs about their problem solving, and how these relate to actual problem solving performance. We are interested in studying problem solvers of different ability levels and styles.

We are looking for adults aged between 20 and 50 years of age to attend two separate sessions, spaced about one week apart, at Massey University. Times can be arranged to suit you. In each session, you will be asked to solve a number of small problems, and to provide predictions and evaluations of your performance. Participants may be eligible for \$10 compensation to cover time and costs.

If you are interested in volunteering, or would like to know more about the study, please contact Shane Palmer on 350-5799 ext 5874.

Shane Palmer

Dr Julie Bunnell



**Massey University**

Te Kunenga ki Purehuroa

## RESEARCH PROJECT ON PROBLEM SOLVING BELIEFS AND PREDICTIONS

### INFORMATION SHEET

My name is Shane Palmer. I am currently undertaking a research project at Massey University towards completion of my Master of Arts degree in psychology. My supervisors are Dr. Julie Bunnell and Dr. John Podd, both senior lecturers in the School of Psychology. You can contact me at the School of Psychology, phone (06) 350 5799 ext. 5874. Dr. Bunnell can be contacted on (06) 350 5799 ext. 5258, and Dr. Podd can be contacted on (06) 350 5799 ext. 4135.

Thank you for your interest in this study about problem solving, and for taking the time to contact me. This study investigates how much knowledge people have about their problem-solving behaviour, and whether this knowledge differs for two types of verbal word problems. The problems we are using do not require you to provide information, or answer questions, about problems or issues from your own life experience.

You will be asked to complete two sets of problem solving tasks on a computer, one set for each type of problem. There will be eight problems in each set. You will have the opportunity to read each problem before you attempt to solve it. You will be asked to provide predictions of your success; periodic ratings indicating how close you are to solution; and a final evaluation at the end of the problem set. You will be given two practice problems before you start solving either problem set. This will familiarise you with the type of problem you will be working with, and help you feel comfortable about what you are expected to do. You do not need to be proficient with a computer to take part; full instructions and guidance will be provided.

You may also be asked to complete a Problem Solving Inventory. This is a short paper-and-pencil test that assesses an individual's knowledge and attitudes about their problem solving in a general sense.

Because we want to assess your knowledge of solving each type of problem separately, we need to have you work on each type on two different occasions, about one week apart. Each session should take about 40 to 50 minutes to complete. In the first session, you will be introduced to the study and procedure, given the opportunity to ask questions and to sign a Consent Form, and will work on one type of problem. During the second session you may be asked to fill out the Problem Solving Inventory, and you will work on the second type of problem.

If you choose to take part in this experiment, it will be conducted at the School of Psychology, Massey University, Palmerston North. The dates and times will be arranged to suit you. Once you have completed the first session, and have given your agreement to take part in the second session, a second date and time

will be arranged at your convenience. Shane Palmer will conduct all aspects of the experiment.

The information that you provide will remain confidential and anonymous at all times. Your information will be identified only by a code number. Findings from the study will be reported in grouped summary form only, for all participants combined. Any information will be used solely for the purpose of this study, and findings will be reported only in Shane's thesis, in a results summary for participants, and in any resultant academic presentations and/or publications. Your information will be stored securely at all times, and will be accessible only to the researcher and his supervisors. Once the project is completed, the researcher will continue to store securely the raw information provided by participants, in accord with accepted psychological research practice. However, any contact information that you have provided will be destroyed.

If you choose to take part in the study, you have the right to:

- Decline to participate at any point;
- Refuse to answer any particular questions;
- Withdraw from the study at any time;
- Ask questions about the study at any time during participation;
- Provide information on the understanding that your name will not be used unless you give permission to the researcher;
- Be given access to a summary of the findings of the study when it is concluded.

Once you have completed both sessions of the study, you will be eligible for a small compensation of \$10, in appreciation for your time, and to cover any costs that you may have incurred. You have the right to accept or refuse this as you wish.

If you have any questions or concerns, or would like more information about this study, please contact either Shane or his supervisors. **If you would like to take part in the study, please contact Shane at (06) 350 5799 ext. 5874. Shane will return your call, and arrange a convenient date and time for your first session.** If you have already arranged a time with Shane to participate, please note your appointment details as given on the enclosed appointment sheet.

This project has been reviewed and approved by the Massey University Human Ethics Committee, PN Protocol 01/41.

Thank you for your consideration.

Shane Palmer  
M.A. Student  
(06) 350 5799 ext.5874

Dr. Julie Bunnell  
Senior Lecturer  
(06) 350 5799 ext.5258

Dr. John Podd  
Senior Lecturer  
(06) 350 5799 ext.4135

School of Psychology  
Massey University  
Palmerston North  
(06) 350 5799

## PROBLEM SOLVING BELIEFS AND PREDICTIONS

### CONSENT FORM

I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I understand that I have the right to withdraw from the study at any time and to decline to answer any particular questions. By signing this Consent Form I agree to be involved in both study sessions, under the condition that I may choose at any point not to take part in the second session if I so wish. I understand that I may refuse to take part in the second session, or any other aspect of the study, without penalty or recrimination of any kind.

I agree to provide information to the researcher on the understanding that my name will not be used without my permission. The information that I provide will be used only for this research and publications arising from this research project.

I agree to participate in this study under the conditions set out in the Information Sheet.

**Signed:** .....

**Name:** .....

**Date:** .....

**RESEARCH PROJECT ON  
PROBLEM SOLVING BELIEFS AND PREDICTIONS**

**PAYMENT FORM**

I have received \$10 as an honorarium in partial recognition for my time participating in this study.

YES ☐ NO ☐

Signed: .....

Name: .....

Date: .....

I would like to receive summary information of the results of this study.

YES ☐ NO ☐

Contact Address: .....

.....

.....

## RESEARCH PROJECT ON PROBLEM SOLVING BELIEFS AND PREDICTIONS

### THANK YOU FOR YOUR PARTICIPATION

Thank you for taking part in this research project on problem solving. Your time and contribution are appreciated.

The main objectives in this study are to examine how well people can predict, monitor, and evaluate their own problem solving performance. These abilities are indexes of **metacognition**, which is your knowledge and awareness of your thinking and problem solving. Metacognition involves both reflecting on, and controlling, your thinking skills and processes.

Previous research has indicated that when people use greater levels of metacognitive reflection and control in their thinking processes, this tends to improve the efficiency and accuracy of their thinking. For example, when you are faced with a problem, a greater understanding and ability to plan and allocate your problem solving skills should help you to deal more effectively with that problem.

For this study, you attempted to solve two types of problems. The order in which you worked on each type differed between people. One type was **well-defined** problems. These are problems in which the initial problem (what is 'wrong' or what needs to be solved), the desired goal (the 'answer', or what you want to work out or achieve), and the steps you need to take to get from the problem state to the goal state, are all well-specified.

The other type of problem is known as **insight** problems. These are different from well-defined problems because although the initial problem and the desired goal are well-specified, the steps needed to reach the goal are ambiguous or not obvious. Often people mention that they solved one of these problems in a sudden flash of 'insight', or an 'Aha!' experience; the correct answer came to you suddenly. Of course, once you realise the answer, it then seems obvious!

In problem solving, metacognition incorporates several beliefs and processes when a person attempts a specific problem. You engaged in several activities, each involving a different process or belief. First, you made predictions about your performance on each specific problem. These are an index of your specific **personal metacognitive beliefs**, your beliefs about your *own specific* ability to solve an immediate problem (such as, "How well can I solve this problem?" "What strategies should I use for this problem?"). We will look at how closely people's predictions of their performance match their actual performance, and whether this is different for different types of problems.



Second, you made Feeling-of-Warmth (FOW) judgements every 15 seconds as you tried to solve a problem, indicating how close you thought you were to a correct solution. These are an index of **metacognitive monitoring**; that is, your ability to track the progress of your problem solving strategies and attempts. We will look at the patterns of FOW ratings across the time it takes people to solve a problem. We expect that these patterns will be different for well-defined and 'insight'-type problems.

Third, you made a final evaluation at the end of each set of problems, indicating how well you think you did on that type of problem as a whole. This is an index of **metacognitive evaluation**, or your ability to assess the successfulness of your strategies and attempts on a problem. Your final FOW rating also acts as an evaluation for each specific problem. We will look at how closely your evaluations match your actual performance (we expect that this match will be closer than the match between predictions and performance), and at how your evaluations are related to your predictions.

Finally, you may have completed the Problem Solving Inventory (PSI), which has been designed to assess an individual's perceptions of his or her own problem solving behaviours and attitudes. This is an index of your **personal metacognitive beliefs generally**, that is, your beliefs about your *own general* problem solving abilities, strategies, and performance (such as "How good am I at solving problems?" "What skills do I have to solve problems?"). We will assess how your general beliefs relate to both your predictions and your actual performance.

Please note that your performance on the problems in this study is not a measure of your intelligence, or of your problem solving ability in a personal or social sense. Indeed, by their nature the insight problems can be 'tricky' for everyone, at least until they have found the 'insight' to the answer! Also, because we are studying metacognition, we are more interested in how your predictions, evaluations, and FOW ratings relate to your performance than in how many problems you actually solved.

If you are interested in finding out more about the results of this study, please feel free to sign up for a results summary. This will be written and sent out by Shane to interested participants at the completion of his research, around February or March 2002. Unfortunately, for the purposes of the study, we cannot yet give you a copy of the problems or their answers. *If someone you know is also taking part, it is important that they do not find out the problems or the answers beforehand. For this reason, please do not discuss the study with them prior to their participation.* However, a copy of all the problems and answers will be included with the results summary, to relieve your curiosity!

Thanks again for your time and participation in this study. We hope that you found the experience interesting, and maybe even enjoyable! We are very grateful for your contribution to our research project.

Shane Palmer  
M.A. Student

Julie Bunnell  
Senior Lecturer

John Podd  
Senior Lecturer

School of Psychology  
Massey University  
Palmerston North  
(06) 350 5799

This project has been reviewed and approved by the Massey University Human Ethics Committee, PN Protocol 01/41.

## **APPENDIX B**

### **MATERIALS**

Well-defined Problems

Insight Problems

Instructions: Session 1

Instructions: Session 2

Problem Solving Inventory (PSI)

## WELL-DEFINED (WD) PROBLEMS

The well-defined problems to be used in the present study will be drawn from the following or similar problems. Each problem has been used in previous research and/or is widely regarded as an example of a well-defined problem.

### WD PRACTICE 1. EGG TIMER PROBLEM

For his breakfast Mr. Shell likes an egg that has been boiled for exactly two minutes. He has two egg timers: one that runs for three minutes and one that runs for five minutes. How can he use these egg timers to make sure that his egg is cooked just the way he likes it?

**Answer: He starts both timers at the same time. When the first timer goes off after 3 minutes, Mr Shell will know that he has two minutes in which to cook his eggs before the second timer goes off after five minutes.**

(Brandreth, 1987)

### WD PRACTICE 2. BROTHERS' AGE PROBLEM

Algernon is older than Basil but younger than Cyril, who is older than Dinsdale, who is older than Algernon. Who is the oldest and who is the youngest?

**Answer: Cyril is the oldest, Basil is the youngest.**

(Brandreth, 1987).

### WD 1. WATER JUGS PROBLEM

Suppose that you have a 21-cup jug (Jug A), a 127-cup jar (Jug B), and a 3-cup jar (Jug C). Drawing and discarding as much water as you like, you need to measure out exactly 100 cups of water. How can this be done?

**Answer: Fill Jug B to top, pour enough water from Jug B to fill Jug A. Then pour enough water from Jug B to fill Jug C. Discard the water from Jug C and refill Jug C from Jug B. There will now be 100 cups of water left in Jug B.**

(Weiten, 1992).

### WD 2. FROG PROBLEM

A frog is at the bottom of a 30-metre well. Each hour he climbs 3 metres and slips back two metres. How many hours does it take him to get out?

**Answer: 28 hours**

(Brandreth, 1987)

### WD 3. HOWARD'S AGE PROBLEM

When Howard is twice as old as he is now, he will be three times as old as he was three years ago. How old is Howard now?

**Answer: Nine years old**

(Brandreth, 1987)

**WD 4. GRANDFATHER AND GRANDSON**

A grandfather is forty-four years older than his grandson. Five years ago he was five times as old as his grandson.

How old are the grandfather and grandson now?

**Answer: The grandfather is 60; the grandson is 16.**

(Brandreth, 1987)

**WD 5. FOUR CITIES**

There are four cities in Ruritania. Alphaville is due north of Betaville. Gammaville is due east from Alphaville. Deltaville is due south from Gammaville.

Which city is due east from Betaville?

**Answer: Deltaville**

(Brandreth, 1987)

**WD 6. THE HARDEST PUZZLE**

If the puzzle you solved before you solved this one was harder than the puzzle you solved after you solved the puzzle you solved before you solved this one, was the puzzle you solved before you solved this one harder than this one?

**Answer: Yes**

(Restle & Davis, 1962: cited in Metcalfe & Wiebe, 1987).

**WD 7. SPOTTY DOGS**

The total number of spots on two spotty dogs is ninety-six. If one spotty dog has eighteen more spots than the other spotty dog, how many spots are there on each of the spotty dogs?

**Answer: One spotty dog has 39 spots, the other has 57 spots.**

(Brandreth, 1987).

**WD 8. RACING RESULTS**

Five girls took part in a race. Alison finished before Bunty but behind Clare. Debby finished before Emma but behind Bunty. What was the order in which the girls finished the race?

**Answer: Clare was first, followed by Alison, Bunty, Debby, and finally Emma.**

(Brandreth, 1987).

## INSIGHT (INS) PROBLEMS

The 'insight'-type problems to be used in the present study will be drawn from the following or similar problems. Each problem has been used in previous research and/or is widely regarded as a problem often solved through a sudden 'insight' process.

### INS PRACTICE 1. FARMER'S PROBLEM

A farmer is asked how many animals he has. He replies "They're all horses but two, all sheep but two, and all pigs but two". How many animals does the farmer have?

**Answer: Three- one horse, one sheep, one pig.**

(Brandreth, 1987)

### INS PRACTICE 2. TWO COINS PROBLEM

You have two coins, totaling 55 cents. One of the coins is not a 50-cent piece. What are the two coins?

**Answer: A 50-cent coin and a 5-cent coin. Just one of the coins is not a 50-cent piece... the other coin is.**

(Brandreth, 1987)

### INS 1. TOWER PROBLEM

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

**Answer: The prisoner separated the two strands of the rope, tying one end of each strand together. The rope would then be twice its original length, long enough to reach the ground.**

(Restle and Davis, 1962: cited in Metcalfe and Wiebe, 1987).

### INS 2. WATER LILIES PROBLEM

Water lilies double in area every 24 hours. At the beginning of summer there is one water lily on the lake. It takes 60 days for the lake to become completely covered with water lilies. On which day is the lake half covered?

**Answer: The lake is half covered on the 59th day.**

(Sternberg & Davidson, 1982: cited in Metcalfe & Wiebe, 1987).

### INS 3. SOCKS PROBLEM

You want a pair of socks from your room, but it is night-time and the light in your room is not working. If you have black socks and brown socks in your drawer, mixed in a ratio of 4 to 5, how many socks will you have to take out to make sure that you have a pair of socks the same colour?

**Answer: If you take out 3 socks, two are bound to be a pair of the same colour.**

(Sternberg & Davidson, 1982: cited in Metcalfe & Wiebe, 1987).

**INS 4. TREE PROBLEM**

A landscape gardener is given instructions to plant 4 special trees so that each one is exactly the same distance from each of the others. How is he able to do it?

**Answer: Plant the trees in a tetrahedron. 3 trees are planted in an equilateral triangle around the fourth tree, which is either on top of a hill, or placed in a hole in the middle of the triangle.**

(deBono, 1967: cited in Metcalfe & Wiebe, 1987).

**INS 5. CHAIN PROBLEM**

A woman has four pieces of chain. Each piece is made up of 3 links. She wants to join the pieces into a single closed loop of chain. To open a link costs 2 cents and to close a link costs 3 cents. She only has 15 cents. How does she do it?

**Answer: Open all 3 links in one of the sections (cost  $3 \times 2 = 6$ ). Use these 3 links to join the remaining three sections (closing cost is  $3 \times 3 = 9$ ) Total cost =  $6 + 9 = 15$ .**

(deBono, 1967: cited in Metcalfe & Wiebe, 1987).

**INS 6. HORSE PROBLEM**

A man bought a horse for \$60 and sold it for \$70. Then he bought it back for \$80 and sold it for \$90. How much did he make or lose in the horse trading business?

**Answer: He made \$20.**

(deBono, 1967: cited in Metcalfe & Wiebe, 1987).

**INS 7. B.C. PROBLEM**

A stranger approached a museum curator and offered him an ancient bronze coin. The coin had an authentic appearance and was marked with the date 544 B.C. The curator had happily made acquisitions from suspicious sources before, but this time he promptly called the police and had the stranger arrested. Why?

**Answer: The coin could not be authentic. It would be impossible for someone who actually lived in 544 B.C. to know that eventually the calendar would change to mark a date that was 544 years in the future.**

(Metcalfe, 1986b).

**INS 8. JACK AND JILL PROBLEM**

Jack and Jill were born on the same day of the same year and they are the children of the same parents, yet they are not twins. How is that possible?

**Answer: They are two of a set of triplets, quadruplets, etc.**

(Brandreth, 1987)

## **RESEARCH PROJECT ON PROBLEM SOLVING BELIEFS AND PREDICTIONS**

### **INSTRUCTIONS – SESSION 1**

Thank you for volunteering to take part in this study. Please read these instructions carefully. They will help you to understand exactly how the study will be carried out. Take your time reading them so that you understand what is involved.

You will be presented with 8 problems of a particular type to solve. Each problem will be presented to you on a computer screen. You will be given 15 seconds to read the problem and to predict how well you think you can solve it. After 15 seconds, the problem will be removed from the screen and you will be directed to record your prediction on a 10-point scale, by pressing a number on the keyboard from 1 to 10. A rating of 1 means you are not at all confident of solving the problem, a rating of 10 means you are completely confident of solving the problem.

When you have made your prediction, the problem will return to the screen so that you can solve it. From this moment, you will have 4 minutes to solve the problem. Every 15 seconds you will be asked to indicate how close you believe you are to solving the problem, by pressing a number from 1 to 10 on the top-row of the keyboard. These are called 'Warmth' ratings. A rating of 1 is 'Very Cold' (you have no idea yet of the answer), a rating of 10 is 'Very Hot' (you are certain you know the answer). An in-between score represents an intermediate feeling of warmth.

The times to make your ratings will be indicated by a "beep" sound from the computer. Make your ratings as quickly as possible, based on your immediate impression about how close you are to the correct solution at that moment. Remember, you will be making your ratings every 15 seconds until either you solve the problem, or the 4 minute time-limit is reached.

When you are sure that you have solved the problem, pressing the 'F' letter key will indicate that you have finished. Record your final 'Warmth' rating as 10 ('Very Hot') by pressing the '10' key on the keyboard. Then record what your final solution is by writing this in the space provided in your answer booklet.

If the 4 minute limit expires before you solve the problem, the problem will be removed from the screen and you will be directed to record your final 'Warmth' rating (according to how close you thought you were to the solution), as well as your best-possible answer, if you can. If you have absolutely no idea of the answer, please indicate this in your booklet (for example, by writing "don't know").

When you are ready to continue, pressing the space bar will start the next problem.

Once you have completed all 8 problems in the this set, you will be asked to record two final evaluations indicating how well you believe you performed on the



problems as a whole. First, indicate how many of the 8 experimental trial problems you believe you solved correctly, by pressing a number on the keyboard from 0 to 8. Second, indicate how well you believe you performed compared to most other people, by pressing a number from 1 to 5 that corresponds to the one option of the 5 presented that you believe best describes your performance.

**Do not worry about being able to solve the problems correctly- Just try your best to solve them as well as you can.**

**Speed is not important. Trying to find the correct solution is more important. Take your time if you think this will help you to solve the problem.**

Before you begin the 8 test problems, you will have the chance to work on two practice problems. These will help to familiarize you with the type of problems you will be solving, and with the procedure of recording responses.

If you have any questions about the study or this procedure, please ask Shane, the researcher, now. When you are ready to begin the experiment, let Shane know and he will start the computer program for you.

Good luck. I hope that you enjoy your participation in this study.

## **RESEARCH PROJECT ON PROBLEM SOLVING BELIEFS AND PREDICTIONS**

### **INSTRUCTIONS – SESSION 2**

Thank you for your continued participation in this study. Please read these instructions carefully. They will help you to understand exactly how the study will be carried out. Take your time reading them so that you understand what is involved.

You will again be presented with 8 problems of a particular type to solve. However, these problems are of a different type to those you solved in your first session. The procedure is the same as what you did in your first session. Once again, each problem will be presented to you on a computer screen. You will be given 15 seconds to read the problem and to predict how well you think you can solve it. After 15 seconds, the problem will be removed from the screen and you will be directed to record your prediction on a 10-point scale, by pressing a number on the keyboard from 1 to 10. A rating of 1 means you are not at all confident of solving the problem, a rating of 10 means you are completely confident of solving the problem.

When you have made your prediction, the problem will return to the screen so that you can solve it. From this moment, you will have 4 minutes to solve the problem. Every 15 seconds you will be asked to indicate how close you believe you are to solving the problem, by pressing a number from 1 to 10 on the top-row of the keyboard. These are called 'Warmth' ratings. A rating of 1 is 'Very Cold' (you have no idea yet of the answer), a rating of 10 is 'Very Hot' (you are certain you know the answer). An in-between score represents an intermediate feeling of warmth.

The times to make your ratings will be indicated by a "beep" sound from the computer. Make your ratings as quickly as possible, based on your immediate impression about how close you are to the correct solution at that moment. Remember, you will be making your ratings every 15 seconds until either you solve the problem, or the 4 minute time-limit is reached.

When you are sure that you have solved the problem, pressing the 'F' letter key will indicate that you have finished. Record your final 'Warmth' rating as 10 ('Very Hot') by pressing the '10' key on the keyboard. Then record what your final solution is by writing this in the space provided in your answer booklet.

If the 4 minute limit expires before you solve the problem, the problem will be removed from the screen and you will be directed to record your final 'Warmth' rating (according to how close you thought you were to the solution), as well as your best-possible answer, if you can. If you have absolutely no idea of the answer, please indicate this in your booklet (for example, by writing "don't know").

When you are ready to continue, pressing the space bar will start the next problem.

Once you have completed all 8 problems in this set, you will be asked to record two final evaluations indicating how well you believe you performed on the problems as a whole. First, indicate how many of the 8 experimental trial problems you believe you solved correctly, by pressing a number on the keyboard from 0 to 8. Second, indicate how well you believe you performed compared to most other people, by pressing a number from 1 to 5 that corresponds to the one option of the 5 presented that you believe best describes your performance.

**Do not worry about being able to solve the problems correctly- Just try your best to solve them as well as you can.**

**Speed is not important. Trying to find the correct solution is more important. Take your time if you think this will help you to solve the problem.**

Before you begin the 8 test problems, you will again have the chance to work on two practice problems. These will help to familiarize you with the type of problems you will be solving in this session, and will re-familiarize you with the procedure of recording responses.

If you have any questions about the study or this procedure, please ask Shane, the researcher, now. When you are ready to begin the experiment, let Shane know and he will start the computer program for you.

Good luck. I hope that you again enjoy your participation in this study.

# The Problem Solving Inventory

## FORM B

P. Paul Heppner, Ph.D.

Name \_\_\_\_\_ Date \_\_\_\_\_

Sex \_\_\_\_\_ Age \_\_\_\_\_ Grade or class (if you are a student) \_\_\_\_\_

### Directions

People respond to personal problems in different ways. The statements on this inventory deal with how people react to personal difficulties and problems in their day-to-day life. The term "problems" refers to personal problems that everyone experiences at times, such as depression, inability to get along with friends, choosing a vocation, or deciding whether to get a divorce. Please respond to the items as honestly as possible so as to most accurately portray how *you* handle such personal problems. Your responses should reflect what you *actually* do to solve problems, not how you think you *should* solve them. When you read an item, ask yourself: Do I ever behave this way? Please answer every item.

Read each statement and indicate the extent to which you agree or disagree with that statement, using the scale provided. Mark your responses by circling the number to the right of each statement.

	1 Strongly Agree	2 Moderately Agree	3 Slightly Agree	4 Slightly Disagree	5 Moderately Disagree	6 Strongly Disagree
1. When a solution to a problem has failed, I do not examine why it didn't work .....	1	2	3	4	5	6
2. When I am confronted with a complex problem, I don't take the time to develop a strategy for collecting information that will help define the nature of the problem .....	1	2	3	4	5	6
3. When my first efforts to solve a problem fail, I become uneasy about my ability to handle the situation .....	1	2	3	4	5	6
4. After I solve a problem, I do not analyze what went right and what went wrong .....	1	2	3	4	5	6
5. I am usually able to think of creative and effective alternatives to my problems .....	1	2	3	4	5	6
6. After following a course of action to solve a problem, I compare the actual outcome with the one I had anticipated .....	1	2	3	4	5	6
7. When I have a problem, I think of as many possible ways to handle it as I can until I can't come up with any more ideas .....	1	2	3	4	5	6
8. When confronted with a problem, I consistently examine my feelings to find out what is going on in a problem situation .....	1	2	3	4	5	6
9. When confused about a problem, I don't clarify vague ideas or feelings by thinking of them in concrete terms .....	1	2	3	4	5	6
10. I have the ability to solve most problems even though initially no solution is immediately apparent .....	1	2	3	4	5	6
11. Many of the problems I face are too complex for me to solve .....	1	2	3	4	5	6
12. When solving a problem, I make decisions that I am happy with later .....	1	2	3	4	5	6

Read each statement and indicate the extent to which you agree or disagree with that statement, using the scale provided. Mark your responses by circling the number to the right of each statement.

	1 Strongly Agree	2 Moderately Agree	3 Slightly Agree	4 Slightly Disagree	5 Moderately Disagree	6 Strongly Disagree
13. When confronted with a problem, I tend to do the first thing that I can think of to solve it .....	1	2	3	4	5	6
14. Sometimes I do not stop and take time to deal with my problems, but just kind of muddle ahead .....	1	2	3	4	5	6
15. When considering solutions to a problem, I do not take the time to assess the potential success of each alternative .....	1	2	3	4	5	6
16. When confronted with a problem, I stop and think about it before deciding on a next step .....	1	2	3	4	5	6
17. I generally act on the first idea that comes to mind in solving a problem .....	1	2	3	4	5	6
18. When making a decision, I compare alternatives and weigh the consequences of one against the other .....	1	2	3	4	5	6
19. When I make plans to solve a problem, I am almost certain that I can make them work .....	1	2	3	4	5	6
20. I try to predict the result of a particular course of action .....	1	2	3	4	5	6
21. When I try to think of possible solutions to a problem, I do not come up with very many alternatives .....	1	2	3	4	5	6
22. When trying to solve a problem, one strategy I often use is to think of past problems that have been similar .....	1	2	3	4	5	6
23. Given enough time and effort, I believe I can solve most problems that confront me .....	1	2	3	4	5	6
24. When faced with a novel situation, I have confidence that I can handle problems that may arise .....	1	2	3	4	5	6
25. Even though I work on a problem, sometimes I feel like I'm groping or wandering and not getting down to the real issue .....	1	2	3	4	5	6
26. I make snap judgments and later regret them .....	1	2	3	4	5	6
27. I trust my ability to solve new and difficult problems .....	1	2	3	4	5	6
28. I use a systematic method to compare alternatives and make decisions .....	1	2	3	4	5	6
29. When thinking of ways to handle a problem, I seldom combine ideas from various alternatives to arrive at a workable solution .....	1	2	3	4	5	6
30. When faced with a problem, I seldom assess the external forces that may be contributing to the problem .....	1	2	3	4	5	6
31. When confronted with a problem, I usually first survey the situation to determine the relevant information .....	1	2	3	4	5	6
32. There are times when I become so emotionally charged that I can no longer see the alternatives for solving a particular problem .....	1	2	3	4	5	6
33. After making a decision, the actual outcome is usually similar to what I had anticipated .....	1	2	3	4	5	6
34. When confronted with a problem, I am unsure of whether I can handle the situation .....	1	2	3	4	5	6
35. When I become aware of a problem, one of the first things I do is try to find out exactly what the problem is .....	1	2	3	4	5	6

	CON	AA	PC		
Page 1 Subtotal	_____	_____	_____		
Page 2 Subtotal	_____	_____	_____		
Score	_____	_____	_____	_____	Total

**APPENDIX C**  
**ADDITIONAL RESULTS TABLES, FIGURES, AND NOTES**

APPENDIX C

ADDITIONAL RESULTS TABLES, FIGURES, AND NOTES

Table 1. *Frequency of solution for individual well-defined (WD) and insight (INS) problems*

Problem	Solution			
	Correct	Incorrect	Unsolved	Excluded
WD 1	27	5	0	0
WD 2	3	29	0	0
WD 3	7	18	7	0
WD 4	7	17	8	0
WD 5	22	10	0	0
WD 6	16	15	1	0
WD 7	19	10	3	0
WD 8	28	4	0	0
Total	129	108	19	0
INS 1	26	1	3	2
INS 2	19	8	0	5
INS 3	13	19	0	0
INS 4	6	21	3	2
INS 5	4	22	6	0
INS 6	12	20	0	0
INS 7	9	15	0	8
INS 8	4	26	1	1
Total	93	132	13	18
Combined Total	222	240	32	18

Note: WD = well-defined problem; INS = insight problem.  
See Appendix B for list of problems.



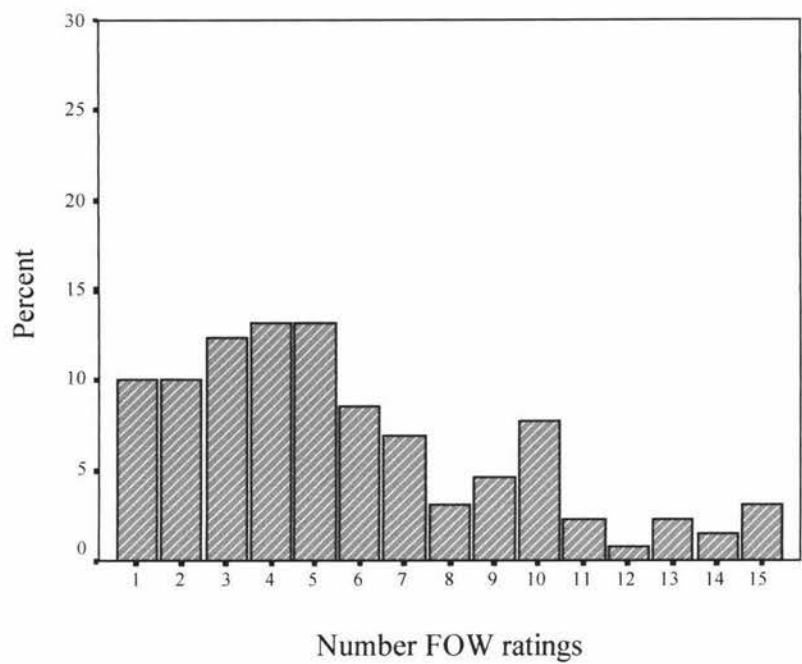


Figure 1(a). Well-defined problem correct solutions

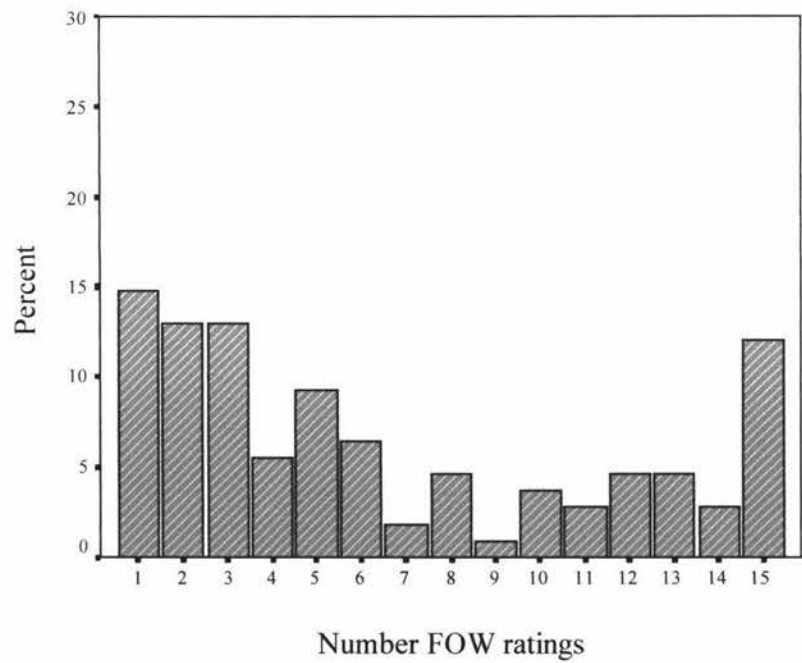


Figure 1(b). Well-defined problem incorrect solutions

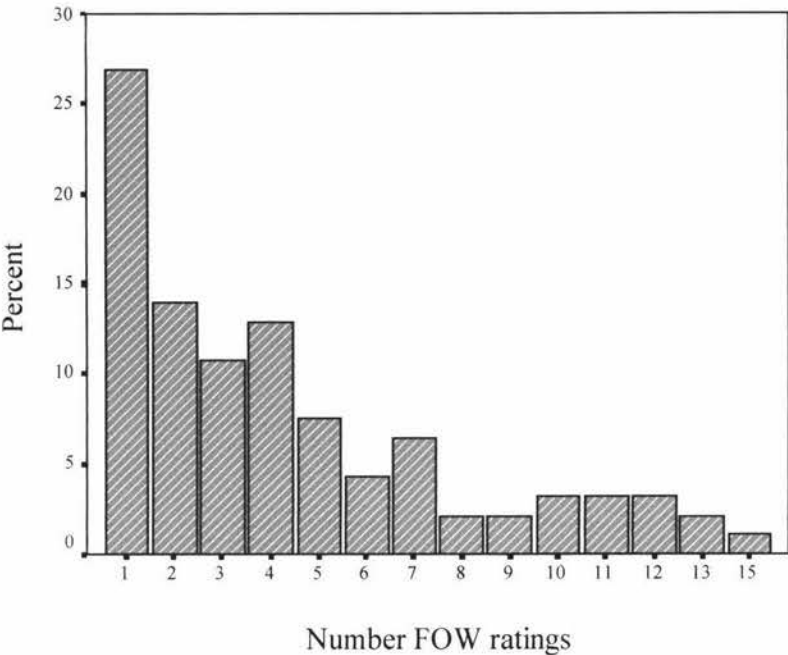


Figure 1(c). Insight problem correct solutions

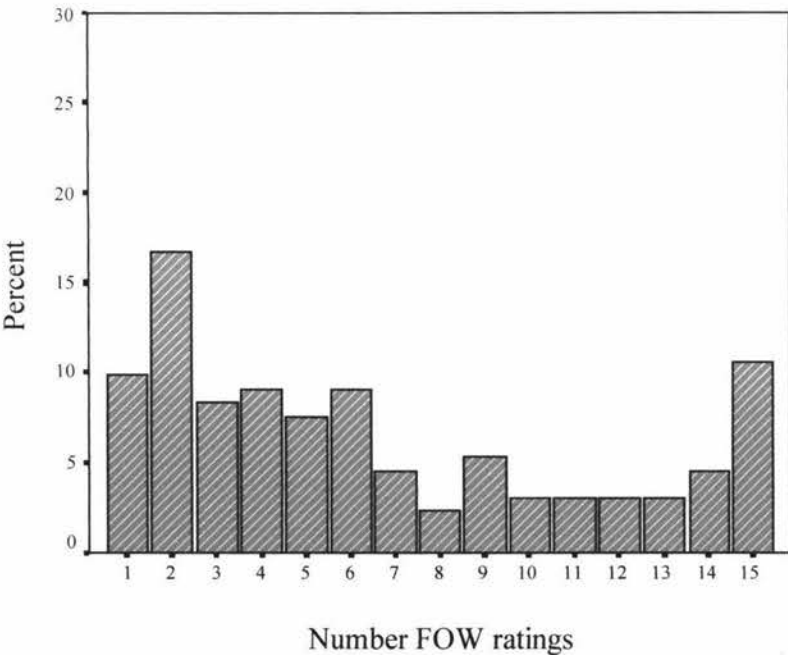


Figure 1(d). Insight problem incorrect solutions

Figures 1(a), (b), (c), (d). Percentage of well-defined correct (a), well-defined incorrect (b), insight correct (c), and insight incorrect (d) solutions solved in each number of FOW intervals.

Table 2. *Frequency of predictions made for correct and incorrect solutions to well-defined and insight problems (all protocols with two or more FOW ratings)*

Prediction	Problem Set	Solution		Total
		Correct	Incorrect	
1	Well-defined	11	9	20
	Insight	22	3	25
	Combined	33	12	45
2	Well-defined	9	5	14
	Insight	5	3	8
	Combined	14	8	22
3	Well-defined	10	8	18
	Insight	6	7	13
	Combined	16	15	31
4	Well-defined	18	19	37
	Insight	25	10	35
	Combined	43	29	72
5	Well-defined	9	16	25
	Insight	23	12	35
	Combined	32	28	60
6	Well-defined	8	13	21
	Insight	9	12	21
	Combined	17	25	42
7	Well-defined	12	17	29
	Insight	9	8	17
	Combined	21	25	46
8	Well-defined	2	7	9
	Insight	10	4	14
	Combined	12	11	23
9	Well-defined	8	11	19
	Insight	4	2	6
	Combined	12	13	25
10	Well-defined	5	11	16
	Insight	6	7	13
	Combined	11	18	29
Total	Well-defined	92	116	208
	Insight	119	68	187
	Combined	211	184	395

Table 3. Mean FOW ratings at each interval for correct and incorrect solutions to well-defined and insight problems (all protocols with four or more FOW ratings)

Problem Set			FOW-14	FOW-13	FOW-12	FOW-11	FOW-10	FOW-9	FOW-8
Well-defined	Correct	Mean	3.60	3.60	3.79	4.14	4.28	4.64	3.99
		N	5	5	7	7	10	16	18
	Incorrect	Mean	3.88	3.10	3.63	3.72	3.78	3.75	4.04
		N	8	10	13	16	17	19	19
Insight	Correct	Mean	2.00	3.00	3.33	4.50	4.31	4.75	27.00
		N	1	1	3	5	8	10	11
	Incorrect	Mean	4.27	4.03	3.75	3.99	3.78	4.02	4.30
		N	11	15	17	17	20	23	25

Problem Set			Solution	FOW-7	FOW-6	FOW-5	FOW-4	FOW-3	FOW-2	FOW-1	Last FOW
Well-defined	Correct	Mean		4.18	4.34	4.40	4.90	5.08	5.71	6.55	9.70
		N		19	23	26	27	29	29	29	29
	Incorrect	Mean		4.11	4.25	4.55	4.61	4.79	5.02	5.80	8.51
		N		21	22	24	27	28	28	28	28
Insight	Correct	Mean		5.27	4.88	4.36	4.24	4.53	5.04	6.18	9.72
		N		13	16	19	20	24	24	24	24
	Incorrect	Mean		4.01	3.88	3.91	4.22	4.48	4.78	5.02	7.57
		N		25	26	28	29	30	30	30	32

Note: Includes means for all participants who had at least one protocol with four or more FOW ratings in each category of problem type and solution. FOW intervals shown in ascending time order; i.e., from earliest (FOW – 14, fourteenth-last) to last (Last FOW) rating in protocols. Top table shows fourteenth-last to eighth-last ratings, bottom table follows with seventh-last to final ratings. N increases with interval as increasingly more participants had protocols with fewer ratings (problem solved faster).

Table 4. *Frequency of first FOW ratings for correct and incorrect solutions to well-defined and insight problems (all protocols with four or more FOW ratings)*

First FOW Rating	Problem Set	Solution		Total
		Correct	Incorrect	
1	Well-defined	13	20	33
	Insight	3	15	18
	Combined	16	35	51
2	Well-defined	6	9	15
	Insight	10	10	20
	Combined	16	19	35
3	Well-defined	8	6	14
	Insight	4	7	11
	Combined	12	13	25
4	Well-defined	15	5	20
	Insight	6	21	27
	Combined	21	26	47
5	Well-defined	18	10	28
	Insight	6	18	24
	Combined	24	28	52
6	Well-defined	9	5	14
	Insight	10	4	14
	Combined	19	9	28
7	Well-defined	9	4	13
	Insight	4	7	11
	Combined	13	11	24
8	Well-defined	7	2	9
	Insight	1	4	5
	Combined	8	6	14
9	Well-defined	1	3	4
	Insight	1	0	1
	Combined	2	3	5
Total	Well-defined	86	64	150
	Insight	45	86	131
	Combined	131	150	281

Table 5. *Frequency of differential warmth values for well-defined problems (all protocols with four or more FOW ratings, and first FOW rating between 1 and 9)*

Solution	Differential Warmth												Total
	-6	-2	-1	0	1	2	3	4	5	6	7	8	
Correct	1	1	2	20	12	14	12	11	4	3	3	3	86
Incorrect	0	0	5	19	8	12	7	6	0	5	2	0	64
Total	1	1	7	39	20	26	19	17	4	8	5	3	150

Table 6. *Frequency of differential warmth values for insight problems (all protocols with four or more FOW ratings, and first FOW rating between 1 and 9)*

Solution	Differential Warmth														Total
	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	
Correct	0	0	0	0	3	10	6	10	6	5	0	4	1	0	45
Incorrect	1	3	3	4	4	25	6	10	11	11	6	1	0	1	86
Total	1	3	3	4	7	35	12	20	17	16	6	5	1	1	131

*Note 1.* Differences in the calculation of angular warmth between Metcalfe and Wiebe's (1987) study and the present study.

In Metcalfe and Wiebe's (1987) visual-analogue rating procedure, rating scales for each interval were arranged as horizontal lines repeated down a page, with the uppermost scale representing the first rating and each successive scale representing a rating in the following interval; participants marked each scale with a slash to indicate their warmth at each interval. Therefore, angular warmth for each protocol represented the angle of a line drawn from the mark on the first scale to the mark on the penultimate scale. Because the present study used computer key presses rather than visual scales to collect warmth data, angular warmth had to be calculated from a trigonometric model based on first and penultimate ratings and the number of rating intervals. For the same protocol, the angle calculated from this method may differ from the angle calculated from a visual-analogue method (the latter depends on the length of page space between successive scale lines as well as the length of space between rating points on each scale line); however, for differing protocols, the present calculation produces angles that differ in relation to each other in the same way that angles calculated with a visual analogue method would differ in relation to each other. For example, regardless of which calculation technique is employed, the larger the angle the more incremental the increase from first to penultimate ratings.

The following formula was used to calculate angles between first and penultimate ratings:

$$\text{Angle (in radians)} = \text{Inverse Tangent} \frac{(\text{penultimate rating} - \text{first rating})}{(\# \text{ rating points between first \& penultimate rating})}$$

$$\text{Angle (in degrees)} = \text{Angle (radians)} \times \frac{180}{\pi}$$

For example, given a protocol with 3 rating points between first and penultimate ratings, a first rating of seven and a penultimate rating of 8:

$$\begin{aligned} \text{Angle (in radians)} &= \text{Inverse Tangent} \frac{(8-7)}{3} = 0.32175 \\ \text{Angle (in degrees)} &= 0.32175 \times \frac{180}{\pi} = 18.43^\circ \end{aligned}$$