Implicit and Explicit Processes in Multiple Cue Judgement

By

Jonathan James Rolison

A thesis submitted to the University of Plymouth in partial fulfilment for the degree of

Doctor of Philosophy

School of Psychology

Faculty of Science

October 2009

Jonathan Rolison

Implicit and Explicit Processes in Multiple Cue Judgement

Abstract

Making judgements often involves integrating multiple pieces of information, or cues, in the environment. While experts, such as physicians, are able to make accurate judgements from multiple cues, they often have poor insight into how they make their inferences. This provides some indication that judgement is influenced by knowledge that is implicit and inaccessible to verbal report. In the present thesis, the cognitive processes involved in multiple cue judgement were explored by training participants on a small number of novel cues using the multiple cue probability learning (MCPL) paradigm. In a training phase, participants predicted a criterion and received outcome feedback in response to each judgement. Learning and judgement in these tasks is often assumed to draw on explicit hypothesis-testing processes. However, a great deal of research suggests that implicit as well as explicit processes can contribute to performance on complex tasks. In eight experiments, several methods were used to examine the role of explicit and implicit processes in multiple cue judgement. While concurrent working memory loads failed to disrupt judgements after learning, we nevertheless found clear evidence that explicit processing is involved in the learning of negative, but not positive cues. Performance on such tasks was correlated with individual differences in working memory capacity, as well as measures of explicit knowledge obtained in the learning process. The results are discussed with respect to dual process theories of learning, judgement, and reasoning. The findings of the present thesis indicate that multiple cue judgement is best viewed within a dual process framework.

Contents

Appendix

List of Tables and Figures

Tables

Figures

Acknowledgments

I would like to thank Jonathan Evans for his continuous support. Over the past three years he has tutored and guided me towards producing what is in this thesis. I am extremely grateful for his supervision. He met regularly with me to discuss ideas and the results of our findings and always valued my opinion. I would also like to thank Jonathan for his help with my writing. He has always encouraged me to write clearly and simply. I am very thankful for his tuition. I would also like to thank Clare Walsh. Especially, Clare provided detailed and constructive criticism of our draft papers and my experimental designs. Ian Dennis was involved in much of the experimental work of this thesis and I am indebted to him for all his help. I would also like to thank Ian for his guidance with statistical analyses. On numerous occasions Ian patiently discussed with me much of the analysis of the studies in this thesis. Finally, I would like to thank my parents for their loving support and encouragement.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

This study was financed with the aid of an internal studentship from the University of Plymouth.

Word count of the thesis (not including references and Appendices):

43,982

Signed..

Date.........2/02/2010.........

Chapter 1

Theoretical Background

Many of the judgements we make in everyday life are based on multiple cues, or pieces of information, in the environment. When putting a car up for sale we may consider factors such as its age, the price of similar cars, and the amount of rust it has, when setting its price. Judgements from multiple cues are also made by experts, such as physicians, stockbrokers, and personnel managers. In these cases the expert must rely on previous experience or training to give the correct weighting in judgement to relevant cues. A physician for instance, may consider a number of symptoms and results of medical examinations when diagnosing a patient. Adding to the difficulty of these kinds of inferences is the probabilistic nature of judgement tasks in real-world uncertain environments.

Multiple cue judgement is of much theoretical interest to cognitive psychology. If an expert is not able verbally to report how they make their judgements this would suggest that they are influenced by knowledge that is not consciously accessible. It is therefore of interest to psychology to understand the cognitive processes that influence judgement from multiple cues. An understanding of the kinds of processes that underlie multiple cue judgement can also be used to inform training programs designed to improve expertise.

A research tradition, known as Social Judgement Theory (SJT), developed as a means for revealing the tacit judgement policies of experts. This approach uses multiple regression methods to uncover the relative weight that experts give to available cues, as well as measuring self-insight into their judgement policies. However, SJT provides a descriptive approach to multiple cue judgement outside of mainstream psychology. SJT research is more concerned with how accurate experts are and their tacit judgement policies than the cognitive processes that underlie their inferences.

The present thesis uses a related paradigm, known as Multiple Cue Probability Learning (MCPL) to explore the cognitive processes involved in multiple cue judgement following training. This approach requires that the individual learns a small number of unfamiliar cues in a novel task environment in order to simulate expertise. Following a training phase, in which outcome feedback is provided in response to each judgement, the individual's learning is assessed in a test phase. A substantial amount of noise is added to the feedback participants receive in MCPL tasks to simulate learning of expertise in real-world environments. This paradigm allows task factors to be manipulated, such as the relation between cues and outcome, in order to explore the cognitive processes involved in multiple cue judgement.

In principle, two types of cognitive processes can contribute to judgement from multiple cues. On the one hand, an individual may acquire explicit verbalisable knowledge as a result of testing hypotheses against the feedback they receive, and use this knowledge in a controlled effortful manner. On the other hand, an individual may acquire implicit knowledge as a result of experiential learning. While they may not be able verbally to report how they made their judgements, implicit knowledge can nevertheless influence the judgements they make. In cognitive psychology some traditions emphasise the role of explicit hypothesis-testing processes, others focus on implicit and associative learning processes, while dual-process theorists suggest that *both* implicit and explicit processes may contribute to learning and subsequent judgements. In the experimental studies of the present thesis we approach multiple cue judgement from a dual process perspective, and apply methods not previous used in the MCPL literature to study the contribution of implicit and explicit processes to judgement from multiple cues.

1.1. SOCIAL JUDGEMENT THEORY

Social judgement theory (SJT) is rooted in Egon Brunswik's ecological approach to psychology (1944, 1955, 1956, see also Hammond & Stewart, 2001). Brunsik's framework, known as *probabilistic functionalism*, describes the environment as inherently uncertain on the one hand (the probabilistic aspect), and cognitive processes as adapted to the structure of the environment on the other (the functional aspect). Making reliably accurate judgements requires both that individuals make use of relevant information (cues) and that the information available reliably predict events in the environment. Brunswik developed the *lens model*, displayed in Figure 1.1, as a means of illustrating these points. In the centre of the model are the available cues. The left side of the model describes the structure of the environment, illustrated by the links between each cue and the criterion (outcome). The right side of the model describes the judgement policy of the individual, shown by the links between the cues and judgement.

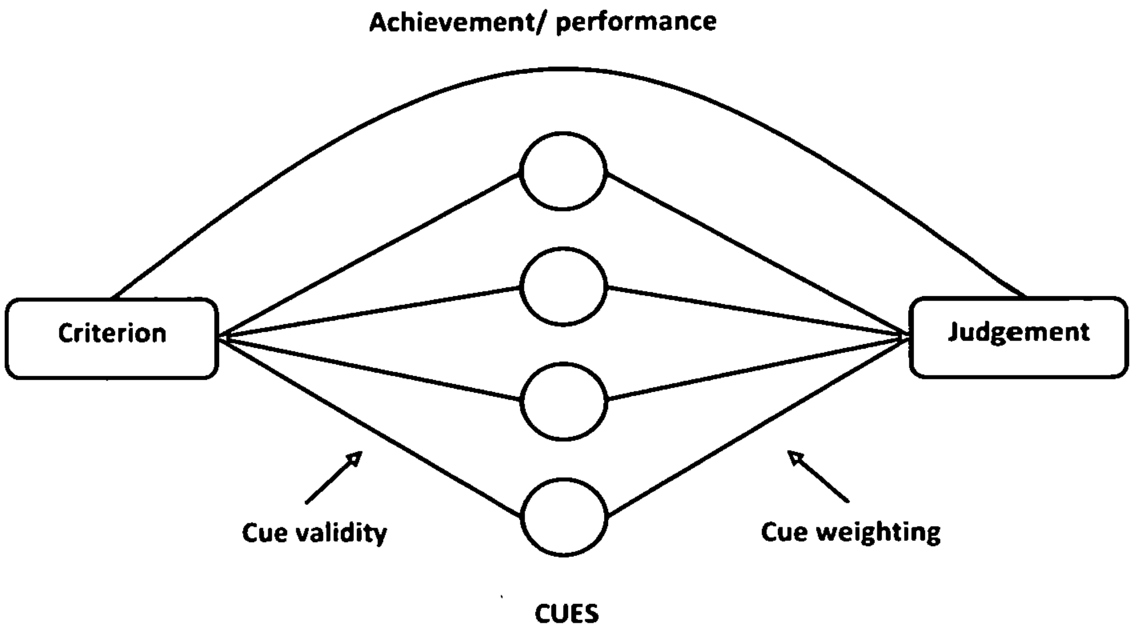

Figure 1.1. The lens model

To illustrate how the lens model works, imagine a physician wishes to diagnose a patient based on the symptoms they display. For the physician to make an accurate diagnosis the available symptoms must together be diagnostic of the patient's condition (left side of Figure 1.1). If the symptoms are not diagnostic, then the physician cannot make an accurate diagnosis regardless of their expertise. If the symptoms are reliable indicators of the patient's condition, however, then the accuracy of the physician's diagnosis depends on their use of the relevant cues (right side of Figure 1.1). The accuracy of the physician's diagnosis, therefore, depends on both the multiple correlation between the cues and the criterion, and the physician's use of the cues.

Kenneth Hammond (1955) was first to apply Brunswik's lens model analysis to the study of clinical judgement. What developed out of Hammond's early work became a methodological approach to judgement analysis, and the beginnings of SJT. Alongside developments of the lens model analysis (Hursch, Hammond, & Hursch, 1964; Tucker, 1964), multiple regression and correlation statistics became popular methods for describing experts' judgement policies and measuring achievement (Cooksey, 1996). In SJT studies, the participant makes a number of judgements based on a set of cues. On each trial, the values of each cue can be generated randomly, while the correlations between cues and the criterion to be judged remain the same throughout the task. *Achievement*, or *performance*, in judging the criterion can be measured by correlating the judgements with the criterion. The cues can then be regressed onto the judgements to provide regression coefficients as an indication of the relative weight that the expert gave to each cue. This process is referred to as *policy capture* because it captures something about the expert's judgement policy. Once the judgement policy of the expert is known, it may also be of interest to know how reliable, or consistent, the expert is in their judgement making. This measure is often referred to as *reliability*, and can be measured by the multiple correlation of the cues with the judgements. Finally, verbal reports provided by the expert about their judgement policy can be compared with their tacit judgement policy as a measure of the degree of insight they have into how they made their judgements. Measuring these factors is referred to as judgement analysis (Wigton, 1996).

Achievement, policy capture, reliability, and self-insight measures describe only the individual side of the lens model. Studies of this type are known as *single-system designs* (Stewart, 1988), and are used to study the judgement policies of experts. However, if criterion values are available the structure of the environment, or task characteristics, can also be described and compared with the expert's model of the environment. As noted earlier, even if an expert has an accurate model of the environment and makes optimal use of the available cues, if the cues are not together diagnostic of the criterion then judgement will be poor. It is therefore of interest also to model the environment. This measure, known as *task predictability*, is the multiple correlation of the cues with the criterion. Finally, in studies where the full lens model is assessed, referred to as *double-system designs*, the model of the environment can be compared with the model of the expert judge. The match between the individual's model and the environment model can be measured by correlating the environment model of the relations between the cues and the criterion with the individual model of the relations between the cues and the judgements. This provides a measure of *task knowledge*, or the 'match' between the environment and the judge (Cooksey, 1996). Recall that achievement in predicting some criterion depends on both sides of the lens model. This can be seen in the equation below (Equation 1.1). Appreciating the full lens model, achievement is a function of reliability, task predictability, and task knowledge. If all three factors are high, then achievement will be good.

Equation 1.1. Achievement = reliability x task predictability x task knowledge.

SJT has been applied to a wide range of judgement domains, including business (Roose & Doherty, 1976; Singh, 1990), medical diagnosis and prognosis (LaDuca, Engel, & Chovan, 1988), weather forecasting (Stewart, 1990), education (Athanasou & Cooksey, 2001), and psychological assessment (Cooper & Werner, 1990). In the vast majority of SJT studies only the individual side of the lens model is considered (Wigton, 1996). In these studies, the tacit judgement policy of the expert is modelled and compared with the policies of other experts to measure variation between experts. Assessing variation between judgement policies is important for exploring whether experts agree on which cues are most important within their judgement domain. However, measuring achievement and task knowledge is also important for assessing expertise. Task characteristics, such as the relation that cues have with the criterion and task predictability are important for comparing achievement across judgement domains.

Studies that make use of criterion values often report high levels of achievement by experts (Brannen, Godfrey, & Goetter, 1989; Goldman et al, 1988; Stewart, Roebber, Bosart, 1997), while others have shown achievement to be disappointingly poor (Faust, 1986). Achievement levels also vary between judgement domains (Kaufmann & Athanasou, 2009), with medical health professionals outperforming those in education, and experts involved in business outperforming both. Differences in reported levels of achievement between studies and between judgement domains are most likely due to task characteristics (Stewart et al, 1997). One important characteristic that puts an upper limit on judgement accuracy is task predictability. When task predictability is low, the cues available to the judge are not sufficient to achieve high levels of judgement accuracy. Accordingly, when task predictability reduces, achievement worsens (Goldberg, 1965; Harvey, 1995). Another important task factor is the number of cues made available to the expert. When the number of available cues is increased, judgement achievement along with reliability reduces (Lee & Yates, 1992; Payne, Bettman, & Johnson, 1993). In sum, achievement in multiple cue judgement environments can be quite high, and in some cases compete with regression models. However, judgement accuracy depends on the characteristics of the task. When task predictability is low or the number of available cues is high, the accuracy and reliability of experts' judgements decrease. If high levels of achievement are possible among experts, this raises the question concerning what type of model experts' use in their judgements, and how closely their judgement policies match the environment.

Can experts' judgement policies be described by linear regression models, or are they sensitive to cue configurations and nonlinearity? If a linear model accounts for a judgement policy then each cue contributes independently to the inferences made. If, however, the judge's use of a cue is dependent on the value of one or more other cues then their judgement policy is configural. One way of assessing whether a judgement policy can be described as linear is by measuring the amount of variance accounted for by a linear model. When specific alternative models are available, these can be pitted against a linear model to assess which best fits the judgement policy (Goldberg, 1971). A third approach involves introducing nonlinear components to a linear model (Millimet & Greenberg, 1973). If significant interactions exist between cues then these can be introduced as interaction terms to assess whether they significantly increase the accounted variance. Despite attempts to describe expert judgement policies with nonlinear and configural models, there is overwhelming support that both judgement tasks and expert judges are best described by linear regression models (Brehmer, 1988, 1994; Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Mear & Firth, 1987; Payne et al, 1993; Ullman & Doherty, 1984). However, there is some indication that nonlinear models can be adopted by judges. Einhorn (1972) for instance, found that a nonlinear conjunctive model best fitted his participants' judgement policies, even though this model did not fit the task structure. SJT research clearly indicates that expert judges in various domains can achieve high degrees of judgement accuracy, and that their judgement policies, as well as the environment can often be described by linear regression models. It is therefore interesting to ask whether experts have insight into how they make their judgements.

If experts make reliably accurate judgements then we may expect them to demonstrate high levels of insight into their own judgement policies. Surprisingly however, studies comparing judges' verbal reports with the actual weight they gave to cues in their judgements provide little evidence that experts have more than a minimal degree of self-insight (Arkes, 1981; Brehmer, 1984; Hoffman, 1960; Roose & Doherty, 1976; Slovic & Lichtenstein, 1971). Many of the studies that have compared self-reported policies with tacit judgement policies inferred by multiple regression ask the participant to give an explicit rating for each cue on a continuous scale. This provides an estimate of participants' explicit beliefs about the weight that they gave to the cues in their judgements. Similar methods include ranking the cues in order of their importance. Arguably, this approach assumes that the tacit judgement policy and the explicit policy that the judge has in mind are both linear, so may fail to take account of nonlinearity or configurality in either policy (Brehmer & Brehmer, 1988). However, even when participants are asked simply to report which cues they used they still fail to show good levels of self-insight, and even choose incorrect cues when asked only to report the single most important one (Brehmer & Brehmer, 1987). Instead, judges appear to overestimate the weight they give to less predictive cues and underestimate the weight they give to more predictive ones, providing a flatter distribution of cue weights in their verbal reports than their judgement policies (Slovic, Fleissner, & Bauman, 1972; Slovic & Lichenstein, 1971). Judges are also shown to report using more cues than they actually gave significant weight to in their judgements (Brehmer & Brehmer, 1987; Evans, Harries, & Dean, 1995). Evans et al. (1995) suggest that the physicians in their study may have listed the cues they believed would be predictive based on medical training, but were unaware that they did not use all the cues they listed when making their judgements.

An alternative approach to asking participants explicitly to identify the cues they used is to ask them to report how they are making their judgements whilst performing a task. Such *process-tracing* methods involve generating a model from the physician's ongoing verbal protocol, which is then formalised as a computer algorithm (Kleinmuntz, 1963, 1975; Newell & Simon, 1961, 1972). Although such model-building is in danger of misinterpretation of protocols by the experimenter, linear regression models appear to be *better* predictors of physicians' judgements than models generated from their own protocols (Einhorn et al, 1979). A novel way of measuring self-insight, introduced by Reilly and Doherty (1992; Harries, Evans, Dennis, & Dean, 1996; Reilly, 1996), requires participants to identify their own judgement policy (e.g. as a set of usefulness indices for each cue) among the judgement policies of other participants. These studies have shown participants correctly to identify their own policies well above chance. Therefore, while judges may have very limited ability verbally to report their judgement policies, they do appear able to distinguish them from those of others. But how important is this kind of self-insight? If an individual makes their judgements in a controlled explicit manner, consciously combining cue values with their subjective weight (importance) of the cues, then we would expect them to show reasonably good levels of insight on measures of verbal report. If an individual is able only to identify their judgement policy then it seems less likely that their explicit beliefs substantially influence their judgements.

Social judgement theory provides a methodological framework for measuring the accuracy and reliability of multiple cue judgement, as well as a means for describing the structure of the environment on the one hand, and the tacit judgement policy of the individual on the other. There is a general consensus across SJT studies that while experts can often achieve good levels of achievement, and can be described well by linear regression models, they lack insight into how they make their judgements. This issue is of much theoretical interest to psychology. If experts do not have conscious explicit knowledge of how they make their judgements, then what kinds of cognitive processes guide the inferences they make? In cognitive psychology, dual-process theories posit that both conscious explicit and unconscious implicit processes can contribute to behaviour. When an individual makes accurate inferences but lacks insight into their judgement policies, this provides some indication that implicit knowledge is contributing to the judgements they make. The cognitive processes involved in multiple cue judgement may, therefore, be explained by dual process theories of thinking.

1.2. DUAL PROCESS THEORIES OF LEARNING AND THINKING

Dual process theories of thinking propose that two types of cognitive processes belonging to separate cognitive systems contribute to the inferences, judgements, and decisions we make in everyday life (Epstein & Pacini, 1999; Evans, 2008; Evans & Over, 1996; Kahneman & Frederick, 2002; Reber, 1993; Sloman, 1996; Stanovich, 1999). *Implicit processes* on the one hand are described as unconscious and automatic, in contrast with *explicit processes* that are conscious and under intentional control. In an unfamiliar city, we may rely on intuition or 'gut feeling' when deciding which restaurant to eat at or we may carefully weigh up the pros and cons of each. In the former case we may not be able to report how we made our decision but intuition, or implicit processing, may have nevertheless guided our decision making. In principle, both implicit and explicit processes can contribute to our thinking. Accordingly, dual process theories have developed in numerous, and quite separate areas of psychology, including implicit learning (Knowlton, Ramus, & Squire, 1992; Reber, 1993), category learning (Ashby & Maddox, 2005), reasoning (Evans, 2003; 2008; Evans & Over, 1996; Stanovich, 1999, 2004), judgement and decision making (Kahneman & Frederick, 2002; 2005) and social cognition (Bargh, 2006; Chaiken & Trope, 1999; Smith & DeCoster, 2000). Before considering how dual process theories have contributed to our understanding of cognitive processes, I first introduce the dual process framework in more detail. This framework will form the theoretical basis of our study of multiple cue judgement in the present thesis.

1.2.1 The dual process framework

While dual process theories differ in terms of how implicit and explicit processes interact to control behaviour, general characteristics of the two modes of thought are shared across theories (Evans, 2008). Explicit processes (also referred to as analytic, Evans, 1984; Evans, 2006, or System 2 processes, Stanovich, 1999) are associated with conscious controlled thought. This is a domain-general type of processing, and is involved in hypotheses-testing, imagining of counterfactual states of the world, and rule-based thinking. Explicit processing and its underlying neurological bases is believed to have developed relatively recently in human evolution (Evans & Over, 1996; Stanovich, 2004). Some even regard this type of thinking as uniquely human. However, explicit processing is effortful and heavily demanding on the individual's limited working memory resources (Barrett, Tugade, & Engle, 2004). For this reason, individual differences in cognitive ability are often positively associated with performance on tasks that draw on explicit modes of thought (Stanovich &

West, 1998a,b,c). Indeed, much conscious thinking is likely constrained by the sequential and capacity limited nature of explicit processing. For instance, medical diagnosis involves adjusting hypotheses about a patients' condition in light of numerous symptoms and medical examinations (Wigton, 1996). Multiple cue judgement of this type may be beyond the cognitive capacity of the physician's conscious processing, but may be supported by less constrained implicit processes.

In contrast with explicit processing, implicit processing (also referred to as System 1 processing; Stanovich, 1999) is described as unconscious and automatic (Reber, 1993, Berry & Dienes, 1993). This mode of thought is related to associative learning and heuristic processing (Evans, 2003; Kahneman & Frederick, 2002; Sloman, 1996). While incremental learning may be slow, and require many instances for learning, implicit associative processes are believed to be fast at generating responses (Ashby & Maddox, 2005). Their underlying mechanisms are often described by neural network models, and are believed to be massively parallel (Dienes, 1992; Gibson, Fichman, & Plaut, 1997; Rumelhart, McClelland, & the PDP Research Group). This affords implicit associative processes a high capacity for dealing with large amounts of information. In contrast, heuristic processes automatically contextualise problems by directing attention to relevant information based on prior beliefs (Evans, 2008; Kahneman & Frederick, 2005; Tversky & Kahneman, 1983). Implicit processing is also involved in the automaticity of learned skills. Actions that initially require effortful thinking, such as performing mental arithmetic or a complex skill can become automated through practice, drawing less of explicit processing (Anderson, 1983; Anderson et al, 2004; Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Logan, 1988). A corresponding decrease in response time is observed when skills become automated by implicit processes. Unlike explicit thinking, implicit modes of thought do not load heavily on limited working memory resources. Tasks, such as judgement from multiple cues, may then be better suited to high capacity implicit learning processes. However, while implicit modes of thought do not burden working memory resources (Barrett et al, 2004), attention does appear necessary for both implicit and explicit learning in complex environments (Cohen, Ivry, & Keele, 1990; Frensch, Buchner, & Lin, 1994).

The implicit system appears to comprise a number of domain-specific subsystems, or modules (Fodor, 1983), including associative and heuristic processes. Although some simple explicit strategies can be described as heuristics (e.g. Bröder, 2003), much heuristic processing appears to be implicit and unconscious (Evans, 2006; Stanovich, 2004; Stanovich & West, 1998b). While the implicit subsystems likely developed for different purposes, the neurological bases of the implicit system are believed to have originated earlier in human evolution than explicit processing and is said to have some commonalities with animal cognition (Epstein & Pacini, 1999). Although there appears to be a consensus among dual process theorists of the defining characteristics of implicit and explicit modes of thought, the attention that theorists give to specific implicit processes, and the interaction between implicit and explicit processes in controlling behaviour differs greatly between research traditions.

Research on reasoning and decision making indicates that heuristic processes provide automatic representations of task contents. When these representations are in conflict with the logic of a task the individual must inhibit a heuristic response in order to reason the task logically (Evans, 2006; Kahneman & Tversky, 1972). This requires effortful explicit thinking on the part of the individual. In contrast, implicit learning research is concerned primarily with associative learning mechanisms, and often implies that explicit processing contributes little to improved performance in complex learning tasks (Cohen et al, 1990; Reber, 1993), and can even impede learning (Turner & Fischler, 1993). The contribution of implicit and explicit modes of thought to other types of learning such as category learning, appear to depend on characteristics of the task. When categories are defined by a small number of cues and category rules are easily verbalised, explicit knowledge can guide correct categorical decisions that are otherwise driven by implicit processes (Maddox & Ashby, 1993; Ashby & Maddox, 2005). Some dual process theorists propose that implicit and explicit learning occurs in parallel, and that both types of processes compete to control behaviour (Sloman, 1996). This is in contrast with other dual process models proposed in reasoning and judgement literatures that describe responses as determined by implicit heuristic processes unless overridden by explicit analytic thinking. Therefore, while the characteristics of implicit and explicit processes are shared by dual process theories across areas of psychology, the contribution of these processes to the inferences we make depends on the types of behaviour that are studied. I now turn to a discussion of how dual process theories of thinking have contributed to our understanding of implicit learning, category learning, reasoning, and judgement and decision making.

1.2.2. Implicit learning

Implicit learning is traditionally the study of learning without awareness. In numerous tasks, including sequence learning (Nissen & Bullemer, 1987; Reed & Johnson, 1994; Stadler, 1995), artificial grammar learning (Pothos, 2007; Reber, 1967, 1993), and dynamic systems control tasks (Berry, 1991; Berry & Broadbent, 1984; Hayes & Broadbent, 1988), improved performance is often not accompanied by explicit knowledge of the task.

Artificial grammar learning (AGL) provides a classic example of learning without awareness. Participants are presented lists of letter-strings as part of a memory task for later recall. Unknown to the participant the letter strings are in fact generated by a complex grammatical rule. In a test phase, participants are then able to categorise novel letter strings as grammatical or non-grammatical well above chance when letter strings are generated by the same grammatical rules used to generate the learning set. This is taken as evidence that participants have learned something about the grammatical rules that underlined the letter strings (Brooks, 1978; Dulany, 1984; Fied & Holyoak, 1984; Reber, 1965, 1967, 1993). Interestingly, in post-task interviews the same participants are not usually able verbally to report any features of the grammar, and are often unaware they have learned anything at all. Explicit processing may not only be irrelevant to performance in some tasks, but can even impede learning of complex rule structures when participants are encouraged to search for the underlying rules of the task (Turner & Fischler, 1993).

Initially, AGL and other types of complex structure learning were taken as evidence that people acquire complex implicit knowledge in the form of abstract rules (Lewicki, Czysewska, & Hoffman, 1987; Reber, 1993). For an AGL task, a rule could take the form "if a letter string begins with the letter H followed by L or M, then the string is grammatical" when judging the grammaticality of letter strings. A number of alternative approaches challenge this assumption and have shown that complex learning can be driven by exemplar memory (Brooks & Vokey, 1991; Vokey & Brooks, 1992) and associative learning processes (Cleeremans & McClelland, 1991; Dienes, 1992). In the case of AGL, explicit knowledge of list fragments may even account for reliable judgements of grammaticality (Perruchet & Pacteau, 1990). Complex structure learning likely involves multiple-memorial systems including abstract and implicit associative processing, as well as explicit learning (Knowlton & Squire 1994, 1996), an idea that is supported by studies with amnesic patients (Knowlton et al, 1992).

Learning of complex grammatical structures and sequences is often shown to occur in the absence of awareness. In these studies, the task is presented as a memory test or a reaction time task, and participants are not made aware that the stimuli are structured by an underlying rule. However, implicit learning of complex rules is also demonstrated when participants are instructed explicitly to discover the rules of a task. In a commonly used rule-discovery task, known as Dynamic Systems Control (DSC), participants are instructed to learn to control the output of a system by manipulating its inputs, a task they are informed can only be achieved by learning the rules that govern the system. These systems can take many forms, such as a simulated person, sugar factory, or economy (e.g. Berry, 1991; Berry & Broadbent, 1984; Broadbent & Aston, 1978; Hayes & Broadbent, 1988). What these studies demonstrate is that while participants tend to improve in their ability to control a complex system by manipulating its inputs, they are not often able verbally to report how they gained control of the system.

Interestingly, participants do acquire accurate explicit knowledge of DSC tasks when given the opportunity to practice with each input variable separately (Broadbent et al, 1986). Explicit learning also occurs when variables are few in number, and when the relations between them are 'salient' (Berry & Broadbent, 1988; Dienes & Fahey, 1995; Geddes & Stevenson, 1997). For example, Hayes and Broadbent (1988) instructed participants to learn to improve the mood of a simulated person to 'friendly', then keep it there. For one group the rule governing the person's character was salient (dependent on the participant's most recent input), whereas for a second group the rule was non-salient (dependent on the participant's previous input). Participants trained on the salient rule showed good explicit knowledge and accordingly performed poorly when a working memory load was applied during a rule-change block, whereas participants learning the non-salient rule demonstrated poor explicit knowledge but were less affected by the concurrent load task.

In sum, learning of complex rule structures often occurs implicitly, even when participants are instructed to search for the underlying rules of a task. Explicit learning does occur, however, when variables are few in number, underlying rules are salient, or when participants are given the opportunity to observe each part of a complex rule. This suggests that explicit hypotheses-testing processes may be constrained by limited processing capacity, which may in part be due to limited working memory resources (Barrett et al, 2004). Multiple learning processes appear to be involved in learning of complex rules, including abstract and associative implicit processes, as well as exemplar memory and explicit learning processes. However, implicit learning research has received strong criticism from sceptics (Redington & Chater, 1996; Perruchet, Gallego, & Savy, 1990). In a comprehensive review of the literature, Shanks and St. John (1994) argue that evidence used to propose separate learning systems is not sufficient, and suggest two criteria must be met before such a claim can be made. The *Information Criterion* urges that the experimenter must be sure that any test of explicit knowledge draws on the same information that was required to perform the task. This comes from their suggestion that many tests of explicit knowledge ask for information that is not directly relevant for performing the task. The *Sensitivity Criterion* argues that any test of explicit knowledge must be as sensitive to information available in consciousness as the measure of performance. This is because a test of explicit knowledge may be less sensitive to the same explicit knowledge that was used to perform the task, which could lead to the false assumption that learning occurred in absence of awareness. While the criteria set by Shanks and St. John are endorsed by some researchers (Berry, 1994), others argue the criteria are too strict (Dienes & Perner, 1994; Holyoak & Gattis, 1994). In the experimental studies of the present thesis we attempt to comply with the criteria set by Shanks and St. John.

1.2.3 Category learning

Dual process theories have also become increasingly popular in the category learning literature (Ashby et al, 1998; Brooks, 1978; Erickson & Kruschke, 1998; Gluck, Shohamy, & Myers, 2002; Smith, Patalono, & Jonides, 1998). These models distinguish between a rule-based system associated with explicit hypotheses-testing on one hand, and an implicit procedural learning system on the other. In category learning tasks participants learn to categorise stimuli as belonging to usually one of two categories based on a small number of cues. These can be either probabilistically or deterministically related to category membership. In most cases cues are represented as visual images, such as a box shape that can change in shape and size. A category rule could follow "if the box is tall and thin then choose category A, else if the box is short and wide choose category B". In other cases a category rule may not be easy to verbalise, such as when the rotation of a bar indicates category membership. Other examples of non-verbalisable tasks include complex visual images such as dot patterns.

In dual process theories of category learning, a rule-based system involved in conscious effortful thinking is believed to load heavily on limited working memory resources (Price, 2006). This is in contrast with an implicit procedural system that learns by associative processing, and is less dependent on working memory (Gluck, Oliver, & Myers, 1996). Similar to studies of implicit learning, explicit rule-based processes are limited to learning simple verbalisable rules for categorisation (Maddox, Filoteo, Hejl, & Ing, 2004). When the underlying rules of the task are not easy to verbalise learning appears to depend on associative processing of the implicit procedural system. Furthermore, in tasks that are best learned implicitly, low levels of self-insight also accompany poor explicit knowledge of the task (Gluck, Shohamy, & Myers, 2002; Price, 2005). Lending support to the idea that distinct cognitive systems are involved in category learning, separate regions of the brain have been found to be associated with category learning tasks that are learned by rule-based and procedural learning processes (Knowlton, Mangels, & Squire, 1996; Knowlton, et al, 1996).

A number of methods have become popular in category learning research for dissociating rule-based and procedural learning systems, revealing some interesting findings. Explicit processing appears capable of learning only a small number of cues when category rules are verbalisable (Waldron and Ashby, 2002). In these tasks the addition of a concurrent working memory load interferes more with learning of one cue than three cues. In an innovative version of the traditional category learning task, Maddox et al (2004) required some participants to categorise stimuli into one of four categories. They found that increasing the number of categories from two to four in tasks that were verbalisable worsened performance. However, increasing category numbers did not affect performance when category rules were not easy to verbalise. This indicates that explicit processing may have been involved in learning of the verbalisable but not the non-verbalisable task.

Multiple memory systems, including explicit rule-based, implicit associative, as well as exemplar memory processes appear to be involved in inferring category membership from multiple cues (Erickson & Kruschke, 1998; Nosofsky & Johanson, 2000). The task characteristics that influence the role of explicit processing are similar in implicit learning and category learning literatures. In easier tasks that contain few variables determined by salient verbalisable rules explicit processing appears to contribute to accurate performance. However, when tasks are made more difficult and contain many variables or underlying rules that are difficult to verbalise, explicit processing contributes little to performance.

1.2.4. Reasoning

In contrast with the implicit learning and category learning literatures, studies of human reasoning focus on default heuristic processing of the implicit system. Heuristic processes are believed to control behaviour unless overridden by effortful explicit thinking. A classic example of how both implicit and explicit processes can influence reasoning is illustrated by the 'belief-bias' effect in syllogistic reasoning tasks (Evans, 2003, 2008; Klauer, Musch, & Naumer, 2000). In these tasks the participant is presented two premises followed by a conclusion and asked to rate whether the conclusion necessarily follows from the premises, that is, whether they logically imply the conclusion. However, the conclusion can be believable or unbelievable, allowing the participant to judge the task according to whether the conclusion logically follows from the premises, or whether it is believable or not. Typically participants are influenced by both the logical validity of the conclusion given the premises *and* the believability of the conclusion, making judgements more in line with the believability of conclusions, especial when problems are not logically valid (Evans, 2003, 2007; Evans, Barston, Pollard, 1983). In accordance with dual process theories of thinking it is argued that participants do reason about the logic of the arguments in syllogistic tasks, drawing on working memory dependent explicit processing, but are also influenced by prior beliefs generated by implicit heuristic processes (Evans, 2003, 2008). Individual differences in cognitive ability (Stanovich & West, 1997), working memory capacity (De Neys, 2006; Feldman, Tugade, & Engle, 2004), and age (Gilinsky, & Judd, 1994) are shown to account for logical reasoning in syllogistic tasks. Individuals of higher cognitive ability are more likely to inhibit and override prior belief and provide the normatively correct response.

The contribution of implicit and explicit processes to reasoning is also demonstrated in the Wason selection task (Evans, 1998, 2006; Wason & Evans, 1975; Wason & Johnson-Laird, 1972). In its abstract form the participant is presented four cards face down, with A, D, 3, and 7 printed on each card. The participant is then given the rule "*If there is an A on one side of the card, then there is a 3 on the other side*", and asked to pick only the appropriate cards that need to be turned over in order to test whether the rule is true. Although turning over cards A and 7 are the only selections that will falsify the rule, that is, prove that there *is* a 3 on the other side of card A, and *not* an A on the other side of card 7, only 10-20% of participants choose these cards. Interestingly, the cards that the majority of participants choose, A and 3, are the same as those stated in the rule. When the rule is changed to "*If there is an A on one side of the card, then there is **not** a 3 on the other side*" cards A and 3 become the correct choices and are still the selections made by the majority of participants (Evans, 1998, 2003). This effect, known as 'matching bias', of selecting the cards stated in the rule indicates that participants are influenced by automatic heuristic processes rather than the logic of the task (Evans, 1998, 2003). In such cases, heuristic processing of the implicit system is believed to influence the inferences made by individuals unless they intervene with effortful explicit reasoning (Evans, 2003).

Belief-bias and matching bias effects demonstrate a competition between implicit heuristic and explicit analytic processes (Evans, 2003, 2008). If participants do not inhibit and override intuition their inferences can be biased by heuristic processes. However, if participants explicitly inhibit automatic intuitive responses they are at least more likely to reason analytically (Evans, 2007). For instance, when instructions strongly encourage participants to think about the logic of syllogistic tasks belief-bias tends to reduce, suggesting task instructions elicit explicit analytic thinking (Evans, 2000). Lending support to this argument, neuropsychological evidence has shown that distinct brain regions become activated when participants reason according to the logic of syllogistic tasks compared with when they are influenced by the believability of the conclusion (Goel & Dolan, 2003). Similarly, matching-bias can be interpreted as a competition between default heuristic and effortful explicit processes. For example, Houdé et al. (2000) found that when participants are trained to watch out for the '*habit we all have of concentrating on the cards with the letter or number mentioned in the rule* (pp. 726)' when performing the Wason selection task, that is, to avoid matching bias, they are more likely to reason logically.

Houdé et al. found this instructional effect was also associated with a shift in neurological activity in separate brain regions.

Whether participants inhibit heuristic processing and reason logically also depends on cognitive ability, indicated by SAT scores and working memory capacity measures (Stanovich & West, 1998a,b). Those of higher cognitive ability are more likely to engage in analytic thinking and inhibit a heuristic response (Stanovich, 2004). They are also more likely to provide normatively correct inferences when they do engage in explicit reasoning (Kokis, Macpherson, Toplak, West, & Stanovich, 2002). However, if an individual does not detect that an explicit effort is required of them then regardless of their cognitive ability, they will likely settle for a default heuristic response (Kahneman, 2000; Stanovich & West, 2008). Indeed, a number of thinking dispositions, including 'open-mindedness' and 'need-for-cognition' are also shown to predict performance in reasoning tasks even when individual differences in cognitive ability are controlled (Stanovich, 1999). Hence, high cognitive ability is not always sufficient for logical reasoning, the individual must also be inclined to think analytically.

Dual process theories often describe a two-stage process of reasoning. Intuitive responses are generated quickly (stage 1), and these will likely influence the inferences an individual makes unless they consciously intervene (stage 2), but this process is often slow and effortful. Alternatively, reasoning could be described as a competition between parallel implicit and explicit systems. Sloman (1996, 2002) explains that people are sometimes aware of a conflict between belief and logic when reasoning about conditional statements and propositions. In this way, conflicting inferences generated by separate cognitive systems can compete to control behaviour. This kind of dual process model is popular in the social cognitive literature (Epstein, 1994; Epstein & Pacini, 1999; Smith & DeCoster, 2000). For instance, prejudice and stereotyping can be thought of as a competition between implicit attitudes and explicit beliefs (Epstein & Pacini, 1999; Kawakami, Dion, & Dovidio, 1999).

1.2.5. Judgement and decision making

Kahneman and Tversky (1972, 1973; see also Gilovich, Griffen, & Kahneman, 2002) in their "heuristics and biases" program of the early 1970s outlined a number of specific heuristics that appeared to describe peoples' judgements of probability and likelihood. Their discovery of such heuristics as availability, representativeness, and anchoring and adjustment, revealed that when making judgements from probabilistic information participants' decision-making is not simply less sophisticated than normative models, but qualitatively different. Of interest here is how these heuristics in recent years have been attributed to implicit processing, generating fast intuitive inferences unless overridden by effortful explicit processing (Kahneman & Frederick, 2002, 2005).

One such heuristic, the representativeness heuristic, influences the way judgements are inferred from diagnostic information. Imagine for example, you are presented a personality description which is highly diagnostic of an engineer, and asked whether the description most likely describes an engineer or a lawyer. You may judge the description as most likely belonging to the former. However, when told that the base-rate proportion of engineers in the population from which the description was taken is very low compared to the proportion of lawyers, your initial estimate should reduce. How close your estimate should be to the base-rate proportion of engineers and lawyers depends on how diagnostic your information is of an engineer than a lawyer, and the proportion of engineers and lawyers in the population. Kahneman and Tversky (1973; see also Koehler, 1996) found surprisingly that participants' judgements took little account of base-rate information, indicating that participants were heavily influenced by the personality descriptions. Thus, participants' judgements were overly sensitive to how 'representative' the description was of the individual.

In the example above, heuristic processing led to bias in judgement by neglect of base-rate information. The same heuristic can also lead to conjunction errors however, in which the probability of the conjunction of two events is rated as higher than the probability of either of the two events alone. Another famous example, known as the Linda problem, illustrates this point. Participants are provided a compelling description of Linda, containing attributes representative of both a bank teller and a feminist. When participants are asked to rate the likelihood that Linda is a 'bank teller', a 'feminist', and a 'bank teller who is active in the feminist movement', the major majority of participants rank the conjunction of the two events as higher than either event. This is because although the conjunction of two events (i.e. bank teller *and* feminist) is equally or less likely than either of the two events occurring alone, based on Linda's description she appears most representative of a feminist bank-teller (Tversky & Kahneman, 1983).

From a dual process perspective, this failure to conform to normative rules and rely on intuitive judgements of representativeness is attributed to implicit heuristic processing. In such cases explicit processes may simply fail to inhibit or replace automatic inferences (Kahneman & Frederick, 2002, 2005). Indeed, while participants with statistical knowledge are generally less likely to fall prey to the conjunction fallacy, participants less educated in statistics are also shown to be aware of the nested set relations between single events and their conjunctions in post-task interviews (Tversky & Kahneman, 1983). Tversky and Kahneman (1983) suggest that while both groups of participants may be aware of the conjunction rule, only participants with statistical knowledge are able to see the significance of the rule and override their intuitive judgements of representativeness. Similarly, participants who do not commit the conjunction fallacy tend be of higher cognitive ability, measured by SAT scores and working memory capacity, indicating that such participants may be more likely, or more able, to inhibit intuitive responding and apply the conjunction rule (Stanovich & West, 1998b). This suggests explicit processes may indeed be capable of overriding heuristic inferences of the implicit system, but often fail to do so.

In recent years, inhibition mechanisms of working memory have received attention in explaining why participants often shown in post-task interviews to be aware of normative rules nevertheless fail to reason logically (Houdé & Moutier, 1996). Similar to reasoning tasks, when participants are trained to avoid making common errors caused by attending to the representativeness of evidence in the Linda problem, participants' judgements tend to conform more with the conjunction rule (Moutier & Houdé, 2003). This indicates that bias is not due solely to an inability to reason with normative rules, but to a failure to inhibit and override intuitive responding.

Participants also underestimate base-rate information in learning tasks that provide outcome feedback in response to trial-by-trial judgements (Edwards, 1968; Kruschke, 1996; Slovic & Lichtenstein, 1971). However, it is unlikely that the same failure to incorporate base-rate information is due to similar cognitive processes. The representativeness heuristic may be well suited to making intuitive one-off judgements, whereas implicit associative processing may explain base-rate neglect in learning tasks. Gluck and Bower (1988) found that an associative learning model accounted for base-rate neglect in experiential learning tasks, suggesting that participants learned direct associations between diagnostic cues and outcomes, which did not take into account the base-rate probabilities of outcomes. Similarly, when making judgements about the probability of outcomes given the presence of cues in experiential learning tasks, participants are shown to commit the conjunction fallacy by assuming the probability of the conjunction of two outcomes to be greater than the probability of each outcome given the cues (Cobos, Almaraz, & García-Madruga, 2003). This appears to be due to learning of direct cue-outcome associations. When the conjunction of two events given the presence of the cues is considered participants sum the associative strength between the cues and each outcome, making inferences more in line with the conjunction fallacy.

When making judgements of probability, the implicit system can generate automatic intuitive inferences unless overridden by effortful explicit processing. However, it is important to note that biases in judgement occur when heuristics are adhered to that do not cohere with normatively prescribed rules. This is not to say that such heuristics do not correspond with the environment, and in indeed may serve as useful rules-of-thumb (Tversky & Kahneman, 1973; Kahneman, 2000). Recent theorists more concerned with the ecological validity of heuristics than their logical coherence suggest that heuristic processing can be highly accurate (Czerlinski, Gigerenzer, & Goldstein, 1999; Gigerenzer, Todd, & the ABC Research Group, 1999). These theorists however, make no claims about whether applying heuristics is driven by implicit or explicit processes.

When participants make one-off judgements from verbal descriptions or numeric information in reasoning and judgement and decision making tasks heuristics such as matching-bias, belief-bias, and the representativeness heuristic appear suited to generating fast intuitive responses. In the case of experiential learning however, in which participants make trial-by-trial judgements and receive outcome feedback, associative learning processes may provide a better account of human judgement.

1.3. CONCLUSIONS

Social Judgement Theory (SJT) provides a methodological framework for measuring experts' judgement policies. As well as describing which cues influence judgement when multiple cues are available, SJT provides a means for modelling the environment and comparing the individual's judgement policy with the structure of the environment. SJT studies have shown that despite the probabilistic nature of judgement tasks, such as medical diagnosis and weather forecasting, experts often attain good levels of achievement. However, this is in contrast with the finding that experts often have poor levels of self-insight, which casts doubt on the extent to which conscious explicit thinking contributes to expertise. Dual process theories in psychology have shown that both controlled explicit and automatic implicit processes can influence our judgements, decisions, and reasoning about information in the environment. When reasoning about the logic of a task, or making a decision based on diagnostic information implicit heuristic processes direct (or even bias) our attention towards relevant information. In contrast, when learning from experience, implicit associative processes acquire complex knowledge incrementally. In both cases, controlled explicit processing can compete to control the inferences we make. Whether explicit processing is successful depends in part on our limited working memory resources. For this reason, logical reasoning is often poor, and knowledge of complex environments that require us to attend to multiple pieces of information is often not acquired explicitly.

The Multiple Cue Probability Learning (MCPL) paradigm provides a controlled environment for studying the cognitive processes involved in judgement from multiple cues. In these tasks, expertise is simulated by training participants on a novel task environment. There is a large body of work dedicated to the study of learning from feedback using the MCPL paradigm, which is reviewed in the next chapter. In the present thesis we apply methods not previously used in the MCPL literature as a means for measuring the contribution of implicit and explicit processes to multiple cue judgement.

Chapter 2

Multiple Cue Probability Learning and the Present Research

Social judgement theory (SJT) demonstrates how lens model analysis can be used to measure the judgement policies of experts. However, in SJT studies, judgement policies are studied only after expertise is acquired. Expertise is likely influenced by the amount of experience an expert has or the types of training they received, in addition to task factors. It is therefore difficult to study the kinds of cognitive processes involved in multiple cue judgement when expertise has already been acquired. In an almost entirely separate field of research, more akin with cognitive psychology, lens model analysis has been used to study learning from multiple cues in novel environments. The multiple cue probability learning (MCPL) paradigm provides an ideal methodology for studying expertise following training in controlled environments. MCPL research is also rooted in Egon Brunswik's ecological approach to psychology, and uses many of the same methods of analysis as SJT for measuring expertise following training (Holzworth, 2001).

In a seminal paper by Smedslund (1955) the application of lens model analysis to learning had begun. In the following years MCPL contributed greatly to the study of complex learning, and is discussed in some detail in the present chapter. In the experimental studies of the present thesis the MCPL paradigm is used to train expertise in order to explore the cognitive processes involved in multiple cue judgement. The rationale of the experimental studies is introduced within the dual process framework, and details are given concerning the experimental methodology.

2.1. MULTIPLE CUE PROBABILITY LEARNING

In MCPL tasks participants are trained on a small number of cues. Their task is to learn to predict criterion values on each trial. Cues and criterion can take continuous or ordinal values and can be presented as visual cues such as the length of a line, or verbal labels. In a learning phase participants are provided outcome feedback (actual outcome) in response to each judgement. This provides participants the opportunity to learn the weight to give to each cue in judgement and the relation between cues and the criterion. In order to simulate learning of expertise in the types of uncertain environments that people usually learn, a noise element is usually included with the outcome feedback participants receive in the learning phase. This provides a probabilistic element to multiple cue tasks. Following training, participants are provided a number of test trials designed to measure their learning. In these trials feedback is not provided in response to judgement. While performance scores are usually measured along with participants' judgement policies, few MCPL studies have considered the degree to which participants have insight into their judgement policies, and whether they acquire accurate explicit knowledge of the task. This may in part be due to an assumption in MCPL studies that learning is conscious and explicit. According to Juslin, Jones, Olsson, and Winman (2003) a common view in the MCPL literature is that people "abstract explicit representations of the cue-criterion relations". This may explain why explicit knowledge and self-insight is rarely assessed, since people should have good insight if judgment is explicit and explicit knowledge should correspond with performance.

2.1.1. Background

One of the main findings to come out of early MCPL research was participants' difficulty in learning certain types of cue-criterion relations. It became evident that although participants were often able to learn non-linear U-shaped and inverted U-shaped relations (see Figure 2.1), their performance in tasks that contained these types of cues was far poorer than for linear cue tasks (Deane et al, 1972; Hammond & Summers, 1965, 1972; Sheets & Miller, 1974; Summers & Hammond, 1966). Furthermore, participants also appeared to perform less well with negative linear than positive linear cues (Evans, Clibbens, Cattani, Harries, Dennis, 2003; Evans, Clibbens, & Harris, 2005; Naylor & Clark, 1968).

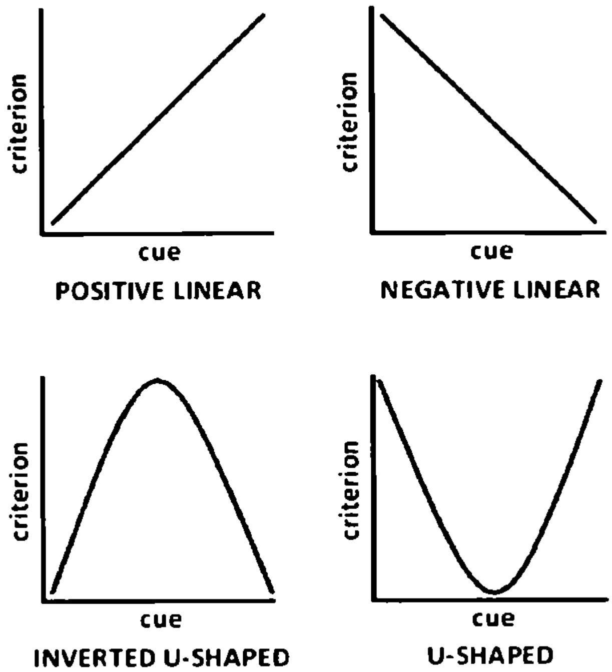

Figure 2.1: The four main cue-criterion relations used in MCPL tasks. Top panel: positive linear and negative linear relations. Bottom panel: Inverted U-shaped and U-shaped non-linear relations.

As evidence of participants' difficulty with learning negative linear and non-linear cues mounted, many theorists began to conclude that people are simply unable to learn from outcome feedback in MCPL tasks (Brehmer, 1980). For some authors learning appeared to be a "very slow and peculiarly ineffective process" (Smedslund, 1955), while others concluded that "subjects are not bad intuitive statisticians, they are not statisticians at all" (Brehmer & Kuylenstierna, 1978). It appeared that participants persistently tested hypotheses that were not supported by outcome feedback (Einhorn & Hogarth, 1978), often retested hypotheses they had previously rejected (Brehmer, 1979), and only performed well when told which cues to use and how to use them (Deane, Hammond, & Summers, 1972; Hammond, 1971). This led some authors to argue that outcome feedback is simply not sufficient for learning in complex MCPL tasks (Hammond, 1971; Hammond & Summer, 1972; Hoffman, Earle, & Slovic, 1981).

A number of alternatives to outcome feedback were developed in an attempt to improve performance in difficult versions of MCPL tasks, namely U-shaped and inverted U-shaped tasks. This involved either providing participants feedback about the environment structure (task information), their own judgement policy (cognitive information), or the match between their judgement policy and the environment model (functional validity information). These can be provided to the participant either verbally, graphically, or as statistical information. Task information (TI) appears to improve learning in complex MCPL tasks containing negative linear, U-shaped, and inverted-U shaped relations, while cognitive information (CI) and functional validity information (FVI) general have little effect on performance (Balzer, Doherty, & O'Conner, 1989; Hoffman et al, 1981; Lindell, 1976; Newton, 1965). This is likely because TI is the only type of feedback that informs the participant explicitly about the relations between cues and the criterion and the relative importance of each cue. There is also some suggestion that a combination of both TI and CI provides the best feedback for learning (Hammond & Boyle, 1971; Schmitt, Coyle, & King, 1976). Indeed, providing information about the environment and the individual's judgement policy allows the participant to compare their judgement policy with the model of the environment. While alternatives to outcome feedback are useful for improving the judgement policies of experts, graphical aids and statistical information about the environment structure and one's judgement policy are arguably rarely available in natural environments.

The above review paints a grim picture of people's ability to learn in MCPL tasks. However, it is important to note that participants are often able to learn nonlinear U-shaped and inverted U-shaped relations, even when combined with linear cues in mixed-cue tasks, albeit less well than linear cues (Hammond & Summers, 1965). It is also consistently shown that participants can achieve very good performance when all relevant cues have a positive linear relation to criterion (Brehmer & Kuylenstierna, 1978; Naylor & Domine, 1981), and are able to distinguish relevant from irrelevant cues (Evans et al, 2003). Although participants perform less well when cues have a negative linear relation to criterion (De Klerk, De Leeuw, & Oppe, 1966; Naylor & Clark, 1968), they nevertheless perform better than chance even when up to 25% of noise is added to outcome feedback in a learning phase of only 80 trials (Evans et al, 2003; 2005). It has also been argued that in real world environments much of the process of learning from multiple cues involves discovering which ones are important, that is, which cues should be added to one's cognitive model of the environment, and which should be removed (Klayman, 1984). Klayman (1984, 1988) suggests that by presenting participants with an explicit list of cues, as is usually the case, this downplays their ability to discover cues in the environment. He found participants are able to discover, after extensive training, which visual cues in a complex visual display are relevant for predicting a criterion.

In sum, learning in MCPL tasks is heavily affected by the types of relation cues have with the criterion. Early research using cues with nonlinear cue-criterion relations suggested participants simply do not learn from outcome feedback. However, this finding stands in contrast with research using linear cues, which often report very good performance. Participants are shown to perform well when cues have a linear relation to criterion, and better still when the relation is positive.

2.1.2. The cognitive processes involved in MCPL

MCPL theorists appear to assume that people learn by consciously testing hypotheses against the feedback they receive. One possibility is that following training, what knowledge people have acquired of the task is explicit and available for verbal report. The general consensus among MCPL researchers appear to be that participants test the hypotheses that most easily come to mind until the outcome feedback they receive confirms an hypothesis, they then begin to abstract rule-based explicit knowledge of each cues' importance (Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Kleinmuntz, 1963, 1975; Newell & Simon, 1961, 1972; Brehmer, 1973, 1974, 1980; Juslin, Olsson, & Olsson, 2008). While many of these theorists do not describe an 'explicit' theory of judgement, people's hypothesis testing strategies are usually assessed using methods of verbal report and questionnaires. For instance, people tend to rate explicitly that positive linear relations are more prevalent in their environment than negative linear ones in post-task interviews (Brehmer, 1974). Hence, it is likely that hypothesis testing in MCPL is conscious and driven by controlled explicit processes. Accordingly, a participant may discover that a positive linear hypothesis accounts for the cue-outcome associations they are observing and begin to formalise explicit rules such as "*high values on Cue A go with high outcome values, and low values on Cue A go with low outcome values*" (Juslin, Jones, Olsson, & Winman, 2003). However, there is some dispute that theorists implicitly equate the multiple linear regression analyses used to reveal participants' judgement policies with the *actual* cognitive processes involved in performing the task (Dawes, 1975). This may occur if we assume that participants have a linear additive model in mind, leading to the assumption that participants explicitly combine cue values with their subjective beliefs about each cue in a controlled and conscious manner (Simon, 1976)

Brehmer (1974, 1979, 1980) suggests that participants approach MCPL by testing hypotheses about cue-criterion relations one-by-one in order that hypotheses most easily come to mind, first testing a positive then a negative linear hypothesis, followed by nonlinear hypotheses. The order that participants test hypotheses against feedback is also influenced by their beliefs about which types of cue-criterion relations are most common in the environment. Participants tend to list more examples of positive cues than negative cues, and more examples of linear relations than nonlinear relations (Brehmer, 1974). Participants also tend to rate positive cues as more prevalent in the environment than negative cues, and test hypotheses in an order consistent with this account (Brehmer, 1974; 1976; 1978; Brehmer & Kuylenstierna, 1978). Similarly, Naylor and Clark (1968) suggest that participants appear to be biased towards looking for a "positive relatedness" between cues and criterion, explaining why they find negative cues harder to learn. Interestingly, learning of negative cues is poorer when added noise is increased or cue validities are decreased, whereas the same factors do not affect learning of positive cues, suggesting that negative cue learning is difficult in part due to participants' inability to reject an initial expectation of positive cues (Brehmer, 1973).

MCPL studies indicate that participants' learning from multiple cues is not only biased towards positive and linear cue-criterion relations but is also surprisingly inefficient. MCPL tasks are arguably very complex when approached by explicit hypothesis testing. As

Hammond (1971) notes, cues may differ in their relation to the criterion, similar outcomes can be generated by different patterns of cue values, and due to the noise element added to outcome feedback similar patterns of cue values can produce different outcomes. However, when participants are given the opportunity to select their own cue values in order to encourage explicit hypothesis testing of each cue-criterion relation, participants often fail to manipulate cue values in an efficient and informative way. They often choose not to change any cue values on many trials, generating similar outcomes, or opt to change all cues on a single trial as would happen if cue values were randomly generated (Hoffman et al, 1981). In tasks containing only linear cues, however, participants do appear to benefit from the opportunity to test hypotheses when selecting cue values to correspond with outcome values (Enkvist et al, 2006). Similarly, when cue-criterion relations are made more salient to facilitate hypothesis testing by holding some cue values or the criterion value constant over a number of trials, learning is facilitated in tasks containing linear relations with deterministic outcome feedback (Uhl, 1960). However, the same manipulations do not appear to aid learning in more complex tasks containing inverted U-shaped cues (Hoffman et al, 1981).

It appears that methods aimed on enhancing participants' explicit hypothesis-testing do not facilitate learning in highly complex tasks, but instead help to make salient the cue-criterion relations that participants find easier to learn (i.e. linear cues). Participants also appear insensitive to the probabilistic nature of MCPL tasks, and are often too quick to abandon hypotheses. When told explicitly that similar cue values can lead to different outcomes due to an added noise element, learning does not improve even in simpler positive linear tasks (Brehmer, 1980; Brehmer & Kuylesntierna, 1978).

In sum, the MCPL literature appears to suggest that learning in multiple cue environments involves consciously testing hypotheses against outcome feedback in a controlled effortful manner. While MCPL theorists do not rule out the possibility that judgement can be more intuitive (See Hammond, 1996), the assumption appears to be that analysis of self-reports and verbal protocols is sufficient to capture the process of judgement. For instance, Hammond and Summers (1972) propose that performance on more difficult versions of MCPL tasks that contain nonlinear cues is poor in part due to people's difficulty in applying rule-based knowledge. For Hammond and Summers judgement can be taxing on cognitive resources even when one knows how the cues are related to the criterion and how to use them. The assumption here appears to be that judgement is constrained by limited cognitive resources, which are usually associated with conscious reasoning. Again, however, the possibility that judgement can be intuitive is also entertained (Hammond & Stewart, 2001). Nevertheless, hypothesis testing in these tasks appears to be heavily biased and deterministic, leading to especially poor performance in probabilistic tasks that contain nonlinear cues. As with studies of expertise, participants are often unable verbally to report their judgement policies (Balke, Hammond, & Meyer, 1973; Hammond, 1971). If judgement in MCPL tasks is influenced only by conscious processing, however inaccurate their explicit knowledge, we would expect participants to be able verbally to report how they made their judgements. Poor levels of self-insight appear to confirm MCPL theorists' suspicions that judgement under some conditions is guided more by intuition than explicit reasoning (Evans et al., 2003; Hammond, 1996).

The implicit learning literature suggests that unconscious implicit processing controls behaviour in complex learning environments. Indeed, similar neural network models used to describe associative processing in implicit learning tasks could be applied to learning of multiple cues (Cleeremans, Destrebecqz, & Boyer, 1998; Dienes, 1992; Ganis & Schendan, 1992). This would involve training a neural network model on the cue-criterion relations via outcome feedback. However, proposing that judgement is instead influenced exclusively by associative processes fails to explain why certain cue-criterion relations are harder to learn. Positive cue learning should have no advantage over negative cue learning for an untrained neural network. Dual-process theories instead propose that both implicit associative and controlled hypothesis-testing processes can contribute to the inferences and judgements we make, and perhaps provide a better account of the cognitive processes involved in multiple cue judgement following training.

2.1.3. Dual process theories of MCPL

Egon Brunswik (1956) was first to become aware that inferences based on multiple cues may be approached cognitively in different ways. He suggested that some kinds of multiple cue judgements may draw on intuitive types of thought, whereas others may be approached more analytically. Brunswik seemed to be making a distinction between perceptual inferences on the one hand, and more effortful judgements on the other (Hammond, 2001). Brunswik's distinction between intuitive and analytic thinking was further developed by Kenneth Hammond as part of his Cognitive Continuum Theory (CCT). Hammond (1990; 1996; Hammond & Stewart, 2001) proposed that multiple cue judgement most likely draws on a blend of intuitive and analytic thinking. The extent to which judgement is intuitive or analytic depends primarily on a number of task characteristics, and to some extent on the cognitive characteristics of the individual. Tasks that are unfamiliar and highly complex are expected to draw heavily on analytic thinking and less on intuition. Such tasks can be placed at an analytic end of a continuum. Judgements that are made more intuitively would instead be placed at the opposite, 'intuitive', end of the continuum. However, CCT theory proposes that most tasks would be placed somewhere between the two poles, drawing on a mixture of both analytic and intuitive thinking.

CCT makes an interesting distinction between multiple cue judgement involving linear and nonlinear cues. It is proposed that while judgement in tasks containing nonlinear cues likely draws on analytic thinking, judgements based on linear cues are made more intuitively. That is, difficult nonlinear judgement tasks may require analytic thinking, whereas judgement in easier linear cue tasks can be guided more by intuition. While few studies have empirically tested CCT (Dunwoody, Haarbauer, Mahan, Marino, & Tang, 2000; Hammond, Hamm, Grassia, & Pearson, 1987), the theory makes an interesting distinction between intuitive and analytic thinking. However, Hammond made sure that although CCT distinguishes different kinds of thinking, he was not endorsing the idea that distinct and separate cognitive processes were intuitive or analytic, and criticised theories that propose separate cognitive processes (i.e. implicit and explicit) can compete to control behaviour (Hammond, 1966; 1996). CCT instead appears to distinguish between types of thinking styles, and makes no distinction between unconscious automatic and conscious controlled processing. For this reason CCT fails to take account of a large body of research (reviewed in Chapter 1) indicating that separate implicit and explicit processes can contribute to the inferences we make in a wide range of task environments.

Recent work by Jonathan Evans and his colleagues (Evans et al, 2003, 2005), indicates that both implicit and explicit processes may contribute to judgement in MCPL tasks. Following training, they measured performance on 40 test trials where no outcome feedback was presented, by correlating judgements with criterion values for each trial. They found that performance scores were poorer (but still above chance) following training in which cues were negatively related to the criterion, compared with positive cue tasks. They also asked participants, after completion of each task, to rate the relevance of each cue. The results showed clear dissociation between explicit knowledge and actual performance. On more difficult tasks, explicit knowledge was very poor even though performance was well above chance. Evans et al. concluded that explicit learning contributes to performance on easy but not difficult versions of the MCPL task, but that implicit learning is present on both.

2.2. RATIONALE OF THE THESIS

MCPL is traditionally viewed as a hypothesis testing task (Brehmer, 1974; 1980). This is despite the possibility that learning may instead be driven by implicit associative processes, similar to those shown to account for other types of skill learning (e.g. Cleeremans & McClelland, 1991). The MCPL paradigm has explored multiple cue judgement following training in controlled environments and suggests that participants' inferences are guided by knowledge acquired via hypotheses-testing. However, this is in contrast with studies of implicit learning that indicate automatic implicit processes can guide accurate inferences in complex environments. Dual-process theorists instead propose that both unconscious implicit and conscious explicit processes can compete to control behaviour.

Recently, Evans et al. (2003, 2005) have shown how judgements from multiple cues following training may indeed be influenced by both implicit and explicit processing. However, their findings rested solely on the relation between performance and participants' explicit beliefs about each cues' relevance. In the experimental studies of the present thesis we use dual-process methods not previously applied to MCPL tasks to explore the contribution of implicit and explicit processes to multiple cue judgement in controlled learning environments. Figure 2.2 provides an illustration of how implicit and explicit processes can compete to control judgements in MCPL tasks. Learning can, in principle, lead to both implicit and explicit knowledge. Outcome feedback provided in response to participants' judgements in a learning phase can foster both incremental learning through implicit associative processes and explicit hypothesis-testing. Implicit knowledge and explicit knowledge can then compete to control judgement following training (see Figure 2.2). Acquisition of expertise in a controlled environment using the MCPL paradigm allows the contribution of implicit and explicit processes to multiple cue judgement to be measured.

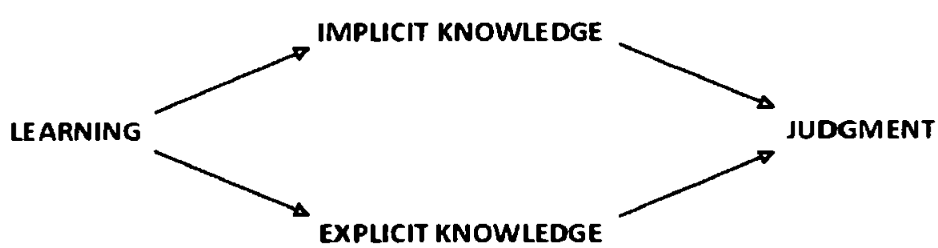

Figure 2.2: MCPL leads to the acquisition of both implicit and explicit knowledge. Both types of knowledge can compete to control judgement.

In the experimental studies of the present thesis we introduce dual process manipulations as well as measure individual differences to study implicit and explicit processes in multiple cue judgement. In Chapter 3, a concurrent working memory load is introduced to the test phase of the MCPL task after training to disrupt the contribution of explicit knowledge to judgement. While we expect a working memory load to inhibit explicit processing due to its dependence on limited working memory resources, the contribution of implicit knowledge to judgement should remain intact. We also introduce instructional manipulations. In Chapter 4, measures of individual differences in working memory capacity are correlated with performance and explicit knowledge scores. While attention is required for both implicit and explicit learning of complex tasks, explicit processing but not implicit processing is believed to load heavily on limited working memory resources. For this reason, correlations between working memory capacity and measures of performance and explicit knowledge can be diagnostic of explicit processing. The experimental studies of Chapter 5 introduce task manipulations designed to improve explicit learning, as well as monitor the contribution of implicit and explicit processes to learning and judgement using novel methods.

2.3. METHODOLOGY: THE MCPL JUDGEMENT TASK

The MCPL task used to train participants in the present thesis was based on Evans et al. (2003, 2005), and is similar to previous studies of MCPL. Participants were trained on three task types, each containing two relevant cues that were linearly related to the criterion, and two irrelevant cues. Positive-cue tasks contained two positive cues and two irrelevant cues (++00), negative-cue tasks contained two negative cues and two irrelevant cues (--00), and mixed cue tasks contained one positive, one negative, and two irrelevant cues (+-00). Relevant cues were maximally predictive of the criterion (before addition of noise), whereas irrelevant cues were entirely non-predictive. Criterion values were calculated using a linear model by entering cue weights as either 1 for positive cues, -1 for negative cues, or 0 for irrelevant cues. We followed the same procedure used by Evans et al. to calculate outcome feedback values. These were generated by adding a random variable from a Gaussian distribution to criterion values. This was done in order to add 25% noise to the outcome feedback participants received to simulate learning of expertise in uncertain environments. After adding the noise component task predictability was reduced to .87 (R^2 = .75). Criterion and feedback values were normalised to accommodate the ordinal scale used to display feedback to participants. On each trial, cue values for each cue were independently randomly generated, and so were theoretically uncorrelated.

Participants completed the MCPL task on a single computer, with up to five participants performing the task at one time. The task was divided into two sections. Participants first completed a learning phase designed to train them on the cue-criterion relations. In the learning phase outcome feedback was provided immediately in response to participants' judgements on each trial. Participants completed 80 learning trials in total. Our primary interest was to study expertise acquired in the learning phase. For this reason, participants completed a further 40 test trials immediately following completion of the learning phase. In the test phase outcome feedback was not provided in response to judgement. We used *performance* correlations as our measure of learning. These were calculated by correlating criterion values with the judgements individually for each participant.

We were also interested in measuring *explicit knowledge* levels. This provides a measure of how much knowledge participants had acquired explicitly about the task, as well as an indication of the extent to which participants were explicitly engaged in the task. Participants' explicit beliefs used to calculate explicit knowledge scores were measured using a cue rating task presented immediately following completion of the test phase. In the cue rating task participants were again presented with the four cues and asked to rate the relevance of each. In Experiments 1 and 2, they were asked to rate whether each was positively, negatively, or unrelated to the criterion, with a confidence rating for each. To the right of the tests were the three labels *positive, irrelevant,* and *negative*, which could be selected by moving the curser over a label and clicking the mouse. To the right of these labels the confidence levels were presented, ranging from *1* to 5, with *1* labeled as *not very confident* and 5 labeled as *very confident*. In Experiments 3-8, participants instead rated each cue on a continuous scale. They did this by moving a slider presented on the screen either towards a *positive* label or a *negative* label. The slider appeared initially in the middle, below an *irrelevant* label, for each cue. Participants did not give a confidence rating for each cue. We calculated explicit knowledge scores by modelling how participants would have performed based on their explicit beliefs alone. This was achieved by using participants' explicit beliefs about each cue's relevance as cue weights to make predictions on each trial using the linear additive rule:

$$J = x_1C_1 + x_2C_2 + x_3C_3 + x_4C_4$$

where predicted judgement, J, for each participant on each trial is the product of the sum of each cue value, C, on each trial multiplied by the participant's explicit rating of each cue, x (Einhorn et al, 1979; Juslin et al, 2003). In Experiments 1 and 2 we entered participants post-task cue ratings (0 for irrelevant, 1 for positive, -1 for negative) and confidence in each rating, ranging from 1 (not very confident) to 5 (very confident). For example, if a participant rated a cue as positive with a confidence level of 4, a rating of +4 was used, whereas if a participant rated a cue as negative with the same confidence, a rating of -4 was used. In Experiments 3-8, we used participants' explicit ratings of each cue on a continuous scale ranging from 1 to -1. The predicted judgements based on the participant's explicit beliefs about each cue were then correlated with criterion values to give the approximate performance score they could have achieved, had they used only their explicit beliefs about each cue.

In all experimental studies except Experiment 3, participants were first provided instructions explaining that it was their task to rate the suitability of a sample of job applicants one at a time, using four personality tests (the cues), which could predict applicant suitability positively, negatively, or not at all. In Experiment 3, participants instead learned to predict house prices based on four attributes of houses. Participants were provided with definitions of positive, negative, and irrelevant relations, with an example of each. The instructions were provided on paper, allowing participants to refer back to the instructions at any time during the task. The procedure for each trial consisted of presenting the four tests on screen in list form, as shown below, with the current applicant's results to the right-hand side of the corresponding tests. The results ranged from *very low, low, average, high,* to *very high.*

Test A: High

Test B: Very Low

Test C: Average

Test D: Low

How suitable is the job applicant?

Very Poor   Poor   Average   Good   Very Good

Below the tests, five corresponding levels of suitability, from *very poor, poor, average, good,* to *very good,* were presented. These could be selected via a mouse click. In the learning phase, following a selection the selected label remained highlighted and below appeared the feedback, referred to as the actual suitability of the applicant. The test phase followed the same procedure, except participants moved immediately on to the next trial after making each judgement without receiving feedback. Following completion of the test phase, participants completed the cue rating task. They were then debriefed and thanked for their participation.

2.4. GENERAL HYPOTHESES AND RESEARCH QUESTIONS

In Chapter 1 (see section 1.2.), I discussed a wealth of research indicating that much of human learning, judgement, and reasoning is influenced by knowledge that is implicit and inaccessible to verbal report. A puzzling finding in the judgement literature is that experts who take account of multiple cues often are not able verbally to describe how they make their judgements. One possibility is that experts are, in part, influenced by knowledge that is implicit. However, this possibility has received little attention in the multiple cue judgement literature. The experimental studies of the present thesis are designed to examine the role of implicit and explicit processes in judgement.

In Chapter 3, the contribution of explicit knowledge to judgement is disrupted by the addition of a concurrent working memory load to test trials after training. If performance on judgement tasks is unaffected under these conditions then this would suggest that judgement is intuitive and guided by knowledge that is implicit. However, this does not rule out the possibility that people engage in deliberative hypothesis-testing during learning. By measuring individual differences in working memory capacity in Chapter 4, we explore whether explicit processing is involved at any stage of training. Hence, there are two key questions of the present thesis. Our first question concerns whether multiple cue judgement should be viewed as an entirely explicit process, or whether both implicit and explicit knowledge is involved. Our second question concerns the role of these processes during the acquisition of expertise. These questions are explored further in Chapter 5.

Chapter 3

The effects of concurrent working memory load and instructional manipulation on multiple cue judgement

In the present chapter I report three experiments designed to explore the role of implicit and explicit processes in judgement from multiple cues. In Experiments 1 and 2 we take a direct approach by introducing a concurrent working memory load during the test trials, when judgements are tested *after* learning from outcome feedback. We did not load working memory in the learning phase, as a great deal of research suggests working memory (or attention) is required for implicit (as well as explicit) learning of complex tasks (e.g. Nissen & Bullemer, 1987; Hayes & Broadbent, 1988). However, we can expect that use of a working memory load in the test phase will selectively interfere with judgements that are mediated by explicit rule-based processes whilst leaving implicit judgemental processes relatively unimpaired (Curren & Keele, 1993; De Neys, 2006; Hayes & Broadbent, 1988). This method should enable us to tell whether such processes are engaged in the MCPL task.

In tasks that draw on effortful explicit processing, we expect performance to be poorer with the addition of a concurrent working memory load to trials in a test phase. An assumption made by some MCPL theorists appears to be that participants combine cue values on each trial with their explicit beliefs of each cues' relevance when making judgements (Brehmer, 1980; Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Juslin, Olsson, & Olsson, 2003; See also Lagnado, Newell, Kahan, & Shanks, 2006 for a theory of MCPL using a different version of the task). For instance, Einhorn et al. (1979) showed how judgement in the MCPL task can be modelled by computer algorithms based on participants' ongoing verbal protocols. Participants verbalised their reasoning at every stage of judgement, allowing their dynamic judgement policy to be formalised. However, such approaches would not pick up on implicit processing, which may also contribute to judgement. Without seriously considering the role of implicit knowledge the assumption among some MCPL theorists appears to be that people's inferences are made entirely (or mostly) explicitly. In this way, judgement is effortful and demanding on the individual's limited working memory resources, even after learning. We expect, therefore, that this process should be disrupted by loading participants' working memory resources with a secondary task. In other tasks, learning of complex rules is often shown to occur implicitly, without any evidence that explicit processing contributes to participants' improved performance (Berry & Broadbent, 1984; Gluck, Shohamy, & Myers, 2002; Price, 2005; Reber, 1993). When performance on a task is driven by implicit processing, the addition of concurrent load tasks designed to disrupt explicit processing have no effects on performance (Hayes & Broadbent, 1988; Waldron & Ashby, 2001). For this reason, we expect that the addition of a working memory load to MCPL tasks will have no effect on performance after learning if judgements are made implicitly.

In Experiment 3 we use an alternative method to measure implicit and explicit processes in multiple cue judgement. We examine whether task instructions have any effect on performance and explicit knowledge following training. If people use effortful explicit processing to make judgements, then these should be influenced by task instructions. In contrast, instructions should have no influence if the inferences participants make are guided solely by implicit processes.

3.1. EXPERIMENT 1

If performance in MCPL tasks relies predominantly on explicit knowledge, as some theorists claim, then the ability to make judgements under working memory load should be severely impaired. If both implicit and explicit processes contribute to judgements, as claimed by Evans, Clibbens, Cattani, Harris, and Dennis (2003), then the implicit component of performance should be preserved. As a second check on the role of explicit knowledge in this task, we also model the performance that participants would achieve given their post-learning explicit beliefs and compare that with actual performance.

3.1.1. Method

Participants

Eighty undergraduate students at the University of Plymouth participated for course credits.

Design

Participants were divided into two groups according to whether they received two positive and two irrelevant cues (++00; positive cue task) or one positive, one negative, and two irrelevant cues (+-00; mixed cue task). Each group was further divided into two groups according to whether they received a working memory load task (or not) on the test trials, generating four independent groups in total. Details of the task methodology and procedure are provided in section 2.3. The only exception to the task methodology in Experiment 1 is the addition of the concurrent working memory load task to the test phase.

Procedure

Each participant completed the judgement task in one of the four conditions lasting up to 30 minutes. Instructions for participants are provided in section 7.3.

Working memory load task. Participants in the load condition performed a simultaneous visual memory load task during test trials. Once the participant had completed the learning trials they were presented with instructions on screen explaining on each trial they would need to hold in mind the location of four dots in a 3x3 grid. The location of each dot was randomly generated. A grid then appeared on screen for 500 ms, after which they were required to judge an applicant's suitability, without receiving feedback. Immediately after making their selection, an empty grid appeared with which they were required to select the appropriate cells of the grid via a mouse click corresponding to the locations of the dots for each trial. They received a new grid pattern for each of the forty test trials.

3.1.2. Results and Discussion

Working Memory Load

Participants achieved an overall average of 3.51 (SD = 0.41) correct dot placements out of a possible 4 across participants. Load task performance (number of correctly recalled dot placements) did not correlate significantly with judgement task performance ($r = .002$, $n = 40$, $p = .992$).

Performance in the test phase

Performance correlations in the test phase were calculated by correlating participants' judgements with criterion values across the 40 test trials. We also calculated performance scores in the 1st 40 and 2nd 40 learning trials in the same way. Analyses of performance scores in learning phases of the experiments in the present chapter are reported in section 7.1. Performance scores in the test phase for each group are presented in Table 3.1. All group performance scores were significantly above zero, except for those performing the +-00 task in the no load condition.

Table 3.1

Mean group performance scores

	++00		+-00		--00	
	M	t	M	t	M	t
Experiment 1						
Load	.66*	18.076	.24*	3.063		
No load	.60*	18.267	.10	1.318		
Experiment 2						
Load	.52*	8.538	.17*	2.372	.44*	5.687
No load	.51*	9.568	.10	1.776	.29*	2.871
Experiment 3						
'High' instructions	.61*	14.268	.30*	3.694	.44*	6.061
'Low' instructions	.62*	12.072	.22*	3.512	.38*	4.708

* p <.05

In order to examine the effects of the concurrent memory load task on performance, a two-way independent analysis of variance (ANOVA) was performed using load-type (no load or load) and task-type (++00 or +-00) as independent factors, and performance as the dependent variable. A Fisher *z* transformation was applied to performance scores to improve sample distributions in this analysis and all further analysis involving performance scores. In contrast with our predictions, judgement was unaffected by the addition of the concurrent load task in the test phase ($F_{(1,76)} = 3.143$, $MSE = .448$, $p = .080$, partial $\eta^2 = .04$). We expected that loading participants' working memory resources would worsen their performance on judgement tasks that require effortful explicit processing. However, in both task types judgement appeared unaffected. This suggests that judgement is guided solely by implicit processing, at least when performance is measured after learning. Consistent with previous research (Evans et al, 2003; 2005), there was, however, an effect of task-type ($F_{(1,76)} = 45.917$, $MSE = 6.549$, $p < .001$, partial $\eta^2 = .38$), as the positive cue set (mean score = .63) was learned much better than the mixed cue set (.17), which included a negative predictor. There were no significant interactions.

Explicit knowledge

We measured explicit knowledge scores by modelling how participants would perform in the test trials based on their explicit beliefs alone. Details of how explicit knowledge scores were calculated are provided in section 2.3. All group explicit knowledge scores were significantly above zero. Mean group explicit knowledge scores are displayed in Table 3.2.

Table 3.2

Mean group explicit knowledge scores

	++00		+-00		--00	
	M	t	M	t	M	t
Experiment 1						
Load	.33*	2.278	.28*	2.558		
No load	.47*	5.392	.33*	3.040		
Experiment 2						
Load	.42*	3.682	.29*	2.319	.19	1.919
No load	.18	1.607	.35*	3.357	.21	1.703
Experiment 3						
'High' instructions	.41*	3.534	.25	1.709	.24	1.831
'Low' instructions	.31*	2.480	.19	1.434	.27*	2.214

* $p < .05$

A two-way independent ANOVA was performed on explicit knowledge scores, using task-type (++00 or +-00) and load-type (no load or load) as independent factors, and explicit knowledge scores as the dependent variable. A Fisher z transformation was applied to explicit knowledge scores in this analysis and all following analysis involving explicit knowledge scores. There was no effect of task-type ($F_{(1,76)}$ = .752, *MSE* = .499, *p* = .389, partial η^2 = .01), and although participants appeared generally to acquire less explicit knowledge in load conditions (.31) than no load conditions (.40), this difference was not significant ($F_{(1,76)}$ = .304, *MSE* = .202, *p* = .583, partial η^2 < .01). There were also no significant interactions. Hence, in contrast with performance scores, participants acquired roughly the same amount of explicit knowledge in the two task types. Furthermore, participants' explicit knowledge of the ++00 tasks appeared below their actual performance, whereas the converse was true for the +-00 task (see Panel A: Figure 3.1). Thus we performed a further ANOVA, in which we entered explicit knowledge and performance scores for each participant on each task as an additional factor, which we call 'measure'. There was no main effect of measure ($F_{(1,78)}$ = .236, *MSE* = .060, *p* = .628, partial η^2 < .01). However, there was an effect of task-type ($F_{(1,78)}$ = 11.371, *MSE* = 5.825, *p* = .001, partial η^2 = .13), and a cross-over interaction between measure and task-type ($F_{(1,78)}$ = 7.822, *MSE* = 2.002, *p* = .006, partial η^2 = .09). Pairwise comparisons confirmed that participants performing the ++00 task achieved a significantly higher performance score than that predicted by explicit knowledge (t = 2.885, df = 39, *p* = .006). Although participants performing the +-00 task appeared to show the reverse effect, this difference did not reach significance (t = -1.884, df = 39, *p* = .067).

In sum, while performance was poorer when tasks contained a negative cue than when both relevant cues were positive, similar levels of explicit knowledge were acquired of both task types. The lack of any effects of working memory load on performance indicates that judgement may be guided solely by implicit learning processes. If judgement in the test phase was guided by working memory dependent explicit processing then we would have expected the concurrent load task to disrupt performance. However, this conclusion is at odds with our analysis comparing performance and explicit knowledge levels. Following positive cue training participants performed far better than expected based on their explicit beliefs alone, suggesting that a substantial implicit component is involved in learning tasks that contain only positive relevant cues. The extent to which implicit processes are involved in learning tasks containing a negative cue is less clear, however. In mixed cue tasks performance appeared slightly lower than explicit knowledge levels. If these tasks are also learned implicitly we would have expected performance levels to again exceed explicit knowledge scores.

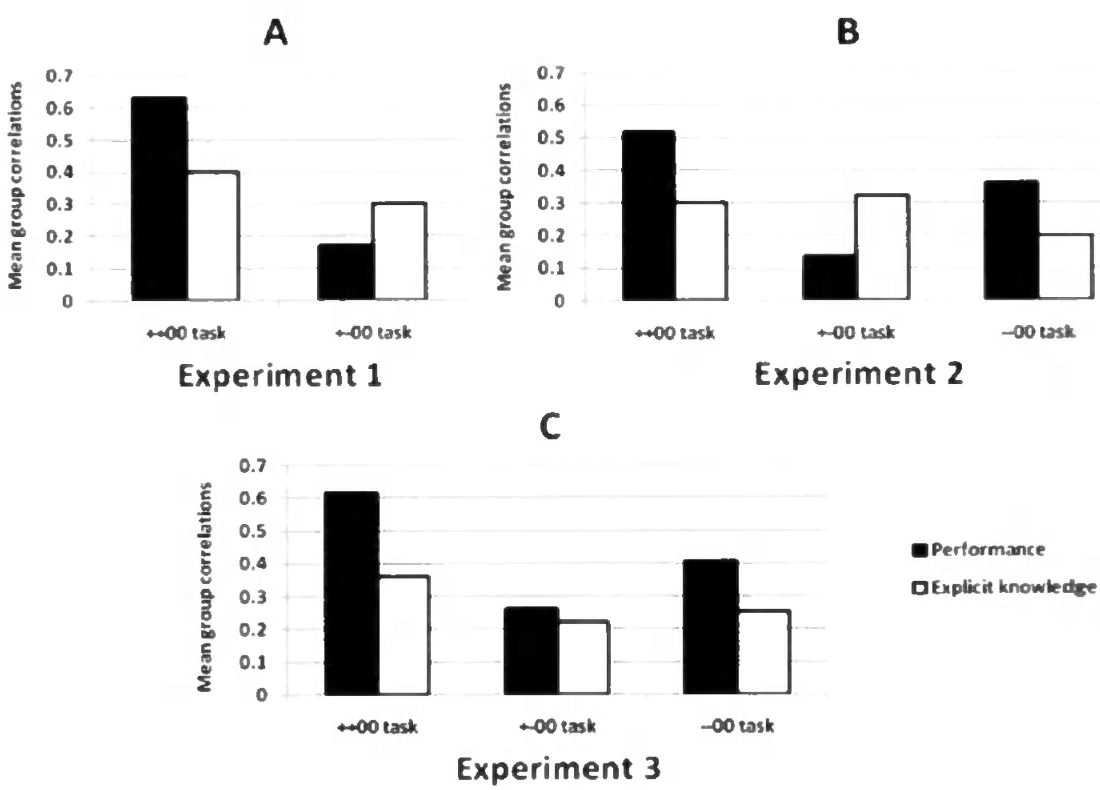

Figure 3.1: Comparison between mean group performance and explicit knowledge scores in Experiments 1, 2, and 3.

3.2. EXPERIMENT 2

In Experiment 1, there was no evidence that working memory load interfered with judgements made after the learning phase. This finding suggests that multiple cue judgement may not depend upon explicit processing, and is in contrast with the proposal of Evans et al. (2003) that both implicit and explicit processes contribute to the inferences that people make. In addition, post-task ratings revealed that while participants had acquired explicit knowledge of positive and mixed cue sets more or less equally, their performance on the two tasks was very different. The working memory load task was, however, of a visual nature. Although this kind of task has been shown to interfere with deductive reasoning (De Neys, 2006) it is arguable that a verbal working memory load task would be more disruptive. This is because the MCPL task presents the participant with verbal cues and cue values, and requires a verbal judgement of applicant suitability to be made on each trial. For this reason, in Experiment 2 we instead used a verbal load task. We also took the opportunity to apply our methods to a third task, in which two negative cues are combined with two that are irrelevant. If explicit processing contributes to judgement in tasks containing negative cues, we should find that loading working memory disrupts performance on these tasks.

3.2.1. Method

Participants

One hundred and twenty undergraduate students at the University of Plymouth participated for course credits.

Design

We used the same general method for the judgement task, except for the addition of a task containing two negative and two irrelevant cues (--00; negative cue task), and the use of a verbal working memory load on test trials instead of the visual load used in Experiment 1. This generated six independent groups.

Procedure

The same general procedure was used (see section 7.3. for task instructions), with the exception that participants in the verbal load condition were required to memorize a list of six digits and letters on each of the test trials rather than the location of dots in a grid. The digits used were integers from 1 to 9, and the letters were B, C, D, F, G, H, J, K, and L. All digits and letters were spoken and recorded on computer and then joined according to their randomly generated ordering, generating forty individual memory lists of six elements. Participants were informed they would be required to repeat each list in order in their head, whilst making their judgement. After making their judgement an empty box appeared on screen into which they were required to type the list.

3.2.2. Results and Discussion

Working memory load

Overall, participants correctly recalled an average of 3.90 (SD = 0.94) digits out of a possible 6. Load task performance (number of correctly recalled digits) did not correlate significantly with judgement task performance (r = .225, n = 60, p = .084).

Performance in the test phase

Mean group performance correlations in the test phase are displayed in Table 3.1. All group performance scores were significantly above zero, except for those learning the +-00 task in the no load condition. A two-way independent ANOVA was performed using working memory load (no load or load) and task-type (++00, +-00, or --00) as independent factors, and performance scores as the dependent variable. As in Experiment 1, there was no effect of working memory load on performance ($F_{(1,114)}$ = 2.760, MSE = .502, p = .099, partial η^2 = .03), but there was a highly significant effect of task-type ($F_{(2,114)}$ = 12.621, MSE = 2.295, p < .001, partial η^2 = .18). Participants performed best on the positive (++00) task, worst on the mixed cue (+-00) task, with the negative cues set (--00) intermediate. Independent t-tests confirmed significant differences between the ++00 (.52) task and --00 task (.36; t = 2.049, df = 78, p = .044), and between the --00 task and +-00 task (.14; t = 2.894, df = 78, p = .005). There were no significant interactions. These results support our conclusions of Experiment 1 that working memory dependent explicit processing contributes little to judgement following training, regardless of the types of cues on which participants are trained. However, the type of task participants perform does influence their learning. Performance appears poorest in mixed cue tasks containing both positive and negative relevant cues, and best when both relevant cues are positive. Performance is intermediate when both relevant cues are negative. Analyses of performance scores in learning phases are reported in section 7.1.

Explicit knowledge

Explicit knowledge scores were calculated in the same way as in Experiment 1, and are displayed in Table 3.2. Three of the six groups acquired explicit knowledge that was significantly above zero. Participants performing the ++00 task in the no load condition, along with both groups performing the --00 task, failed to achieve significant explicit knowledge.

A two-way independent ANOVA was performed, using task-type (++00, +-00, or --00) and load (no load or load) as independent factors, and explicit knowledge scores as the dependent variable. There was no effect of task-type ($F_{(2,114)}$ = .993, MSE = .544, p = .374, partial η^2 = .02), or load ($F_{(1,114)}$ = .355, MSE = .195, p = .552, partial η^2 < .01), and no significant interactions. Hence, explicit knowledge acquisition was broadly independent of the type of cue set on which people were trained, even with the addition of the negative cue task (--00).

As in Experiment 1, a three-way mixed ANOVA was also performed, using task-type (++00, +-00, or --00) and load-type (load or no load) as independent factors, measure (performance vs explicit knowledge) as the within-subjects factor, and correlations as the dependent variable. There was no significant effect of task-type ($F_{(1,117)}$ = 2.596, MSE =

1.212, p = .079, partial η^2 < .04), or measure ($F_{(1,117)}$ = .349, MSE = .092, p = .556, partial η^2 < .01). However, there was a cross-over interaction between measure and task-type ($F_{(2,117)}$ = 6.180, MSE = 1.627, p = .003, partial η^2 < .10), illustrated in Panel B of Figure 3.1. As in Experiment 1, participants performed better than the level predicted by explicit knowledge for the positive cue set and worse for the mixed set. On the newly added negative cue sets, performance again exceeded that predicted by explicit knowledge. Pairwise comparisons confirmed the reliability of the three trends. Whereas participants performed significantly better compared with their explicit knowledge in the ++00 task (t = 2.480, df = 39, p = .018), and --00 task (t = 2.039, df = 39, p = .048), they performed significantly worse in the +-00 task than their explicit knowledge (t = -2.555, df = 39, p = .015).

In sum, Experiment 2 confirmed the findings of Experiment 1 that working memory load has no effect on performance after learning, using a verbal rather than visual load task. Interestingly, Experiment 2 also supports our finding that participants perform better than predicted based on explicit knowledge alone in positive cue tasks, but perform worse compared with their explicit knowledge in mixed cue tasks. With the addition of the --00 task in Experiment 2 it was also shown that participants outperformed their explicit knowledge scores in tasks containing only negative relevant cues. The lack of an effect of a concurrent load task on performance in the test phase suggests that explicit knowledge contributes little to judgement once expertise has been acquired. One possibility is that explicit hypothesis-testing contributes to learning only in early stages of the training phase and that explicit knowledge becomes automated through practice. If this is the case, then judgement should draw less on working memory resources once explicit knowledge has become automated, and would be less affected by the addition of a working memory load to the test phase. However, if explicit processing is involved in the learning of expertise then instructional manipulations should influence performance. In Experiment 3, we attempt to direct participants' attention to negative cues as a means of testing for the role of explicit processing in learning.

3.3. EXPERIMENT 3

In Experiment 3 we explored the effects of instructions on participants' judgements. We used a scenario in which participants must learn which cues predict house price values. It was predicted that participants may be more likely to consider negative cues when they are asked to look for low criterion values, rather than high criterion values. This would therefore affect participants' explicit hypothesis testing without affecting implicit processes that respond to cue-feedback associations alone. In a post-task questionnaire, participants were also asked to list as many examples of positive and negative cues in their environment as they could, in order to examine the effects of instructions on the relative number of cues they consciously bring to mind. Previous research (Brehmer, 1974, 1980) suggests participants perform better with positive cues than negative cues because examples of positive cues come to mind more easily, and are perceived as occurring with a higher frequency in the environment. Therefore, we were interested in whether the relative number of examples of positive and negative cues listed by participants would relate to their performance. It may be that when participants are required to predict low criterion values they instead think of more negative than positive cues, and accordingly show improved performance on tasks containing negative cues. However, in line with the findings of Experiments 1 and 2, this may have no effect on performance if learning is largely implicit.

3.3.1. Method

Participants

One hundred and twenty undergraduate students at the University of Plymouth participated for course credits.

Design

The same general method was used, except in contrast with Experiments 1 and 2 no groups completed a concurrent working memory load task. Instead, an instructional manipulation was added prior to participants beginning the judgement task. In Experiment 3, the cue-rating task used to measure participants' explicit beliefs about each cue was different to Experiments 1 and 2. Details of how explicit beliefs were measured using the cue-rating task are provided in section 2.3. Each of the three task groups (++00 task, +-00 task, and --00 task) were either asked to look for high criterion values or low criterion values, generating a total of six independent groups. All participants then completed a post-task questionnaire.

Procedure

The general task and instructions were used, except that instead of rating the suitability of job applicants participants rated the price level of a sample of houses from *very low* to *very high*, based on four features, labelled **Feature A, Feature B, Feature C,** and **Feature D,** which could vary in cue value on a 5 point scale from *very low* to *very high* (see section 7.3. for task instructions). Task instructions explained that each feature could represent some attribute about houses or their local area, and similarly to Experiments 1 and 2, could be positively, negatively, or not at all predictive (irrelevant). One group was told to imagine they worked for an executive estate agent that only sold expensive houses, and so their job was to predict 'high' house prices. A second group was told to imagine they were buying a house and could only afford a cheap house, so their job was to predict 'low' house prices. Once participants had completed the judgement task, those told to look for 'high' house prices were asked to list as many attributes related to houses and their local area as they could, that may positively predict high house prices, and those that may negatively predict high house prices. Those participants told to look for 'low' house prices were instead asked to list as many attributes as they could that may positively predict 'low' prices, and those that may negatively predict 'low' house prices.

3.3.2. Results and Discussion

Performance in the test phase

Mean group performance scores are displayed in Table 3.1. All group performance scores were above zero and significant. We performed a two-way independent ANOVA on performance scores in test trials, using task-type (++00, +-00, and --00) and instruction-type ('high' or 'low') as independent factors. There was a significant effect of task-type ($F_{(2,114)}$ = 14.845, MSE = 2.432, p <.001, partial η^2 = .207), with participants performing significantly better in the ++00 (.62) than --00 task (.41; $F_{(1,76)}$ = 10.913, MSE = 1.885, p =.001, partial η^2 = .126), and marginally significantly better in the --00 than the +-00 task (.26; $F_{(1,76)}$ = 3.810, MSE = .653, p =.055, partial η^2 = .048). However, there was no effect of the instructional manipulation ($F_{(1,114)}$ = .531, MSE = .087, p = .468, partial η^2 = .005), nor any significant interactions. Our finding that the instructional manipulation had no effect of performance following training suggests that either our manipulation was not sufficient to direct participants' attention to the negative cues, or that explicit hypotheses-testing processes contribute little to judgement even during the acquisition of expertise. While performance is strongly affected by the presence of one or more negative cues, attempting to direct participants' attention to the presence of these cue types does not appear to influence explicit hypotheses testing.

Explicit knowledge

Mean group explicit knowledge scores are displayed in Table 3.2. Only three of the six groups acquired explicit knowledge of the task that was significantly above zero. Both groups that were trained on the +-00 task, along with participants performing the --00 task under 'high' instructions failed to achieve significant levels of explicit knowledge.

A two-way independent ANOVA using task-type (++00, +-00, or --00) and instruction-type ('high' or 'low') as independent factors, was performed on explicit knowledge scores. This yielded no significant effects of task-type ($F_{(2,114)}$ = 1.729, MSE = 1.717, p = .182, partial η^2 = .029), instruction-type ($F_{(1,114)}$ = .238, MSE = .236, p = .627, partial η^2 = .002), or any significant interactions. Therefore, participants' explicit knowledge of the multiple cue tasks was not affected by the types of cues they learned, or by an instructional manipulation designed to improve their explicit learning of negative cues. This indicates that our instructional manipulation did not influence explicit learning of the tasks.

In order to compare performance and explicit knowledge scores across the different task types, we performed a three-way mixed ANOVA using task-type and instruction-type as independent factors, and measure (performance in the test trials and explicit knowledge) as a within-subjects factor. There was no significant effect of instruction-type ($F_{(1,114)}$ = .446, MSE = .305, p = .505, partial η^2 = .004), or measure ($F_{(1,114)}$ = 2.986, MSE = 1.414, p = .087, partial η^2 = .026), but a significant effect of task-type ($F_{(2,114)}$ = 5.840, MSE = 3.992, p = .004, partial η^2 = .093). Surprisingly there was no significant interaction between task-type and measure ($F_{(2,114)}$ = .332, MSE = .157, p = .718, partial η^2 = .006). Consistent with Experiments 1 and 2, however, in the ++00 task performance scores (.62) did appear substantially higher than explicit knowledge scores (.36), and similarly in the --00 task, performance (.41) also appeared higher than levels of explicit knowledge (.25). In contrast, in the +-00 task performance (.26) and explicit knowledge (.22) scores appeared similar (see Panel C: Figure 3.1). There were no other significant interactions. These trends are consistent with our findings across Experiments 1 and 2, that performance levels exceed what is expected based on explicit knowledge alone in positive cue and negative cue tasks.

Performance scores were again similar to explicit knowledge scores in mixed cue tasks, suggesting that implicit processing is not contributing to learning on these tasks.

Post-task questionnaire

The mean number of positive and negative cue examples listed by participants in the post-task questionnaire for each group is presented in Table 3.3. We performed a two-way independent ANOVA to explore the effects of task instructions on the number of positive and negative cue examples listed by participants. For this, we used task-type (++00, +-00, or --00) and instruction-type ('high' or 'low') as independent factors, and mean difference in number of positive and negative cue examples listed as the dependent variable. There was a significant main effect of instruction type ($F_{(1,114)}$= 22.580, MS_e= 243.675, p <.001, partial η^2 = .165), with participants under 'high' instructions listing more examples of positive cues than negative cues (mean difference = 2.9), compared with participants under 'low' instructions (.02). There was no effect of task type ($F_{(2,114)}$= 2.855, MS_e= 30.808, p =.062, partial η^2 = .048), or any significant interactions. Hence, instructions affected the number of examples listed by participants without appearing to affect their performance on the judgement task. This suggests that asking participants which cues in the environment predict low house prices eliminates their conscious bias towards positive cues, whereas asking them to *predict* low house prices in the judgement task has no effect on their performance with negative cues. When participants were asked to list examples of cues in the environment that predicted low criterion (house price) values, they listed similar numbers of positive and negative cues in the post-task questionnaire, suggesting that their conscious bias towards expecting negative cues was eliminated. In contrast, learning still appeared biased towards positive cues when participants were asked to predict low criterion values, with higher performance scores in positive cue tasks than tasks containing one or more negative cues. This suggests that the conscious explicit processing used by participants to list examples of cues in the environment may not have been involved in learning in the judgement tasks.

Table 3.3

Experiment 3: Mean number of examples of positive and negative cues listed in post-task questionna

	++00		+-00		--00	
	M	t	M	t	M	t
'High' instructions						
Positive cues	8.4	9.505	11.1	12.672	11.8	11.101
Negative cues	6.3	6.232	8.7	9.032	7.6	7.612
'Low' instructions						
Positive cues	9.0	7.738	7.2	7.834	6.8	10.780
Negative cues	9.0	7.522	7.9	7.907	6.1	8.767

3.4. GENERAL DISCUSSION

The experimental studies of the present chapter were designed to investigate the role of implicit and explicit processes in learning of multiple cues and subsequent judgements. The introduction of concurrent working memory load tasks in Experiments 1 and 2 was expected to disrupt performance either drastically (on the typical view that MCPL requires explicit hypothesis testing) or partially and selectively, according to the proposal of Evans et al. (2003) that both implicit and explicit learning contributes to performance on the task. In fact, there was no evidence in either experiment that working memory load interfered with performance at all. However, our second method of investigating the question also produced an intriguing result. In both experiments, participants acquired sufficient explicit knowledge to have performed most of the tasks above chance, but with no evidence that the level of explicit knowledge depended on the cue set learned. This shows a very strong dissociation between explicit knowledge and performance, in that the task type substantially and significantly affected performance levels but not explicit knowledge. However, our evidence that participants acquire similar levels of explicit knowledge in each task type is inconsistent with those reported by Evans et al. (2003). They found that participants did not acquire significant levels of explicit knowledge when a negative cue was introduced to judgement tasks.

The current findings seem consistent with two possible conclusions. First, it may be that the explicit knowledge acquired did contribute to performance, in spite of the lack of disruption by concurrent working memory load. That is, our load task may not have sufficiently loaded participants working memory resources to disrupt explicit judgement. This is a possibility; especially in Experiment 1 where performance on the load task was close to ceiling (3.51 dot locations were recalled on average out of a possible 4). If the load task imposed in Experiment 1 was simply too easy then participants may have had enough working memory resources left over to perform the judgement task explicitly. However, this is less likely to be the case in Experiment 2 where the verbal load imposed on participants was clearly demanding (3.90 digits were recalled on average out of a possible 6). Despite this, we can not rule out the possibility that participants were allocating explicit effort to both the load task and judgement task.

Recall that if multiple cue judgement is explicit then this should load heavily on working memory resources. Participants must combine the values of four ordinal cues with their subjective beliefs about their importance for explicit judgment. Hence, we should have expected working memory load to disrupt performance drastically. A second possibility is that although participants acquired explicit knowledge during learning, this knowledge was not used to generate judgements in the test trials. A popular view in the learning literature is that working memory dependent explicit processes may be involved initially to acquire task knowledge which, as skills are rehearsed, becomes automated, drawing less on explicit processes (Anderson et al, 2004; Logan, 1988). Since the tasks in the present research presented participants with 80 trials from which to learn, it is likely that learning in multiple cue tasks is initially assisted by explicit processing, but becomes automated prior to the test trials. This would account for the lack of effect of working memory load on those trials. However, as a cautionary note, these conclusions are based on a null effect of load in Experiments 1 and 2 and are only speculative. We cannot rule out the possibility that our load tasks did not disrupt explicit judgement.

We found that participants performed better than expected in positive cue tasks based on their explicit knowledge levels, suggesting that a substantial implicit component is involved in learning tasks that contain only positive relevant cues. While the findings of Experiment 3 are less clear, the reverse effect was shown in mixed cue tasks containing a negative cue. Implicit learning processes do not appear to boost performance levels above what is attainable by explicit knowledge in these tasks. Hence, it seems that implicit processing may contribute less to learning of tasks that contain a negative cue. We can explain this in a dual-process learning model where the former can be learned implicitly, but the latter requires explicit intervention to learn the negative cue. However, the picture is complicated by the further observation that when *both* relevant cues are negative, performance once again exceeds explicit knowledge. When both cues are negative participants may learn quickly to simply reverse the feedback given, which becomes quickly automated. The idea that explicit processing is required to learn negative, but not positive, cues is consistent with studies showing that explicit processes are required for negating the values of stimuli (Deustch, Gawronski, & Strack, 2006), and providing verbally incongruent responses to stimuli (Kornblum, Hasbroucq, & Osman, 1990).

Our conclusions that explicit processing may contribute to MCPL early in training are inconsistent, however, with the findings of Experiment 3. In Experiment 3 participants were instead instructed to predict either 'high' criterion values or 'low' criterion values. Our rationale for this manipulation was that by asking participants to predict low criterion values their attention would be directed towards the presence of negative cues. Previous research has shown participants to list more examples of positive than negative cues in post-task questionnaires, and accordingly perform better in positive than negative cue tasks (Brehmer, 1974, 1980; Brehmer & Kuylenstierna, 1978). This indicates that participants may be biased in their explicit hypothesis-testing towards expecting positive cues. While instructing participants to predict low criterion values eliminated their conscious bias towards positive cues in a post-task questionnaire, we found that their performance in the MCPL task was unaffected. If learning from multiple cues draws on explicit processing in tasks containing negative cues we would have expected some effect of task instructions on performance. Therefore, learning and judgement may in all task types be guided solely by implicit processing. An alternative explanation is that participants' attention was indeed directed towards thinking of negative cues, but that this effect was not sufficient to influence their explicit learning of negative cues in the judgement tasks.

A number of factors may contribute to participants' difficulty in learning negative cues. Participants tend to assume that cues are deterministically related to the criterion, and continue to perform poorly in some tasks even when told that the same cue values can lead to different outcomes due to the probabilistic nature of the task (Brehmer & Kuylesntierna, 1978). Methods designed to make each cue-criterion relation more salient, such as varying the values of one cue at a time, are also shown to have little effect of learning in MCPL tasks (Hoffman, Earle, & Slovic, 1981). Hence, even if participants approach learning by consciously testing hypotheses against the feedback they receive, directing their attention to specific cue-criterion relations can have little effect on learning of probabilistic multiple cue tasks. Our failure in the present chapter to improve explicit learning by directing participants' attention to the presence of negative cues may be due to participants' difficult in testing hypotheses against feedback in probabilistic tasks. In this way, explicit processes may be involved in initial learning of MCPL tasks, despite the lack of effect of task instructions.

The present chapter provides some indication that both implicit and explicit processes do indeed contribute to learning in multiple cue environments. While it appears that implicit processes contribute heavily to learning of positive cue tasks, it is less clear the extent to which explicit processes are involved in learning tasks that contain negative cues. Our failure to find any effect of concurrent working memory load on judgement does allow us to speculate, however, that explicit processing may not be involved in multiple cue judgement following sufficient practice. The findings of the present chapter stand in contrast with previous research on MCPL. It is commonly assumed that judgement in multiple cue tasks, even after learning, is driven by deliberative explicit processes (Juslin et al, 2003; Lagnado et al, 2006). According to these theorists, participants consciously make judgements by combining their explicit beliefs about each cue with the cue values they receive on each trial, in a controlled explicit manner. However, if this were the case, then the addition of working memory load tasks to the test phase should have had some effect on performance. Instead, performance scores were entirely unaffected by the load tasks, suggesting that if explicit processing is involved in judgement then this may occur only in earlier stages of learning, and not after sufficient practice with the task. The present chapter, therefore, provides some suggestion that if explicit processing contributes to learning of tasks containing one or more negative cues, this explicit processing may become automated through practice and contribute little to judgement after expertise is acquired. Recall, however, that our conclusions based on a null effect of working memory load are only speculative as we have no clear evidence that our load tasks were sufficient to disrupt explicit judgement. For this reason, in the next chapter we introduce methods designed to explore whether explicit processing is involved at any stage of learning in MCPL tasks, by introducing measures of working memory capacity.

Chapter 4

The role of working memory capacity in multiple cue judgement

Chapter 3 indicated that while explicit processes may be involved in learning in multiple cue environments that contain negative cues, these processes contribute little to judgement after sufficient practice with a task. While performance was generally above chance in the MCPL tasks, the addition of a concurrent working memory load to a test phase after extensive training had no effect on performance. This suggests that if explicit processes are involved in learning from multiple cues, they only contribute to early stages of learning.

We introduce measures of working memory capacity to explore whether explicit processing is involved in judgement at any stage of learning. While attention may be necessary for both implicit and explicit learning to occur, explicit processing but not implicit processing is believed to be demanding on the individual's limited working memory resources (Barrett, Tugade, & Engle, 2004). We measure individual differences in working memory capacity (WMC) and correlate these scores with performance and explicit knowledge in the MCPL task. According to dual process theory, explicit rule-based processing selectively loads working memory so a correlation is expected when such processing is important on the task (Evans, 2008). Research on deductive reasoning and decision making has already shown that aspects of performance attributed to the explicit system typically correlate with measures of working memory capacity, or high correlates of it such as general intelligence or SAT scores (see De Neys, 2006; Evans, 2008, Stanovich & West, 2000). Similarly, performance on *rule-based* versions of category learning tasks that are purported to draw on explicit processes are both affected by the addition of a secondary task (Waldron & Ashby, 2001) and correlated with individual differences in WMC (DeCaro, Thomas, & Beilock, 2008). Since explicit learning in a MCPL task requires the testing of hypotheses against feedback, it is also important to note that participants of higher WMC are shown to generate more alternative hypotheses in probability judgement tasks than participants of lower WMC (Dougherty & Hunter, 2003a).

In Chapter 3, performance scores were well in excess of explicit knowledge levels in tasks containing only positive relevant cues, indicating that a substantial implicit component contributes to learning of positive cue tasks. Hence, we expect that individual differences in WMC will not predict performance levels in task types containing only positive relevant cues. Our predictions are less clear for tasks containing negative cues. In Chapter 3, participants performed slightly worse than expected based on their explicit knowledge of mixed cue tasks, containing both positive and negative cues. If implicit processes also contribute to learning in these tasks then we would have expected performance to again exceed explicit knowledge levels. This provides some indication that implicit processing is less involved in learning tasks that contain a negative cue. For this reason we suspect that learning in mixed cue tasks depends on explicit processing, and expect that learning will be associated with individual differences in WMC. For tasks containing only negative relevant cues however, performance again exceeded explicit knowledge of the task. It is therefore less clear whether explicit processing is involved in learning negative cue tasks. If explicit processes do contribute to learning, however, this must occur early in training.

4.1. EXPERIMENT 4

The findings of Chapter 3 suggest that explicit processing may be involved in learning multiple cue tasks containing a positive and a negative cue (+-00), in spite of the lack of effect of working memory load. However, it is unclear to what extent explicit processing may contribute to learning in ++00 and --00 tasks. Given the lack of effect of working memory load on the test trials, it seems that any contribution of explicit processing must occur early in the learning phase. If, however, the learning depends upon explicit processing *at any stage*, then we should be able to detect this by examining individual differences in working memory capacity. Essentially, any task that depends upon explicit learning should be performed better by those with higher working memory capacity. In the present experiment we hence introduce measures of working memory capacity to examine whether explicit processing is involved in judgement at any stage of learning.

We also examine the possibility of transfer effects in learning. In Chapter 3, we found that performance on tasks that contained negative cues was poorer than for positive cue tasks. One possibility is that prior training on negative cues (--00 task) could improve learning of difficult mixed cue tasks, because these also contain a negative cue. We test this by training one group of participants on a negative cue task immediately before performing a mixed cue task. Prior training on positive cues (++00 task) may also influence learning of a mixed cue transfer task. Based on our findings of Chapter 3, we have good reason to believe that learning in these tasks occurs implicitly. Hence, it is interesting to know whether implicit learning of positive cues has any transfer effects on learning of mixed cues. This is tested by a second group who instead perform a positive cue task immediately before a mixed cue transfer task.

4.1.1. Method

Participants

Eighty undergraduate psychology students at the University of Plymouth participated for course credits, forty in each of the two groups.

Design

All participants first completed a working memory task. Participants were then divided into two groups according to whether they first completed a judgement task containing two positive and two irrelevant cues (++00; positive cue task), or two negative and two irrelevant cues (--00; negative cue task). All participants then completed a second judgement task containing one positive, one negative, and two irrelevant cues (+-00; mixed cue task). Details of the MCPL task and procedure are provided in section 2.3.

Materials and Procedure

Each participant first completed a working memory task, followed by two judgement tasks (see section 7.3. for task instructions). Participants completed a cue-rating task only for the second task and were informed on screen that their ratings should be made only for the second task.

Working memory task. All participants first completed a working memory task. This was an operation span task, similar to that used by Engle, Cantor, and Carullo (1992; see also Unsworth & Engle, 2007) except that participants did not read aloud any features of the task. Instead, they were required to work through the task individually and in silence, by holding in mind each word that was presented on screen whilst verifying arithmetic identities (De Neys, Schaeken, & d'Ydewalle, 2005). Participants were instructed that the task was a memory test and would require them to hold in mind sets of words whilst verifying arithmetic identities (e.g. (6 X 3) – 5 = 13). On each trial, an identity appeared on screen with a word attached to the end of the identity. Participants were instructed to decide as quickly as possible whether the identity was correct or incorrect by clicking on the corresponding label with a mouse, and then to click on the word attached. It was explained that whilst judging the identities, they were to hold in mind each of the previous words in order. At the end of each set they were instructed to then write down the set of words on a piece of paper provided, ensuring that the set was in the order that the words appeared. Participants completed sets ranging from 3-6 words in length, three times for each set length. The length of each set was determined randomly, whilst ensuring that each set-length was presented three times. This meant that participants could not anticipate the length that each set would be. Working memory capacity scores were calculated by summing the number of recalled words for each set recalled completely. If an entire set was not recalled, or one or more words in the set were recalled in an incorrect order, the set was not counted.

4.1.2. Results

Working memory task performance

The mean score of the 80 participants on the working memory capacity measure was 28.39 (SD = 11.50) out of a possible 54.

Task 1 performance in the test phase

For the first task, participants performed either a positive cue (++00) or a negative cue (--00) task. Both mean group performance scores were significantly above zero. These are displayed in Table 4.1. We also measured performance scores in the 1st 40 and 2nd 40 learning trials, again by correlating judgements with criterion values for each participant. These analyses are reported in section 7.1.

Table 4.1

Experiment 4: Mean group performance and explicit knowledge scores

	Performance scores		Explicit knowledge scores	
	M	t	M	t
++00 task	.56*	13.141		
--00 task	.32*	5.222		
+-00 task (++00 transfer)	.26*	4.675	.30*	3.267
+-00 task (--00 transfer)	.21*	3.759	.33*	3.738

*$p < .05$

In order to make use of the full range of scores on working memory capacity, we performed a moderated regression analysis on performance and explicit knowledge scores. Moderated regression is a multiple regression analysis with one or more interaction terms (Aiken & West, 1991). Following the suggestion of Aiken and West (1991), all continuous variables included in an interaction term were mean-centred to reduce collinearity. To examine whether the type of task participants were required to perform moderated an association between WMC and performance, we performed a moderated regression analysis on performance scores in the first task. For this, we used task-type (++00 or --00) and WMC as predictors, and performance correlations as the dependent variable. More details of how we performed the moderated regression analyses are provided in section 7.2.

The regression analysis ($R^2 = .239$) revealed a significant main effect of the type of task participants performed ($\beta = -.329$, $t(77) = -3.153$, $p = .002$), with participants performing generally better in the ++00 task (.56) than the --00 task (.32). There was also a significant main effect of WMC ($\beta = .221$, $t(77) = 2.123$, $p = .037$), with participants of higher WMC performing better than participants of lower WMC. Interestingly, however, there was also a significant interaction (see Figure 4.1) between the type of task participants performed and their WMC (β = .381, $t(76)$ = 2.678, p = .009). Simple slope analysis (Aiken & West, 1991) showed that performance was positively predicted by WMC in the --00 task (β = .483, $t(38)$ = 3.397, p = .002) but not in the ++00 task (β = -.057, $t(38)$ = -.353, p = .726). This indicates that working memory dependent explicit processing is indeed involved in learning tasks containing negative cues, but not when both relevant cues are positive. In positive cue tasks explicit processing appears to contribute very little to judgement at any stage of learning. While implicit learning processes appear to guide judgement at all stages of learning in these tasks, learning of negative cues initially require intervention by effortful explicit processing. The explicit knowledge used to make judgements in learning phases of negative cue tasks likely becomes automated through practice, explaining our lack of effects of concurrent load tasks on test trials in Chapter 3. Analyses of performance scores in learning phases of this chapter are reported in section 7.1.

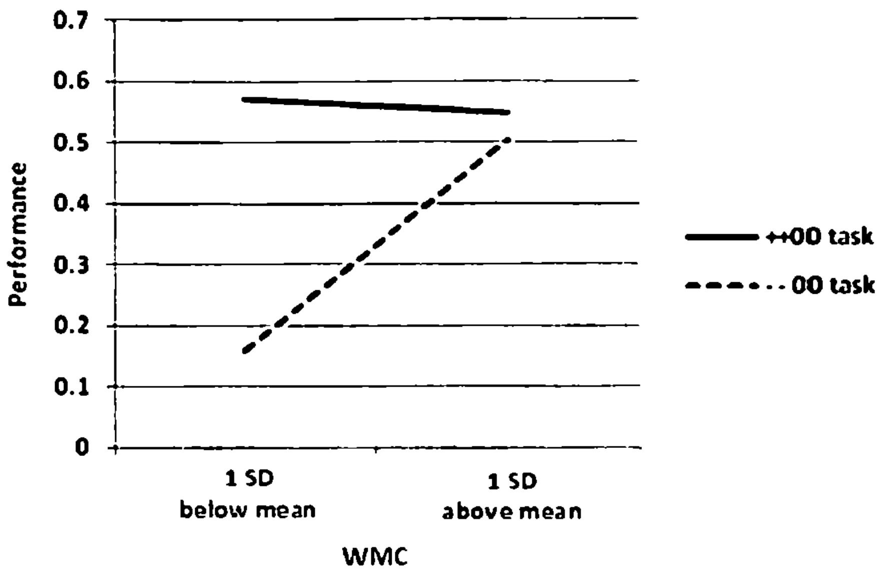

Figure 4.1. Experiment 4: Interaction effect between type of task and WMC on performance scores. The regression slopes plotting the association between performance and WMC for each task type are cut off at 1 SD below the mean WMC score of all participants, and 1 SD above the mean.

Mixed cue transfer task performance in the test phase

As can be seen in Table 4.1, both groups achieved mean group performance scores that were significantly above zero in the +-00 task. Independent t-tests confirmed our earlier findings that participants perform significantly better in the ++00 task (.56) than the +-00 task (.26; t = 4.413, df = 39, p <.001). Although performance appeared lower in the +-00 task (.21) than the --00 task (.32), this difference did not reach significance (t = 1.889, df = 39, p = .066). See section 7.1. for analysis of performance scores in learning phases.

We then conducted a moderated regression analysis on performance scores in the transfer task (+-00) to investigate whether the type of task participants had previously performed (++00 or --00) had any moderating effects on an association between WMC and performance in this task. The regression analysis (R^2 = .162) revealed no significant main effect of the type of task participants had previously performed (β = -.061, $t(77)$ = -.553, p = .582), but a significant main effect of WMC (β = .270, $t(77)$ = 2.458, p = .016), indicating participants of higher WMC generally performed better in the +-00 task than those of lower WMC. Interestingly, however, there was also a highly significant interaction (see Figure 4.2) between the type of task participants had previously performed and their WMC (β = .411, $t(76)$ = 2.750, p = .007). Simple slope analysis confirmed that WMC positively predicted performance in the +-00 task for those who had previously completed the --00 task (β = .610, $t(38)$ = 4.748, p <.001), but *not* for those who had previously completed the ++00 task (β = -.021, $t(38)$ = -.127, p = .900).

This analysis suggests that the mode of processing used to learn a task carries over when a second task is performed. When participants are required to first perform a positive cue task, which can be learned implicitly, this mode of processing is used to learn the mixed cue transfer task. In contrast, explicit processing appears involved in learning tasks containing negative cues, and this explicit mode of processing transfers to learning of the mixed cue task. This is indicated by the positive association between WMC and performance in both the negative cue and mixed cue transfer tasks.

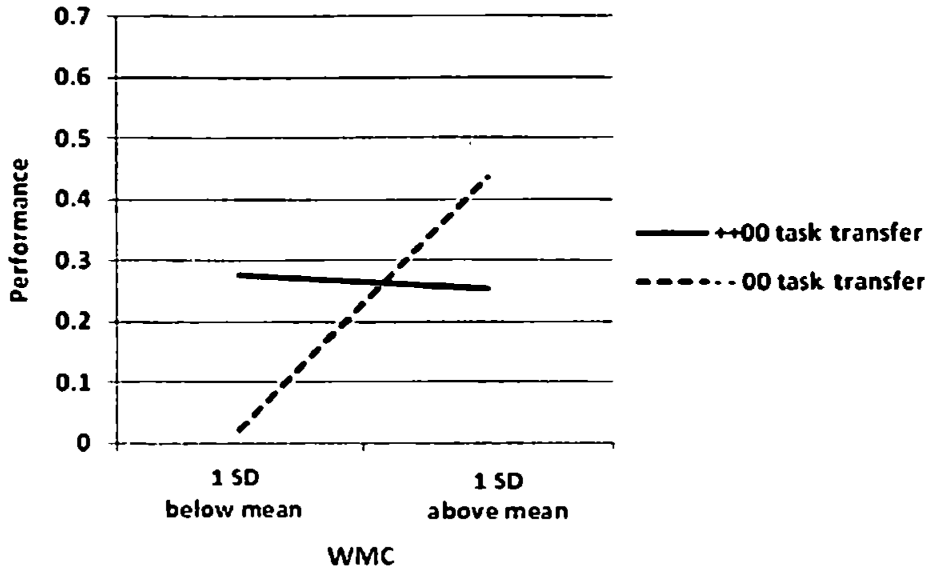

Figure 4.2

Experiment 4: Interaction between prior task type and WMC on performance scores in the +-00 task

Explicit knowledge in the mixed cue transfer task

Explicit knowledge scores were calculated in the same way as in Chapter 3, by correlating predicted judgements based on participants' explicit beliefs with criterion values. Further details of how explicit knowledge scores were calculated are provided in section 2.3. Displayed in Table 4.1 are the mean group explicit knowledge scores in the +-00 task. For both groups these are significantly above zero. We found in Chapter 3 that explicit knowledge scores often exceed performance on mixed cue tasks. Consistent with those findings, comparing participants' performance with their explicit knowledge scores in the mixed cue task, the results confirm that participants did not perform better than expected based on their explicit knowledge. In fact, explicit knowledge scores (.32) were significantly higher than performance scores (.24; t = 2.045, df = 79, p = .044).

We performed a final moderated regression analysis (R^2 = .072) on explicit knowledge scores in the +-00 transfer task, using previous task-type (++00 or --00) and WMC as predictors. There was no significant effect of the previous task participants performed (β = .020, $t(77)$ = .177, p = .860), or WMC (β = .107, $t(77)$ = .944, p = .348). However, there was again a significant interaction between previous task type and WMC (β = .348, $t(76)$ = 2.216, p = .030). Simple slope analysis revealed that WMC was significantly associated with explicit knowledge in the +-00 task for those who had previously completed the --00 task (β = .348, $t(38)$ = 2.291, p = .028), with higher WMC participants performing better than those of lower WMC, but not for those who had previously completed the ++00 task (β = -.140, $t(38)$ = -.875, p = .387). These results confirm our analysis of performance scores. When individual differences in WMC are positively associated with performance in this task, higher WMC is also associated with more accurate explicit knowledge of the task. This supports our conclusions that explicit processing is involved in learning of mixed cue tasks, but not when preceded by a task that can be learned well implicitly.

4.1.3. Discussion

The findings of Experiment 4 indicate that negative cue learning (both in +-00 and --00 tasks) but not positive cue learning (++00) is associated with individual differences in WMC (see Figures 4.1 and 4.2), suggesting that this learning benefits from explicit processing. As already stated, we believe this processing must occur in the early phase of learning as any use of explicit judgement in the test phase should have been disrupted by working memory load in Chapter 3. In Experiment 4, we measured explicit beliefs only for the mixed cue (+-00) tasks but again the results are consistent with our hypothesis. Under conditions where those of higher working memory capacity performed better on this task, they also showed higher levels of explicit knowledge.

An interesting novel finding, however, was that under some conditions this explicit learning apparently did *not* occur. Performance (and explicit knowledge) on the mixed cue task was only related to working memory capacity when preceded by the negative cue task. If the positive cue task was performed first, working memory capacity became irrelevant. We already have reason to believe that the positive cue task benefits little from explicit processing. First, in Chapter 3 we showed that performance was comfortably in excess of explicit knowledge levels for this task. Second, in Experiment 4 we have shown that performance on this task is unrelated to working memory capacity. It now also appears that performing this task first induces an implicit mode of processing that transfers to the subsequent mixed cue task, even though this results in little drop in performance and explicit knowledge levels observed (see Table 4.1). One explanation for why explicit knowledge scores did not reduce in the mixed cue task following positive cue training is that participants acquired explicit knowledge as a result of observing their own behaviour during the test trials. This is likely since explicit knowledge scores are based on participants' explicit beliefs only after completing the 40 test trials. Our finding of a transfer effect clearly merits further investigation.

4.2. EXPERIMENT 5

Experiment 4 provides strong evidence that explicit processing is normally involved in learning tasks with negative or mixed cues. However, we have also found evidence of transfer effects, such that explicit processing of the mixed cue task is inhibited if preceded by training on the positive cue task. One limitation of Experiment 4 was that we only studied mixed cue learning after performance of another task. In Experiment 5, we therefore presented participants with mixed cue learning as their sole task. The first aim, therefore, was to show that performance on this task would still be related to individual differences in working memory capacity. The second was to see if we could enhance performance by provision of partial task information. We know that performance will be very high if participants are told precisely what each cue does in advance (Balzer, Doherty, & O'Conner, 1989; Evans, Clibbens, & Harris, 2005), but this is of limited interest. The information we supplied (to one group) was that one of the cues was a positive predictor, one negative and two irrelevant. They still had to work out which cues were which. If participants are engaging in hypothesis testing, as many authors have suggested for MCPL in general (e.g. Brehmer, 1974; 1994; Juslin, Karlsson, & Olsson, 2008), then we might well expect this to improve their performance.

4.2.1. Method

Participants

Seventy-two undergraduate students at the University of Plymouth participated for course credits, 36 in each of the two groups.

Design

The design was similar to Experiment 4, except participants only completed the second (transfer) judgement task (+-00; mixed cue task) and instead were divided into two groups according to whether they received an instructional manipulation. The instructional manipulation involved presenting one group with task information. In this condition they were informed explicitly that one of the cues would be positive, one negative, and two irrelevant.

Materials and Procedure

All participants first completed the same working memory task as used in Experiment 4. All participants then completed a single judgement task (see section 7.3. for task instructions). The judgement task was the same as the transfer task used in Experiment 4, containing one positive, one negative, and two irrelevant cues (+-00). Importantly, one group was presented with the additional task information immediately before beginning the judgement task. All participants then completed the cue-rating task immediately after completing the judgement task.

4.2.2. Results

Working memory task performance

The mean score of the 72 participants on the working memory capacity task was 25.93 (*SD* = 12.28) out of a possible 54.

Mixed cue task performance in the test phase

Both mean group performance scores in the +-00 task were significantly above zero. These are displayed in Table 4.2.

Table 4.2

Experiment 5: Mean group performance and explicit knowledge scores

	Performance scores		Explicit knowledge scores	
	M	t	M	t
+-00 task	.33*	5.349	.37*	3.856
+-00 task (with task information)	.46*	6.868	.44*	4.951

*p <.05

A moderated regression analysis (R^2 = .152) was carried out on performance scores in the +-00 judgement task in order to investigate whether there were any moderating effects of task information on an association between WMC and performance. There was a significant main effect of task information (β = .233, $t(69)$ = 2.103, p = .039), with participants performing better in the +-00 task when provided task information (.46) than for those who were not provided task information (.33). There was also a significant main effect of WMC (β = .314, $t(69)$ = 2.825, p = .006), confirming that participants of higher WMC perform better than those of lower WMC in the mixed cue task. There were no significant interactions.

The association between WMC and performance is displayed in Figure 4.3, and confirms that higher WMC is ordinarily associated with better performance in mixed cue tasks. In addition, we expected that providing task information would improve performance only if explicit processing contributes to learning of these tasks. Our finding that performance levels were increased by task information provides further support that performance in mixed cue tasks usually depends on explicit processing. These findings thus confirm our conclusions of Experiment 4 that prior training on positive cues has a negative transfer effect on learning of a second task containing a negative cue. While higher WMC affords an advantage in learning these tasks, explicit processing contributes little to learning if a prior task can be learned well implicitly. In this way, explicit processing is switched off by positive cue training.

Interestingly, participants of all levels of WMC benefited equally from task information. Under these conditions, participants are not required to generate their own hypotheses, but must still evaluate hypotheses against the feedback to work out which cues are which. Our finding that performance is associated with WMC even with the addition of task information indicates that both generating and evaluating hypotheses against feedback is demanding on working memory.

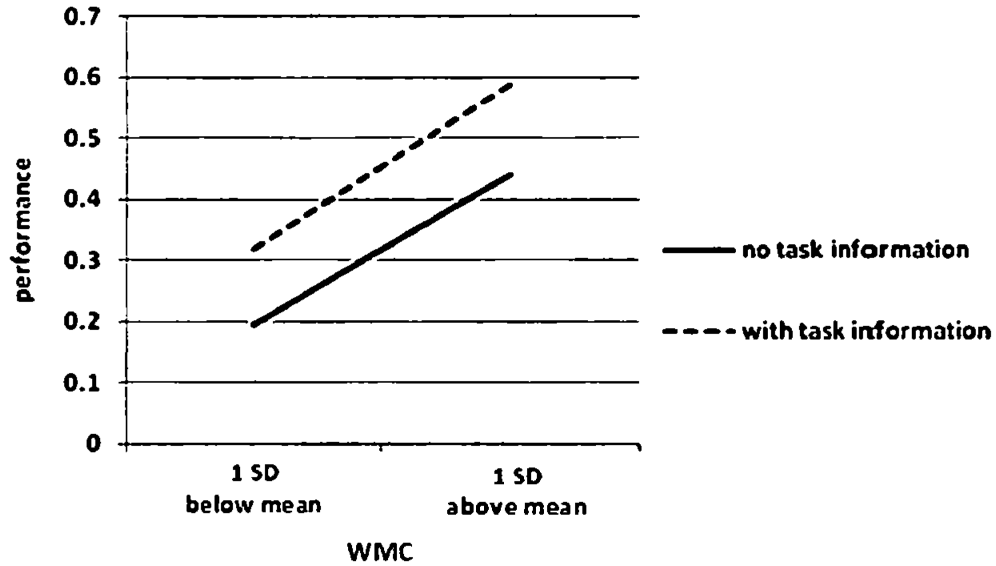

Figure 4.3

Experiment 5: Association between WMC and performance scores in the +-00 task

Mixed cue task explicit knowledge

Mean group explicit knowledge scores are displayed in Table 4.2. These were both significantly above zero. As in our previous experiments, participants did not perform better than expected based on their explicit knowledge. A pairwise comparison confirmed that performance scores (.40) were not significantly different from explicit knowledge scores (.41; t = .303, df = 71, p = .763).

We performed a moderated regression analysis (R^2 = .111) on explicit knowledge scores in the +-00 task. Although participants who were provided task information appeared to acquire more accurate explicit knowledge (.44) than those who were not provided task information (.37), this effect was not close to being significant (β = -.048, $t(69)$ = -.423, p = .674). However, there was a significant main effect of WMC (β = .327, $t(69)$ = 2.877, p = .005), with higher WMC participants acquiring more accurate explicit knowledge of the task than those of lower WMC. There were no significant interactions. The main effect of WMC on explicit knowledge can be seen in Figure 4.4, and confirms our analysis of performance scores in the test phase of the mixed cue task. When WMC is positively associated with performance, higher WMC is also associated with higher levels of explicit knowledge. In addition, explicit processing appears involved in learning mixed cue tasks when no prior task is performed, indicating that such tasks are ordinarily learned by effortful explicit processing.

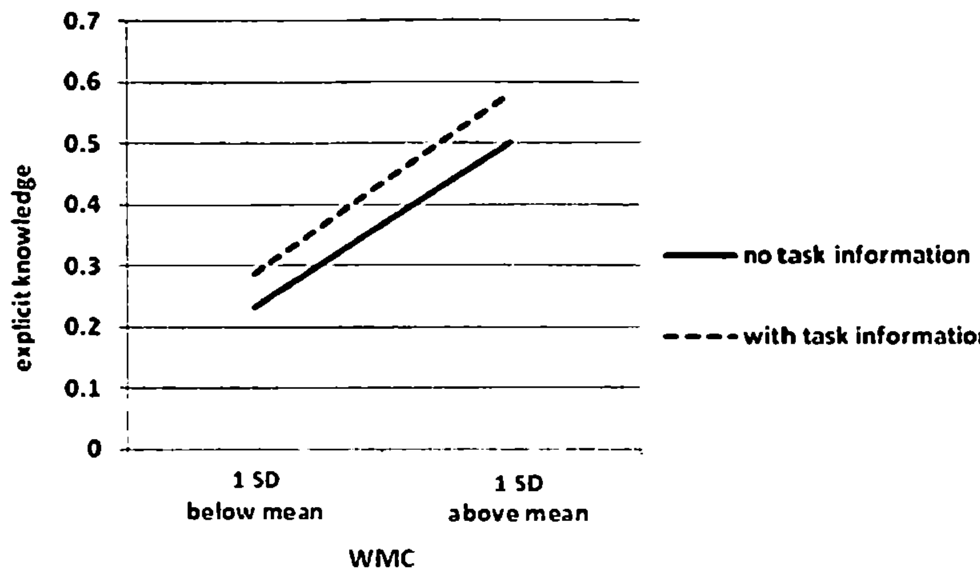

Figure 4.4

Experiment 5: Association between WMC and explicit knowledge scores in the +-00 task

4.2.3. Discussion

Experiment 5 supports our interpretation of findings to date. That is, it seems that performance on the mixed cue task normally benefits from explicit processing. In this experiment, both performance and explicit knowledge levels were related to individual differences in working memory capacity, just as they were in Experiment 4, when preceded by negative cue learning. Hence, we can say with some confidence that it is not the case that explicit processing on this task is induced by prior learning of negative cues. Rather it is *inhibited* (in Experiment 4) by prior experience of the positive cue task. Interestingly, telling participants that there would be one positive and one negative cue among the four presented produced a significant improvement in performance without significantly affecting explicit knowledge. It is important to note, however, that explicit knowledge did appear higher with the addition of task information, though the effect did not reach significance (see Figure 4.4). The fact that task information improved performance suggests again that explicit processes are actively involved in learning the mixed cue task.

4.3. EXPERIMENT 6

The results of Experiments 4 and 5 combined enable us to infer a negative transfer effect. Mixed cue tasks normally benefit from explicit processing *unless* preceded by learning of a positive cue task. Our final experiment was designed both to replicate and further investigate this finding. It seems that participants of higher working memory capacity, who would normally use explicit processes on this task, get these 'switched off' by prior performance of a task on which implicit learning is very effective. In this experiment we attempt to switch them back on again.

In Experiment 6, we aimed to replicate our previous findings of the transfer effects of positive cue training on mixed cue learning and our earlier evidence of the independence of individual differences in WMC in positive cue learning. This was done by again requiring participants to first perform a positive cue task containing only positive relevant cues before performing a mixed cue task containing one positive, one negative, and two irrelevant cues.

Here we attempted to *undo* the suppressing effects of positive cue training on explicit processing in the mixed cue task by introducing explicit instructions to participants before beginning the mixed cue task. If positive cue training reduces the role of explicit processing then providing explicit instruction to expect at least one negative cue in the mixed cue task may encourage participants to approach the task explicitly. If this is the case then we would expect performance in the mixed cue task to be dependent on individual differences in WMC following explicit instruction.

4.3.1. Method

Participants

Seventy-two undergraduate students at the University of Plymouth participated for course credits, 36 in each of the two groups.

Design

The design was similar to Experiment 4, except that all participants first completed the same positive cue task, containing two positive and two irrelevant cues (++00). All participants then completed the same mixed cue transfer task, containing one positive, one negative, and two irrelevant cues (+-00). Importantly, one group of participants were given additional task instructions immediately before beginning the transfer task, explaining that the task will contain one or more negative cues, whilst the other group was not informed that negative cues would be present.

Materials and Procedure

The procedure was similar to Experiment 4, with all participants first completing the working memory task, then the two judgement tasks, followed by the cue-rating task (see section 7.3. for task instructions). However, immediately before beginning the transfer task an additional task instruction appeared on screen for one group of participants, explaining that one or more of the cues would be negatively predictive. All participants then completed the cue-rating task and were reminded that their ratings should be based only on the cues in the second task.

4.3.2. Results

Working memory task performance

The mean score of the 72 participants on the working memory capacity measure was 25.90 (SD = 13.13) out of a possible 54.

Task 1 performance in the test phase

Combining all 72 participants' performance scores in the ++00 task, participants achieved a mean score of .61, which was significantly above zero ($t = 23.707$, $df = 71$, $p < .001$). Confirming the results of Experiment 3, there was no significant association between performance and WMC ($r = .116$, $p = .333$). That is, explicit processing contributes little to learning of tasks containing only positive relevant cues. Furthermore, implicit processing appears to lead to accurate judgement in positive cue tasks, with participants performing well above chance (see Table 4.3). Performance scores in learning phases are provided in section 7.1.

Table 4.3

Experiment 6: Mean group performance and explicit knowledge scores

	Performance scores		Explicit knowledge scores	
	M	t	M	t
++00 task	.61*	23.707		
+-00 task (++00 transfer)	.31*	6.928	.33*	3.676
+-00 task (++00 transfer with explicit instruction)	.31*	6.172	.40*	4.564

*p <.05

Mixed cue transfer task performance in the test phase

Both mean group performance scores in the +-00 task were significantly above zero. These are displayed in Table 4.3. An independent t-test confirmed that participants performed significantly better in the ++00 task (.61) than the +-00 task (.31; t = 7.817, df = 71, p <.001).

A moderated regression analysis (R^2 = .028) was carried out on performance scores in the +-00 task using the addition of explicit instruction (no explicit instruction or with explicit instruction) and WMC as independent predictors. There was no main effect of WMC (β = .162, $t(69)$ = 1.364, p =.177), or explicit instruction (β = -.010, $t(69)$ = -.086, p =.932), with higher WMC affording no advantage for learning in the +-00 task following positive cue training and with explicit instruction doing nothing to reduce the transfer effects. In addition, there were no significant interactions. Our finding that WMC was not associated with performance in the transfer task even when participants were instructed explicitly to expect at least one negative cue, suggests that prior training with positive cues has a strong effect on learning in a second task. Explicit instruction simple does not appear sufficient to override the effects of positive cue training.

Mixed cue transfer task explicit knowledge

Mean group explicit knowledge scores are displayed in Table 4.3. These were above zero and significant for both groups. In accordance with our previous experiments, participants did not perform better than expected based on their explicit knowledge alone. Whilst explicit knowledge scores (.37) appeared slightly higher than performance scores (.31), this difference did not reach significance (t = 1.330, df = 71, p = .188). Therefore, in contrast with our findings of positive cue tasks, performance levels do not exceed explicit knowledge scores when tasks contain both positive and negative relevant cues.

We performed a moderated regression analysis (R^2 = .004) on explicit knowledge scores in the +-00 task, using explicit instruction (no explicit instruction or with explicit instruction) and WMC as predictors. Again, there was no significant main effect of explicit instruction (β = .033, $t(69)$ = .276, p =.784), WMC (β = .050, $t(69)$ = .419, p =.677), or any significant interactions, confirming that higher WMC capacity does not afford any advantage in explicit knowledge acquisition in the +-00 task following positive cue training (see Figure 4.5). In addition, explicit instruction did not reduce the negative transfer effect of positive cue training on explicit processing. If explicit processes were involved in learning we would have expected a positive association between WMC and explicit knowledge levels. Despite telling participants to expect one or more negative cues in the transfer task, explicit processing appeared to contribute little to learning.

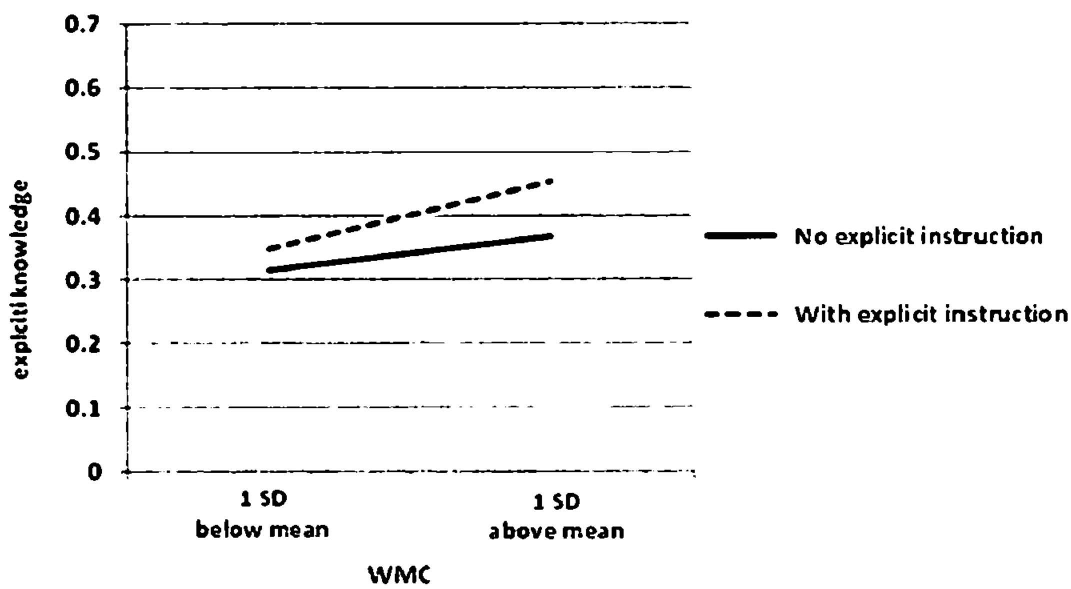

Figure 4.5

Experiment 6: Lack of association between WMC and explicit knowledge scores in the +-00 task following positive-cue training

4.3.3. Discussion

Experiment 6 supports our earlier findings that positive cue training reduces the role of explicit processing in the mixed cue transfer task, suggesting this is a robust effect. Furthermore, telling participants to expect at least one negative cue in the transfer task failed to increase explicit processing following positive training. Prior experience with positive cues thus has a strong effect on reducing explicit processing in learning of the transfer task. Although there was some tendency for higher WMC participants to acquire more explicit knowledge than participants of lower WMC following explicit instruction (see Figure 4.5), this manipulation did not appear to have any substantial effect. This suggests that explicit instruction is simply not sufficient to elicit explicit processing following positive cue training.

4.4. GENERAL DISCUSSION

The experimental studies of Chapter 3 indicated that explicit processing may be required for learning in multiple cue environments that contain negative cues, but that explicit knowledge in any case does not contribute to judgement following training. In the present chapter we introduced measures of WMC and correlated these with performance and explicit knowledge scores to see whether explicit processing is involved at any stage of learning in multiple cue tasks. Our rationale for measuring individual differences in WMC was that explicit processing but not implicit processing loads heavily on the individual's limited working memory resources. For this reason, tasks that draw on explicit processes tend to correlate with measures of working memory capacity (DeCaro et al, 2008; De Neys, 2006; Stanovich & West, 2000). Confirming our conclusions of Chapter 3, individual differences in WMC were indeed associated with learning in tasks that contained one or more negative cues, with higher WMC affording an advantage in learning these tasks, but not in tasks that contained only positive relevant cues. Performance in positive cue tasks was instead entirely independent of individual differences in WMC.

A possible interpretation of our findings so far is as follows. Positive cue tasks can be performed well on the basis of implicit learning, which is why they are the easiest and above the level expected from explicit beliefs acquired. Negative and mixed cue tasks do require explicit processing effort to acquire. However, this benefit likely occurs in the early stages of learning and explicit knowledge may become automated by the time the test trials are administered. This explains the lack of effect of working memory load found in Chapter 3 at this final stage. In the present chapter we instead measured individual differences in WMC, rather than administering working memory load. Our hypothesis was that such individual differences would be predictive of performance in negative and mixed cue tasks, but not positive cue tasks. With the exception of a negative transfer effect (discussed below) this prediction was confirmed throughout the experimental studies of the present chapter. Moreover, our measures of explicit knowledge were significantly related to individual differences in WMC for tasks where those differences predicted performance, but independent of WMC for tasks where it did not. These findings provide compelling evidence that working memory dependent explicit processing is necessary for learning judgement tasks that contain one or more negative cues, but contributes little to learning of positive cue tasks. While positive cue tasks can be learned implicitly, learning of negative cues draws heavily on limited working memory resources associated with the explicit system, such that higher WMC affords an advantage in learning such tasks.

Another novel finding of the present Chapter was a negative transfer effect. Performance on mixed cue tasks (which have one positive and one negative cue) was associated with WMC and explicit knowledge except when preceded by the learning of a positive cue task. In other words, performing a task where one can rely on implicit processing for effective learning (the positive cue task) appears to carry that mode of processing over to one where explicit processing is normally involved. This meant that implicit processing contributed to learning on a task this is usually dominated by explicit processes, such that higher WMC no longer afforded an advantage in performing the task. Instead, when participants had prior experience of a task that requires effortful explicit processes (negative cue task) an explicit mode of processing continued to contribute to participants' learning in the second task. However, we are not aware of any previous report of a negative transfer effect, in which explicit processing is inhibited by previous experience of the type reported here.

It has been generally assumed by dual process theorists in the psychology of reasoning and decision making (e.g. Evans, 2008) that explicit processing will be responsive to experimental instructions. In Experiments 5 and 6 we presented participants with partial task information. In Experiment 5, we found that telling participants in advance that the mixed cue task would contain exactly one positive and one negative cue (but not telling them which) produced a significant benefit in performance on the task. This suggests that participants were able to use the task information to direct their search for relevant cues in the task, confirming that explicit processes are involved in learning the mixed cue task. However, we found in Experiment 6 that telling people that the mixed cue task would include at least one negative cue was ineffective in reducing the negative transfer effect. In spite of this instruction, participants failed to provide any evidence of explicit processing on the task, when it was preceded by positive cue learning. This suggests that explicit instruction is insufficient to undo the effects of prior training on explicit processing. It appears that explicit processing is switched off by prior training on positive cues and is insensitive to both the presence of a negative cue in a transfer task and explicit instruction about the presence of negative cues.

We also found that when participants were provided partial task information, performance in mixed cue tasks was still associated with individual differences in WMC. This suggests that even when participants are told specifically which types of cues will be present, evaluating hypotheses against feedback is heavily demanding on working memory. Hence, there are two aspects of explicit learning in multiple cue tasks. First, individuals must explicitly generate hypotheses about the cue-criterion relations, and second they must evaluate whether the hypotheses are supported by the cue-outcome values they observe.

The present chapter supports our predictions that learning in multiple cue tasks that contain one or more negative cues requires effortful explicit processing. However, explicit processes only contribute to judgement during early stages of learning in these tasks, and any explicit knowledge that guides judgement quickly becomes automated, drawing less on limited working memory resources. In contrast, learning environments that contain only positive relevant cues can be learned effectively by implicit processes. Furthermore, experience with positive cue judgement tasks can induce an implicit mode of thinking that transfers to tasks that would ordinarily be learned explicitly. Telling participants specifically that a task will contain one positive, one negative, and two irrelevant cues before performing the more difficult mixed cue task (+-00) improved performance in the present chapter. However, participants of lower WMC were still at a disadvantage in learning these tasks. This provides some indication that mixed cue tasks are particularly difficult because participants must hold in mind their hypotheses about each cue during learning. In the following chapter we attempt to improve peoples' learning in multiple cue environments by allowing participants to keep a note of their current hypotheses about each cue during the learning phase. As well as introducing other task manipulations, we use participants' trial-by-trial explicit hypothesis-testing to further explore the contribution of implicit and explicit processes to multiple cue judgement.

Chapter 5

The role of explicit hypothesis testing in learning from multiple cues

Our findings so far indicate that while multiple cue judgement draws little on explicit processing after sufficient practice, learning in tasks that contain one or more negative cues is heavily demanding on working memory. Furthermore, many individuals lack sufficient working memory resources to make accurate judgements in tasks that contain negative cues.

The present chapter introduces task manipulations designed to improve peoples' learning in difficult multiple cue environments. In Experiments 7 and 8 we provide a 'hypotheses notepad' on screen for participants during the learning phase of judgement tasks. On each trial participants are instructed to note their current hypotheses about each cue (as positive, negative, or irrelevant). They are encouraged to use the outcome feedback they receive to guide their hypotheses selections, and to use their selections to help them make their judgements. Our rationale is that by displaying participants' hypotheses on screen and allowing them to change their selections in response to feedback, this should reduce the demands of explicit hypothesis-testing on working memory. This is because when participants are provided the hypotheses notepad they do not need to hold in mind their current hypotheses about the cues when making judgements. They can also update their hypotheses in response to feedback by changing their selections on screen, rather than test hypotheses in mind. In Experiment 8, we go one stage further and provide one group of participants the output of their hypotheses selections when feedback is provided on learning trials. In this way, participants are able to compare their actual judgements with the judgements they could have made had they used their explicit beliefs of each cue's relevance. We expect that by providing participants the output of their hypotheses selections they can use the feedback they receive to improve their selections, and in turn use their selections to guide their judgements.

A secondary aim of the present chapter is to measure explicit knowledge levels during learning phases of multiple cue tasks. In previous experiments we asked participants to provide a single rating of each cue's relevance following completion of the judgement task, and used their ratings to predict the judgements they would have made in a test phase had they used only their explicit beliefs about each cue. Introducing the 'hypotheses notepad' to learning phases allows us to use participants' explicit hypothesis-testing on each trial to measure explicit knowledge during training.

5.1. EXPERIMENT 7

In Experiment 7 we attempt to improve explicit learning in difficult versions of multiple cue judgement tasks, containing both positive and negative relevant cues. Participants are provided a 'hypotheses notepad' on learning trials to note on screen their explicit beliefs about each cue. This way they do not need to hold in mind their hypotheses about the cues when making judgements and assessing the feedback. We expect that performance in difficult mixed cue tasks will be less associated with individual differences in working memory capacity when the hypotheses notepad is provided. We also expect the hypotheses notepad to facilitate explicit hypothesis-testing and improve explicit knowledge of the task.

In Experiment 7 we also apply our methods to a judgement task containing only positive relevant cues. We know from our previous studies that this type of task can be learned well implicitly, and does not load on working memory dependent explicit processing. It is of interest to know whether explicit processing can be increased on tasks that are ordinarily learned implicitly, or whether implicit knowledge will continue to dominate judgement.

5.1.1. Method

Participants

Seventy six undergraduate psychology students at the University of Plymouth participated for course credits, with 38 participants in each of the two groups.

Design

All participants first completed a working memory task. Participants were then divided into two groups depending on whether they performed two judgement tasks each containing two positive and two irrelevant cues (++00; positive cue task) or one positive, one negative, and two irrelevant cues (+-00; mixed cue task). It was ensured that the order that cues were presented on screen differed between the two judgement tasks participants performed. In one of the judgement tasks participants were provided an additional 'hypotheses notepad'. We refer to this task as the 'hypotheses notepad' task, in contrast with the 'standard' task that does not contain the hypotheses notepad (see section 2.3.). The order that participants completed the two judgement tasks was randomly determined. Following completion of each judgement task participants completed the cue-rating task, and were reminded that their ratings should be made only for the task that they had just completed.

Materials and procedure

Each participant completed a working memory task, followed by two judgement tasks.

Working memory task. In Experiment 7 we introduced an automated version of the operation span task used to measure working memory capacity in Chapter 4. The automated version was based on Unsworth, Heitz, Schrock, & Engle (2005). As with the operation span task, participants were told that the task was a memory test and that they would be required to hold in mind lists of words whilst verifying arithmetic identities. Participants first completed ten practice trials. On each trial they were presented with an arithmetic identity (e.g. (2 X 1) + 1 = ?) and were required to click on an icon once they had solved the identity. They were then shown a number in the centre of the computer screen and asked to verify whether the number was the correct solution of the identity by clicking on either a 'correct' or 'incorrect' label with a mouse. Participants were instructed to work through the task as quickly and as accurately as possible. Following completion of the practice phase, participants were again required to verify arithmetic identities, except this time a word was displayed for 800 mscs immediately after participants verified each identity. Participants were instructed to hold in mind each word whilst performing the task. At the end of each set of words participants were asked to write the list down on a piece of paper in order that the words were presented. They then began the next list. Lists ranged from 4-7 words in length and each list length was used twice, in random order. If participants took longer than their average time taken to solve arithmetic identities in the practice phase plus 2.5 standard deviations, the next word in the set was displayed immediately and the procedure continued. Thus the task imposes a strict time limit on verifying each identity, designed to prevent participants from rehearsing word lists. Using an automated version of the working memory task allows us to validate our earlier findings using a slightly different method of measuring individual differences in working memory capacity. As in Chapter 4, working memory scores for each participant were calculated by summing the number of correctly recalled words only for lists that were recalled completely.

Hypotheses notepad task. All participants completed two judgement tasks. We followed the same task methodology used in our previous experiments (see section 2.3) with two exceptions. First, participants were given five practice trials before beginning each judgement task. On these trials all four cues were irrelevant. Following completion of the practice trials participants were informed that the cues were entirely non-predictive of the criterion and that in the following task one or more of the cues may be predictive. The practice trials were introduced in Experiment 7 to familiarise participants with the judgement task and the hypotheses notepad. Second, a 'hypotheses notepad' was added to the learning phase of one of the judgement tasks. Figure 5.1 provides a screen shot of one of the learning trials in the hypotheses notepad task. On the first learning trial the hypotheses selections were set to irrelevant for each cue, and participants were informed that they could not change the selections at this point. After making their first judgement, outcome feedback was displayed on screen and participants were explained that they could change the hypotheses selections if they wished. When participants clicked on an icon to move to the next trial they were instructed to use their hypotheses selections to help them make their judgement (see Figure 5.1). This procedure continued for the 80 learning trials. On each trial participants were able to change their hypotheses selections only after making a judgement, and were instructed on each trial to use their hypotheses selections when making their judgements (see section 7.3. for task instructions).

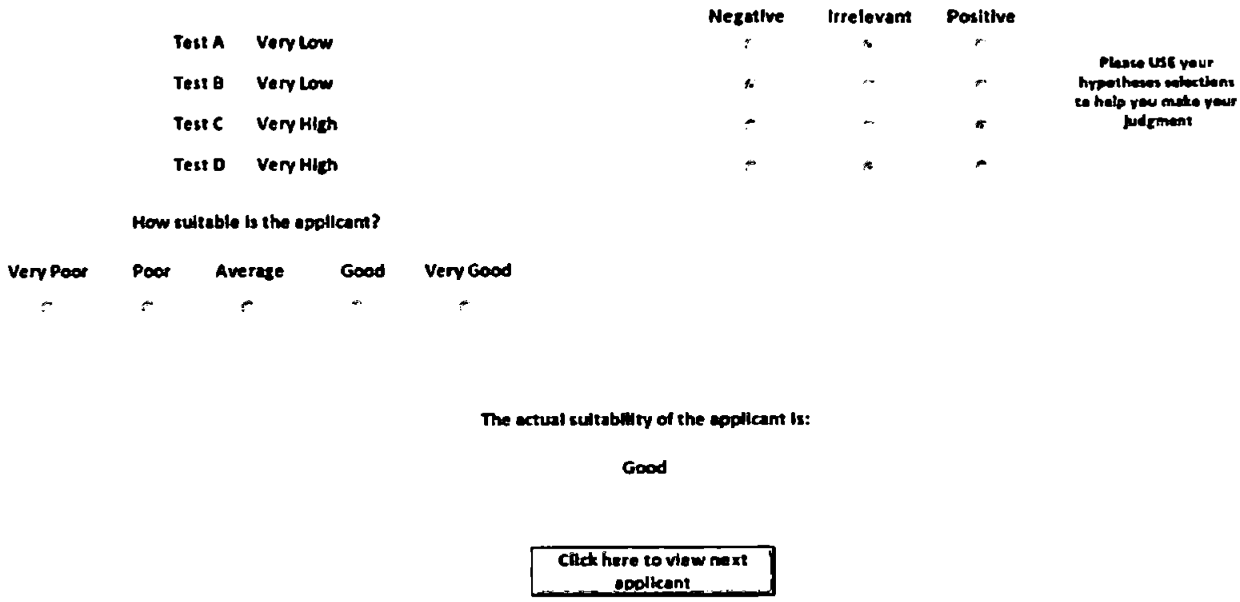

Figure 5.1. Screen shot of a learning trial in the 'hypotheses notepad' condition

5.1.2. Results

Performance in the test phase

Performance scores in the test phase were calculated by correlating judgements with criterion values in the 40 test trials for each participant, as in previous chapters. Mean group performance scores are displayed in Table 5.1, and were all above zero and significant.

Table 5.1

Experiment 7: Mean group performance scores

	1st 40 learning trials		2nd 40 learning trials		Test Trials	
	M	t	M	t	M	t
++00						
Standard task	.54*	14.860	.56*	17.759	.61*	17.840
Hypotheses notepad task	.48*	14.108	.50*	13.165	.59*	20.268
+-00						
Standard task	.21*	4.838	.21*	4.284	.16*	3.668
Hypotheses notepad task	.21*	4.395	.25*	4.655	.19*	3.588

*p = .05

In our analyses of performance scores Chapter 4 we entered our main effects as predictors into a moderated regression model. This allowed us to examine whether the type of task participants performed moderated an association between WMC, as a continuous variable, and performance. However, multiple regression treats predictor variables as independent factors, so is not suitable for experimental designs that contain both within-subjects factors and continuous predictor variables (Hoffman & Rovine, 2007). In our analysis of performance scores in test phases of the present experiment we were interested in whether an association between WMC and performance was moderated by condition (standard task and hypotheses notepad task) as a within-subjects factor, and task-type (++00 or +-00) as an independent factor. For this reason, in the present chapter we entered our main effects as predictors into a multilevel model, using the SPSS MIXED procedure. Multilevel analysis allowed us to explore variation within participants (within-subjects factors) on a first level, and variation between participants (between-subjects factors) on a second level. More details about the multilevel procedure used in the present chapter are provided in section 7.2.

In our multilevel analysis of performance scores in test phases we entered task-type (++00 or +-00), condition (standard task and hypotheses notepad task), and WMC as predictors, and a random intercept (χ^2 difference (1) = 13, p <.01; see section 7.2. for more details). Adding random slopes did not improve the model. There was a significant effect of task-type (β = -.550, SE = .073, t(76) = -7.563, p <.001), indicating that participants performed better in ++00 (.60) than +-00 (.18) tasks, but no effect of WMC (β = .003, SE = .004, t(76) = .708, p = .481). We expected that providing the hypotheses notepad to participants during learning phases would improve their performance on test trials. However, participants performed similarly in hypotheses notepad (.39) and standard (.39) tasks (β = -.015, SE = .048, t(76) = -.319, p = .750). Hence, providing a notepad on screen to allow participants to keep track of their explicit hypothesis-testing does not appear to improve performance. As in Chapter 4, there was however, a significant interaction between task-type and WMC (β = .016, SE = .007, t(76) = 2.145, p = .035). The simple slope for the +-00 task was positive and significant (β = .010, SE = .005, t(38) = 2.192, p = .035), confirming that higher WMC was associated with better performance on mixed cue tasks, but not in ++00 cue tasks (β = -.006, SE = .006, t(38) = -.978, p = .334). There were no other significant interactions. The interaction effect between task-type and WMC can be seen in Figure 5.2, and replicates our findings of Chapter 4 that working memory dependent explicit processing is associated with performance in mixed cue but not positive cue tasks. Mixed cue learning is clearly demanding on the individual's limited working memory resources. However, contrary to our expectations, performance in these tasks does not appear to benefit from displaying participants' hypotheses about each cue on screen during learning phases. We expected that the hypotheses notepad would reduce the burden of explicit hypothesis-testing on working memory.

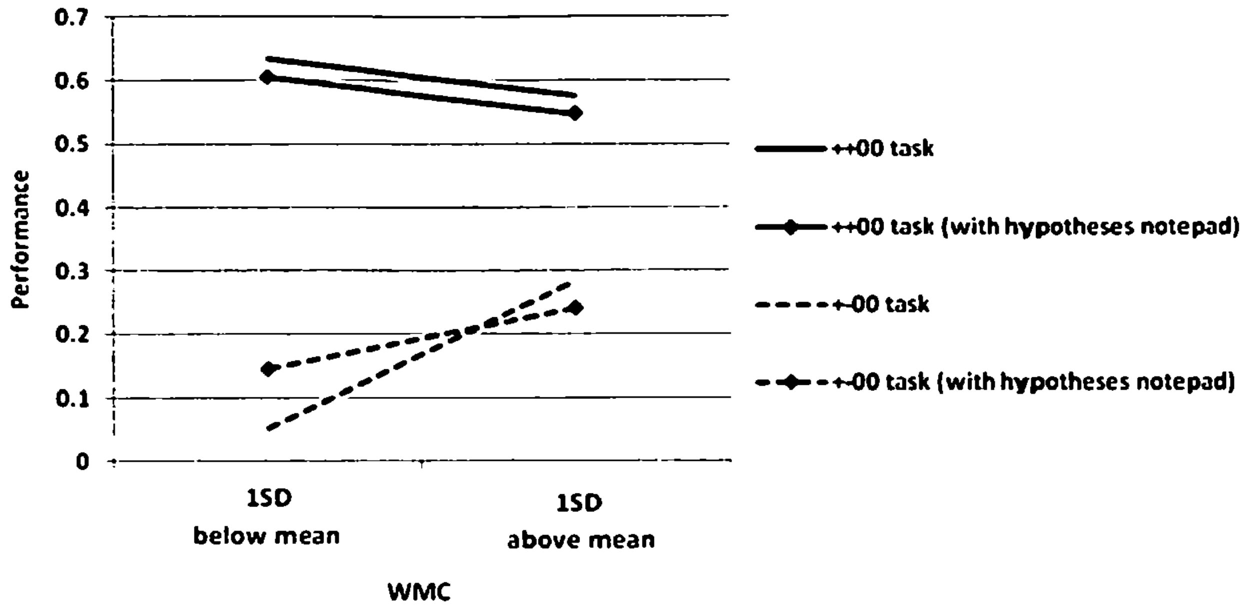

Figure 5.2.

Experiment 7: Interaction effect between task type and WMC on performance scores in the test phase

Performance in the learning phase

The hypotheses notepad task was provided to participants only during learning phases, and may have improved performance on learning trials without affecting performance scores in the test phase. For this reason, we also present analyses of performance scores in the 1st 40 and 2nd 40 learning trials. As with performance in the test phase, performance scores in learning phases were calculated by correlating judgements with criterion values for each participant. Mean group performance scores are displayed in Table 5.1, and were all significant and above zero.

We followed the same multilevel procedure used to analyse performance in the test phase, this time entering task-type (++00 or +-00), condition (standard task and hypotheses notepad task), WMC, and block (1st 40 and 2nd 40 learning trials) as predictors, and a random intercept (χ^2 difference (1) = 82, p <.01) and a random slope for condition (χ^2 difference (1) = 6, p = .01). There was a significant effect of task-type (β = -.401, SE = .066, t(72) = -6.084, p <.001), with participants performing better in the ++00 task (.52) than the +-00 task (.22). Again, there was no effect of the hypotheses notepad on performance (β = -.022, SE = .031, t(193) = -.704, p = .482), confirming that the hypotheses notepad has no effects on performance at any stages of the tasks. There was also no effect of WMC (β = .002, SE = .003, t(72) = .492, p = .624), nor block (β = .035, SE = .029, t(194) = 1.217, p = .225), indicating that learning largely occurred within the 1st 40 learning trials. Consistent with our previous findings there was, however, a significant interaction between task-type and WMC (β = .015, SE = .007, t(72) = 2.178, p = .033). Simple slope analyses confirmed that higher WMC was associated with better performance in learning phases for participants performing the +-00 task (β = .009, SE = .003, t(38) = 2.814, p = .006), but not the ++00 task (β = -.006, SE = .005, t(38) = -1.289, p = .205). While slopes for condition varied across participants there were no significant interactions involving condition. Hence, the slopes did not vary according to WMC, block, or the type of task participants performed.

We concluded in Chapter 4 that while positive cue tasks can be learned well implicitly, learning in tasks containing one or more negative cues requires explicit processing effort. We found that individual differences in WMC were positively associated with performance in mixed cue but not positive cue tasks, suggesting that negative cue learning benefits from working memory dependent explicit processing. The present analysis confirms these findings, this time measuring performance scores in learning phases of the tasks. However, displaying participants' hypotheses selections on screen did not appear to improve performance, even during learning phases.

Explicit knowledge in the test phase

Explicit knowledge scores in the test phase were calculated the same way as in previous chapters, by correlating predicted judgements based on participants' explicit beliefs reported in the cue-rating task with criterion values. By viewing Table 5.2 it can be seen that all groups achieved significant levels of explicit knowledge.

Table 5.2

Experiment 7: Mean group explicit knowledge scores

	1st 40 learning trials		2nd 40 learning trial		Test Trials	
	M	t	M	t	M	t
++00						
Standard task					.38*	5.286
Hypotheses notepad task	.36*	7.663	.35*	5.505	.34*	4.003
+-00						
Standard task					.28*	3.283
Hypotheses notepad task	.25*	4.810	.29*	5.325	.19*	2.178

*$p < .05$

We followed the same multilevel procedure used to analyse performance scores in our analyses of explicit knowledge, entering task-type (++00 or +-00), condition (standard task and hypotheses notepad task), and WMC as predictors, and a random intercept (χ^2 difference (1) = 7, p = .01). There were no significant main effects. Participants appeared to acquire more accurate explicit knowledge in ++00 (.36) than +-00 (.24) tasks. However, this difference did not reach significance ($\beta = -.213$, $SE = .148$, $t(76) = -1.440$, $p = .154$), and is consistent with the findings of our previous studies. While performance levels appear substantially higher in positive than mixed cue tasks, explicit knowledge scores are statistically similar. We found previously that the hypotheses notepad did not appear to improve performance. We also expected that explicit knowledge levels would improve in hypotheses notepad conditions by allowing participants to test their hypotheses on screen against the feedback. However, again there was no effect of the hypotheses notepad ($\beta = -.134$, $SE = .108$, $t(76) = -1.236$, $p = .220$). There was also no association with WMC ($\beta = .010$, $SE = .008$, $t(76) = 1.284$, $p = .203$).

As with our analysis of performance scores, there was however, a significant interaction between task-type and WMC ($\beta = .043$, $SE = .015$, $t(76) = 2.931$, $p = .004$). Simple slope analyses confirmed that higher WMC was associated with more accurate explicit knowledge in the test phase for participants performing the +-00 task ($\beta = .031$, $SE = .010$, $t(38) = 3.060$, $p = .004$), but not for those performing the ++00 task ($\beta = -.013$, $SE = .011$, $t(38) = -1.173$, $p = .248$). The present analysis confirms our earlier findings that higher WMC is associated with more accurate explicit knowledge only when WMC is positively associated with performance, indicating that explicit processing is actively involved in learning tasks that contain a negative cue. As with our analyses of performance scores, explicit knowledge levels after training were not affected by the addition of the hypotheses notepad. If the hypotheses notepad facilitates explicit learning then we would have expected this manipulation to have some effect on participants' explicit knowledge of the tasks.

We also wanted to know how performance scores in the test phase compared with levels of explicit knowledge of the tasks. In our previous experiments participants have been shown to perform better than expected based on their explicit beliefs alone in ++00 tasks, but the reverse effect in +-00 tasks. We thus performed a three-way mixed ANOVA on correlations, using task-type (++00 or +-00) as an independent factor, and condition (standard task and hypotheses notepad task) and measure (performance and explicit knowledge) as within-subjects factors. There was a significant effect of task-type ($F_{(1,74)}$ = 20.335, *MSE* = 5.755, p <.001, partial η^2 = .216), and a marginally significant effect of measure ($F_{(1,74)}$ = 3.838, *MSE* = .545, p = .054, partial η^2 = .049), but no effect of the hypotheses notepad ($F_{(1,74)}$ = .558, *MSE* = .074, p = .457, partial η^2 = .007). However, as in our previous experiments there was also a significant interaction between task-type and measure ($F_{(1,74)}$ = 11.195, *MSE* = 1.590, p = .001, partial η^2 = .131). Two-way mixed ANOVAs, using measure and condition as within-subjects factors, revealed that in ++00 groups participants' performance levels (.59) were significantly higher than their explicit knowledge scores (.37; $F_{(1,37)}$ = 12.131, *MSE* = 1.998, p = .001, partial η^2 = .247). The reverse was shown for the +-00 task, with explicit knowledge scores (.24) instead exceeding performance levels (.17). However, this effect did not reach significance ($F_{(1,37)}$ = 1.145, *MSE* = .137, p = .292, partial η^2 = .030).

Our analysis comparing performance and explicit knowledge levels in positive and mixed cue tasks confirms that performance levels are reliably in excess of explicit knowledge scores in positive cue tasks. In contrast, participants do not perform better than expected in mixed cue tasks based on their explicit beliefs alone. These results confirm that there is a substantial implicit component involved in learning positive cue tasks, raising performance scores above what is expected based on participants' explicit beliefs alone.

Explicit knowledge in the learning phase

In our previous experiments we measured participants' explicit knowledge of the judgement tasks by correlating predicted judgement based on their explicit beliefs about each cues' relevance, with criterion values in the test phase. Participants provided their explicit beliefs in a cue-rating task following completion of the test phase. Hence, explicit knowledge scores were based on a single rating for each cue after training. In the present experiment we were able to measure explicit knowledge levels during learning, using participants' explicit beliefs about the cues on each trial in the learning phase. For this, we entered participants' selections for each cue into a linear model to make predicted judgements on each of the learning trials. These estimate the judgements that participants would have made had they used only their explicit beliefs about each cue. For each trial, we entered 1 if a cue was selected as positive, -1 if selected as negative, and 0 if selected as irrelevant, for each cue. We then correlated predicted judgements with criterion values in the 1^{st} 40 and 2^{nd} 40 learning trials as a measure of explicit knowledge during learning. This provides a measure of how participants would have performed during the learning phase had they used only their explicit hypothesis-testing. Mean group explicit knowledge scores are displayed in Table 5.2, and were all above zero and significant.

Following the multilevel procedure we developed a model that contained task-type (++00 or +-00), WMC, and block (1^{st} 40 and 2^{nd} 40 learning trials) as predictors, and a random intercept (χ^2 difference (1) = 41, p <.01), and random slope for block (χ^2 difference (1) = 11, p <.01). There was no effect of task-type (β = -.107, SE = .082, $t(76)$ = -1.314, p = .193), or block (β = .036, SE = .041, $t(76)$ = .889, p = .377), indicating that explicit knowledge levels were statistically similar in positive (.36) and mixed cue (.27) groups, and did not improve from the 1^{st} 40 (.31) to the 2^{nd} 40 (.32) learning trials. However, the significant random slope for block indicates that there was individual variation in change from the 1^{st} 40 to 2^{nd} 40 learning trials.

Our findings to date indicate that individual differences in WMC are unrelated to explicit knowledge levels in positive cue tasks. However, in the present analysis there was a significant positive association between WMC and explicit knowledge in learning phases across task types (β = .013, SE = .004, $t(76)$ = 2.999, p = .004). Furthermore, there were no significant interactions, indicating that higher WMC was associated with more accurate explicit knowledge in both positive and mixed cue tasks when the hypotheses notepad was provided (see Figure 5.3). This provides our first evidence of explicit processing in positive cue tasks. However, we found no evidence that performance was associated with individual differences in WMC in these tasks. In addition, our previous experiments suggest that learning of positive cues occurs almost entirely implicitly. One possibility is that the hypotheses notepad encouraged participants to think explicitly about their hypotheses selections in positive cue tasks, without influencing their judgements in the task. An alternative possibility is that participants ordinarily think explicitly about positive cue tasks early in training, even though this explicit processing does not influence the judgements they make.

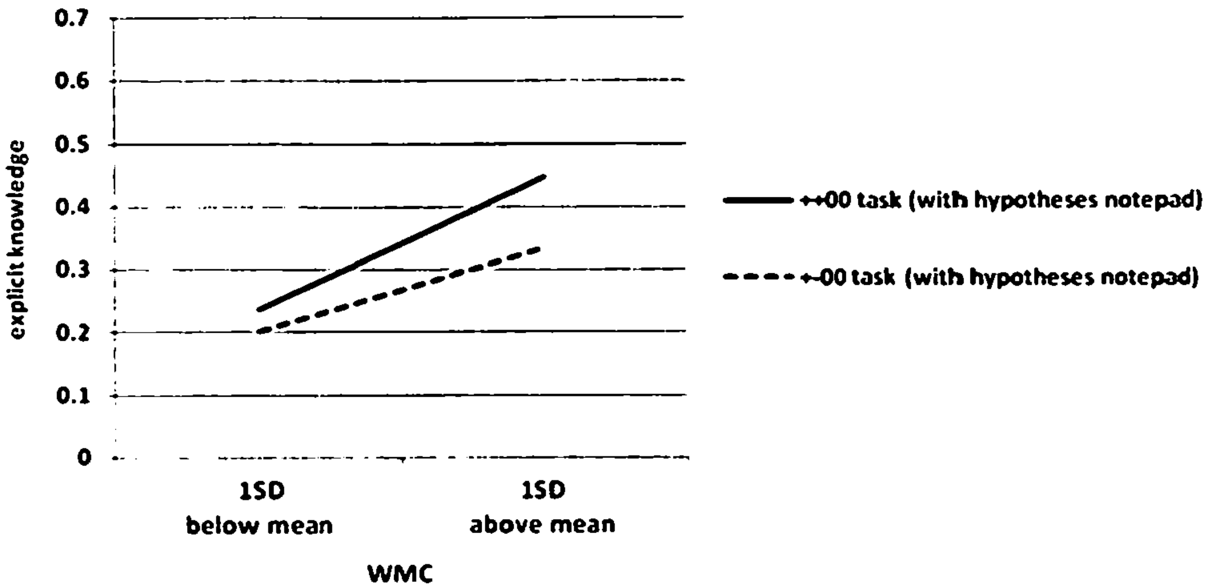

Figure 5.3

Experiment 7: Main effect of WMC on explicit knowledge scores in the learning phase of hypotheses notepad tasks

Explicit knowledge scores derived from participants' hypotheses selections also allows us to compare explicit knowledge levels with performance scores in the learning phase. For this analysis we conducted a three-way mixed ANOVA on correlations, using task-type (++00 or +-00) as an independent factor, and measure (performance and explicit knowledge) and block (1st 40 and 2nd 40 learning trials) as within-subjects factors. There was a significant effect of task-type ($F_{(1,74)}$ = 10.458, *MSE* = 2.293, *p* = .002, partial η^2 = .124), but no effects of block ($F_{(1,74)}$ = .834, *MSE* = .003, *p* = .364, partial η^2 = .011) or measure ($F_{(1,74)}$ = 2.004, *MSE* = .178, *p* = .161, partial η^2 = .026). Again, there was a significant interaction between task-type and measure ($F_{(1,74)}$ = 7.121, *MSE* = .634, *p* = .009, partial η^2 = .088). Two-way mixed ANOVAs, using measure and block as within-subjects factors, confirmed that in the ++00 task performance scores (.49) exceeded explicit knowledge scores (.35; $F_{(1,37)} = 6.224$, $MSE = .742$, $p = .017$, partial $\eta^2 = .144$). In contrast, in the +-00 task we again found that performance scores (.23) did not exceed explicit knowledge scores (.27; $F_{(1,37)} = 1.189$, $MSE = .070$, $p = .283$, partial $\eta^2 = .031$). We found previously that higher WMC was associated with better explicit knowledge in both positive and mixed cue tasks during training, but was only associated with better performance in mixed cue tasks. It now also seems that performance levels in positive cue tasks are substantially higher than what participants could have achieved based on their hypotheses selections alone. This indicates that implicit processes dominate learning in these tasks. One reason why explicit processing did not influence performance in positive cue tasks is that participants could make more accurate judgements if they did not use their explicit hypothesis-testing.

5.1.3. Discussion

Experiment 7 supports our findings of Chapters 3 and 4 and our conclusions so far. While tasks containing only positive relevant cues can be learned well implicitly, effortful explicit processing is required for learning in tasks that contain at least one negative cue. In Experiment 7 we also had the opportunity to measure explicit knowledge levels during learning phases. Our analyses confirmed that higher WMC is associated with better performance in mixed cue but not positive cue tasks, even in learning phases. In positive cue tasks we found that performance scores were well in excess of explicit knowledge levels during training, confirming that implicit learning processes outperform explicit hypothesis-testing during training. The reverse effect was shown in mixed cue tasks containing both a positive and a negative cue. An interesting novel finding of Experiment 7 was that higher WMC was associated with more accurate explicit knowledge during learning phases of both positive and mixed cue tasks when participants were provided the hypotheses notepad. This provides some suggestion that the hypotheses notepad encouraged participants to think explicitly about their selections in positive cue tasks, without affecting their performance. However, the hypotheses notepad did not improve mixed cue learning. We have substantial evidence to suggest that these tasks are learned explicitly, but found no evidence that allowing participants to note their hypotheses on screen had any effect on explicit knowledge or performance levels.

5.2. EXPERIMENT 8

Experiment 7 provided some indication that asking people to note their hypotheses during training encourages explicit hypothesis-testing in positive cue tasks. We found that higher WMC was associated with more accurate explicit knowledge during learning phases of both positive and mixed cue tasks when the hypotheses notepad was provided. However, WMC was only associated with performance in mixed cue tasks, suggesting that participants' explicit hypothesis-testing in positive cue tasks did not influence their performance.

While mixed cue learning appears to benefit from explicit processing and load heavily on working memory, performance was not improved by the hypotheses notepad. One possibility is that participants used the hypotheses notepad only to register their beliefs about each cue, but did not use their selections to help them make their judgements. In Experiment 8, we again provide participants the hypotheses notepad, but this time we present participants the output of their selections on each of the learning trials. In this way, participants view the judgements they could have made had they used their explicit hypothesis-testing. This allows participants to compare the output of their hypotheses selections with outcome feedback on each trial, and how their actual judgements compare with these values.

5.2.1. Method

Participants

Seventy eight undergraduate psychology students at the University of Plymouth participated for course credits, with 39 in each of the two groups.

Design

All participants first completed the same working memory task used in Experiment 7, followed by two judgement tasks. For one group the two judgement tasks each contained two positive and two irrelevant cues (++00; positive cue task), and for a second group contained one positive, one negative, and two irrelevant cues (+-00; mixed cue task). As in Experiment 7, we ensured that the order that cues were presented differed between the two judgement tasks that participants performed. All participants completed both the 'hypotheses notepad' task used in Experiment 7, and a 'hypotheses notepad output' task which provided them the output of their hypotheses selections on each of the learning trials. The order that participants performed the two tasks was randomly determined. All participants completed the cue-rating task following completion of each task (see section 2.3).

Materials and Procedure

All participants first completed the working memory task, followed by two judgement tasks. In both judgement tasks participants were provided the same hypotheses notepad in the learning phase used in Experiment 7, but in one of the tasks they were also provided the output of their hypotheses selections on each trial (see section 7.3. for task instructions). The output of their selections was displayed directly below the hypotheses notepad and to the right of their actual judgement. Output was calculated on each trial by entering participants' hypotheses selections as either 1 for positive, -1 for negative, or 0 for irrelevant as cue weights into a linear model. Output values were then normalised on the same 5 point scale used to present outcome feedback, and were displayed in the same way as participants' actual judgements. The output was provided immediately following participants' judgement on each trial, at the same time as outcome feedback was displayed. All participants completed the cue-rating task following completion of each judgement task.

5.2.2. Results

Performance in the test phase

Mean group performance scores in the test phase are displayed in Table 5.3. All group performance scores were significantly above zero.

Table 5.3

Experiment 8: Mean group performance scores

	1st 40 learning trials		2nd 40 learning trials		Test Trials	
	M	t	M	t	M	t
++00						
Hypotheses notepad	.46*	11.739	.51*	12.791	.49*	12.004
Hypotheses notepad output	.45*	11.676	.50*	12.404	.52*	11.341
+-00						
Hypotheses notepad	.25*	7.060	.34*	7.576	.26*	4.473
Hypotheses notepad output	.20*	4.837	.31*	6.269	.30*	4.973

*p = .05

Following the same multilevel procedure used in Experiment 7, we analysed performance scores in test phases by entering task-type (++00 or +-00), condition (hypotheses notepad and hypotheses notepad output), and WMC as predictors, and a random intercept (χ^2 difference (1) = 37, p <.01). Consistent with our previous studies, there was a significant effect of task-type (β = -.285, SE = .093, $t(78)$ = -3.055, p = .003), with participants performing better in ++00 (.51) than +-00 (.28) tasks. In the present experiment we provided participants the output of their hypotheses selections on learning trials in addition to hypotheses notepad used in Experiment 7. We expected that this would improve performance by encouraging participants to compare the output of their hypotheses selections with the feedback. However, performance levels in test phases were similar (β = .068, SE = .046, $t(78)$ = 1.482, p = .142), regardless of whether participants were provided the output of their hypotheses selections (.41) or not (.38). There was also no association with WMC (β = .008, SE = .005, $t(78)$ = 1.714, p = .091). There was, however, a significant interaction between task-type and WMC (β = .024, SE = .009, $t(78)$ = 2.608, p = .011). Simple slope analyses confirmed that higher WMC was associated with better performance in the +-00 task (β = .021, SE = .007, $t(39)$ = 3.045, p = .004), but not the ++00 task (β = -.003, SE = .006, $t(39)$ = -.518, p = .608).

The interaction effect between task-type and WMC can be seen in Figure 5.4, and confirms that individual differences in WMC are positively associated with performance only when tasks contain at least one negative cue. Surprisingly, however, there was no effect of providing participants the output of their hypotheses selections on performance. We expected that by displaying participants the judgements they could have made had they used their hypotheses selections, this would encourage participants to use the hypotheses notepad to improve their learning. While there was some suggestion that higher WMC participants performed better in the mixed cue task (+-00) when provided the output of their selections (see Figure 5.4), the present analysis provides no evidence that this manipulation had any substantial effects.

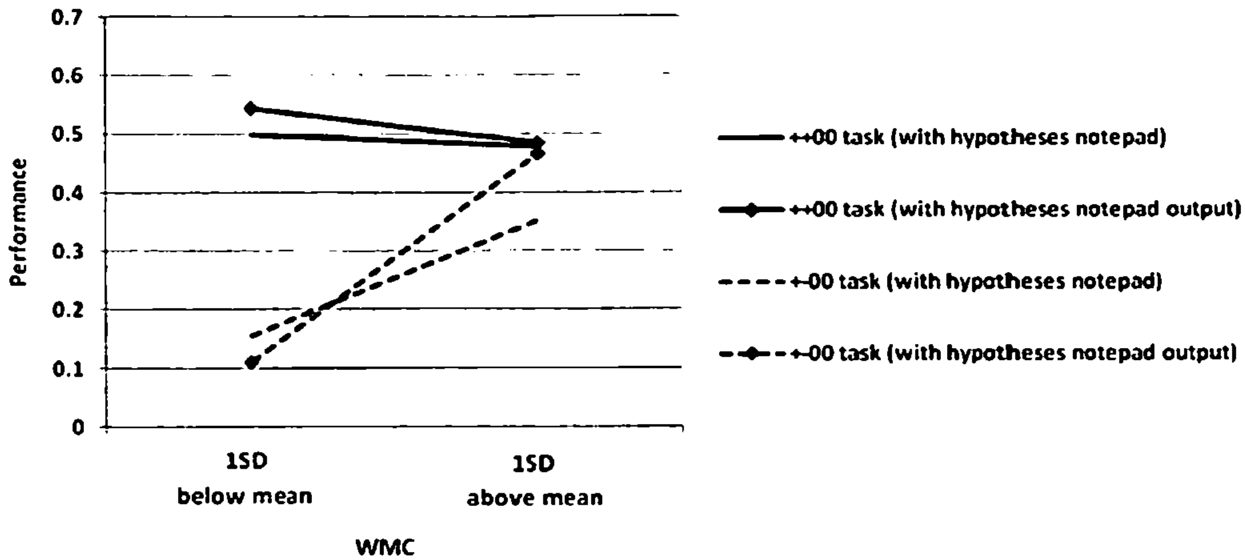

Figure 5.4

Experiment 8: Interaction effect between task type and WMC on performance scores in the test phase

Performance in the learning phase

As in Experiment 7, we also present analyses of performance scores in learning phases of the tasks. This is because the hypotheses notepad and output was provided only on learning trials. Mean group performance scores in the 1st 40 and 2nd 40 learning trials are displayed in Table 5.3, and were all above zero and significant.

Following the multilevel procedure we entered task-type (++00 or +-00), condition (hypotheses notepad and hypotheses notepad output), WMC, and block (1st 40 and 2nd 40 learning trials) as predictors, and a random intercept (χ^2 difference (1) = 67, p <.01). There was a significant main effect of task-type (β = -.268, SE = .061, $t(78)$ = -4.366, p <.001), confirming that participants performed better in ++00 (.48) than +-00 (.28) tasks. There was also a significant effect of block (β = .126, SE = .031, $t(234)$ = 4.139, p <.001), indicating that performance generally improved from the 1st 40 (.34) to the 2nd 40 (.42) learning trials. However, as in our analysis of performance scores in test phases, providing participants the output of their hypotheses selections in addition to the hypotheses notepad did not improve performance on learning trials (β = -.026, *SE* = .031, $t(234)$ = -.857, p = .392). There was also no association with WMC (β = .005, *SE* = .003, $t(78)$ = 1.701, p = .093). Replicating our previous findings, there was an interaction between task-type and WMC (β = .014, *SE* = .006, $t(78)$ = 2.225, p = .029). This interaction was confirmed by simple slope analyses, which indicated that higher WMC was associated with better performance in +-00 (β = .012, *SE* = .004, $t(39)$ = 3.126, p = .003), but not in ++00 (β = -.001, *SE* = .005, $t(39)$ = -.242, p = .810) tasks.

There was also an interaction between WMC and condition (β = .009, *SE* = .003, $t(234)$ = 2.865, p = .005). Simple slope analysis indicated that higher WMC was associated with better performance across task types when participants were provided the output of their hypotheses selections (β = .009, *SE* = .004, $t(78)$ = 2.172, p = .033) but not when provided only the hypotheses notepad task (β <.001, *SE* = .004, $t(78)$ = -.041, p = .967). However, by viewing Figure 5.5 it can be seen that the association between WMC and performance in hypotheses notepad output conditions was substantially higher for the +-00 (β = .015, *SE* = .004, $t(37)$ = 4.028, p <.001) than the ++00 (β = .002) task, which was far from significant (*SE* = .004, $t(37)$ = .444, p = .659). While there may have been some tendency for individuals of higher WMC to perform better in ++00 tasks when provided the output of their hypotheses selections, this effect appears to be relatively small.

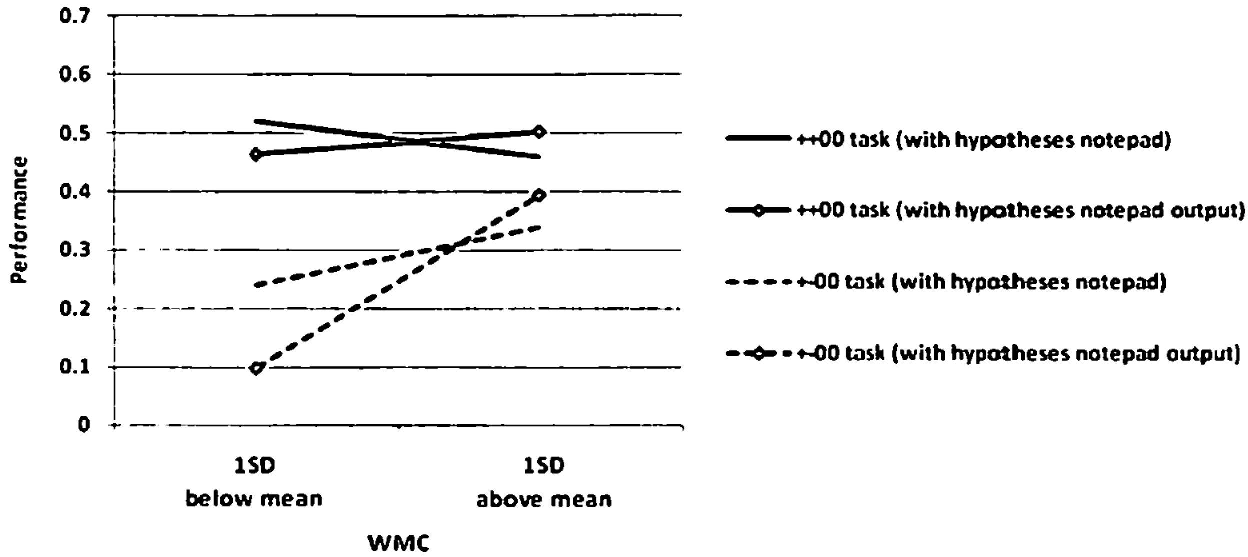

Figure 5.5

Experiment 8: Interaction effect between condition and WMC on performance scores in the learning phase

Providing participants the output of their hypotheses selections was designed to improve learning at least in mixed cue tasks that is shown to load heavily on working memory. We expected that performance would improve in hypotheses notepad output tasks, and to especially benefit individuals of lower WMC. In contrast, performance appears generally unaffected even in mixed cue tasks. However, there was some suggestion that the relationship between WMC and performance was affected by the addition of output in mixed cue tasks (see Figure 5.5).

Explicit knowledge in the test phase

Mean group explicit knowledge scores in the test phase are displayed in Table 5.4. All group scores were significantly above zero.

We followed the same multilevel procedure used in our previous analyses, this time entering task-type (++00 or +-00), condition (hypotheses notepad and hypotheses notepad output), and WMC as predictors, and a random intercept (χ^2 difference (1) = 15, $p < .01$) and random slope for condition (χ^2 difference (1) = 7, $p = .01$). The random slope for condition indicates that the effects of condition varied between individuals. There was a marginally significant positive effect of WMC ($\beta = .020$, $SE = .011$, $t(78) = 1.909$, $p = .060$), but no effect of the type of task participants performed ($\beta = .332$, $SE = .206$, $t(78) = 1.610$, $p = .111$). As in our analyses of performance scores, providing participants the output of their hypotheses selections in addition to the hypothesis notepad had no main effect on explicit knowledge ($\beta = -.011$, $SE = .139$, $t(78) = -.082$, $p = .935$). However, consistent with our earlier findings, there was a significant interaction between task-type and WMC ($\beta = .058$, $SE = .020$, $t(124) = 2.859$, $p = .005$). Simple slope analysis confirmed that this was because higher WMC was associated with more accurate explicit knowledge in the +-00 task ($\beta = .049$, $SE = .016$, $t(39) = 2.989$, $p = .005$), but not in the ++00 task ($\beta = -.003$, $SE = .014$, $t(39) = -.237$, $p = .814$). Hence, explicit learning in tasks that contain a mixture of positive and negative cues is heavily demanding on WM, such that lower WMC individuals acquire less accurate explicit knowledge. However, providing participants the output of their hypotheses selections does not appear to reduce the demands of explicit learning on working memory in these tasks.

As in our previous experiments, we also compared explicit knowledge and performance levels in the test phase. For this analysis we performed a three-way mixed ANOVA on correlations, using task-type (++00 or +-00) as an independent factor, and condition (hypothesis notepad and hypothesis notepad output) and measure (performance and explicit knowledge) as within-subjects factors. There was no significant effect of task-type ($F_{(1,76)}$ = .571, MSE = .267, p = .452, partial η^2 = .007), or condition ($F_{(1,76)}$ = .008, MSE = .001, p = .930, partial η^2 <.001), but a significant effect of measure ($F_{(1,76)}$ = 6.103, MSE = .594, p = .016, partial η^2 = .074). As expected, there was also a significant interaction between task-type and measure ($F_{(1,76)}$ = 21.765, MSE = 2.119, p <.001, partial η^2 = .223). Two-way mixed ANOVAs, using condition and measure as within-subjects factors, confirmed that in the ++00 task performance scores (.51) were significantly higher than explicit knowledge scores (.26; $F_{(1,38)}$ = 22.638, MSE = 2.479, p <.001, partial η^2 = .373). In the +-00 task, performance (.28) did not exceed explicit knowledge scores (.36; $F_{(1,38)}$ = 2.752, MSE = .235, p = .105, partial η^2 = .068). There were no other significant interactions. These findings again support our conclusions that a substantial implicit component is involved in learning positive cue tasks. In these tasks, implicit learning processes appear to boost performance above what is attainable based on participants' explicit beliefs alone. In contrast, implicit processing appears to contribute little to learning of mixed cue tasks, with performance scores not exceeding explicit knowledge levels.

Table 5.4

Experiment 8: Mean group explicit knowledge scores

	1st 40 learning trials		2nd 40 learning trials		Test Trials	
	M	t	M	t	M	t
++00						
Hypotheses notepad	.25*	4.313	.38*	6.237	.22*	2.744
Hypotheses notepad output	.26*	4.688	.38*	5.281	.28*	3.092
+-00						
Hypotheses notepad	.27*	5.312	.43*	7.509	.44*	5.079
Hypotheses notepad output	.32*	6.900	.45*	7.760	.28*	2.848

*p = .05

Explicit knowledge in the learning phase

We measured explicit knowledge levels in the 1st 40 and 2nd 40 learning trials in the same way as in Experiment 7 for hypotheses notepad groups. This was done by correlating predicted judgements based on participants' hypotheses selections with criterion values. In tasks where participants were provided the output of their selections this meant correlating the output they received with criterion values. Mean group scores are displayed in Table 5.4, and were all significantly above zero.

In our analyses of explicit knowledge scores in learning phases we followed the multilevel procedure, entering task-type (++00 or +-00), condition (hypotheses notepad and hypotheses notepad output), WMC, and block (1st 40 and 2nd 40 learning trials) as predictors, and a random intercept (χ^2 difference (1) = 36, p <.01). There was no effect of task-type (β = .083, SE = .112, $t(78)$ = .741, p = .461), but a significant effect of block (β = .339, SE = .067, $t(233)$ = 5.089, p <.001), indicating that explicit knowledge levels increased from the 1st 40 (.28) to the 2nd 40 (.41) learning trials. A significant positive effect of WMC (β = .014, SE = .006, $t(78)$ = 2.418, p = .018) demonstrated that higher WMC was associated with more accurate explicit knowledge in the learning phase. As in our previous analysis, there was no effect of providing participants the output of their hypotheses selections on explicit knowledge levels in learning phases (β = -.015, SE = .067, $t(233)$ = -.222, p = .825).

In Experiment 7 we found that higher WMC was associated with more accurate explicit knowledge across task types when participants were provided the hypotheses notepad. The present analysis instead yielded a significant interaction between task-type, WMC, and condition (β = .012, SE = .004, $t(311)$ = 3.238, p = .001). Simple slope analyses indicated that WMC was significantly positively associated with explicit knowledge only in the hypotheses notepad output task for those performing the +-00 task (β = .038, *SE* = .009, $t(37)$ = 4.096, p <.001). WMC was not significantly associated with explicit knowledge in the hypotheses notepad condition for participants performing the +-00 task (β = .009, *SE* = .010, $t(37)$ = .899, p = .374), nor in the hypotheses notepad (β = .004, *SE* = .011, $t(37)$ = .341, p = .735) and hypotheses notepad output conditions (β = .005, *SE* = .008, $t(36)$ = .608, p = .547) for participants performing the ++00 task. Hence, we did not replicate our findings of Experiment 7 that higher WMC is associated with more accurate explicit knowledge in positive cue tasks when participants are provided the hypotheses notepad. The present analysis suggests that higher WMC is only associated better explicit knowledge for participants performing the mixed cue task, and only when provided the output of their selections.

In both the judgement tasks that participants performed they provided trial-by-trial explicit ratings of each cue's relevance in the learning phase. Hence we also compared performance and explicit knowledge levels in learning phases. This was done by performing a four-way mixed ANOVA on correlations, using task-type (++00 or +-00) as an independent factor, and condition (hypothesis notepad and hypothesis notepad output), measure (performance and explicit knowledge), and block (1st 40 and 2nd 40 learning trials) as within-subjects factors. There was no effect of task-type ($F_{(1,76)}$ = 2.335, *MSE* = .914, p = .131, partial η^2 = .030), or condition ($F_{(1,76)}$ = .002, *MSE* <.001, p = .961, partial η^2 <.001), but a significant effect of block ($F_{(1,76)}$ = 33.272, *MSE* = 1.697, p <.001, partial η^2 = .304), and an interaction between block and measure ($F_{(1,76)}$ = 6.837, *MSE* = .141, p = .011, partial η^2 = .083). There was also a significant interaction between task-type and measure ($F_{(1,76)}$ = 30.956, *MSE* = 2.516, p <.001, partial η^2 = .289). Three-way mixed ANOVAs, using condition, measure, and block as within-subjects factors confirmed that in the ++00 task performance scores (.48) were significantly higher than explicit knowledge scores (.32; $F_{(1,38)} = 16.826$, $MSE = 2.005$, $p < .001$, partial $\eta^2 = .307$). In contrast, performance in the +-00 task (.28) was significantly lower than explicit knowledge levels (.37; $F_{(1,38)} = 15.769$, $MSE = .684$, $p < .001$, partial $\eta^2 = .293$). There were no other significant interactions.

The present analysis confirms that participants perform far better in positive cue tasks than expected based on their hypotheses selections alone. This suggests that a substantial implicit component contributes to judgement in positive cue tasks even when participants are provided the output of their selections. This meant that on learning trials participants' judgements were substantially more accurate in predicting the criterion than those based on their hypotheses selections, even when they were displayed with both their judgements and the output of their selections.

5.2.3. Discussion

In Experiment 7 we found that individual differences in WMC positively predicted explicit knowledge levels in learning phases of both positive and mixed cue tasks when participants were provided with the hypotheses notepad. This provided some indication that the hypotheses notepad may have elicited explicit thinking in positive cue tasks, without affecting participants' performance on the task. However, we did not replicate this effect in Experiment 8. Furthermore, we found that providing the output of participants' hypotheses selections in addition to the hypotheses notepad had no substantial effects on explicit learning in either positive or mixed cue tasks.

We did find that WMC is unrelated to performance in positive cue tasks, and that performance levels are far in excess of explicit knowledge levels in these tasks even during learning phases. This was not the case for tasks that contained a mixture of positive and negative cues. Instead, performance scores did not exceed levels of explicit knowledge in mixed cue tasks, and learning is heavily demanding on WMC. Hence, the findings of Experiment 8 confirm our conclusions to date. Learning in positive cue tasks occurs largely implicitly, and is unrelated to individual differences in WMC. This explains why performance levels are higher than what participants can achieve based on their explicit beliefs alone. Learning in tasks that contain one or more negative cues, in contrast, benefits from explicit processing effort. For this reason, performance is often poorer than explicit knowledge, and both measures are associated with individual differences in WMC.

5.3. GENERAL DISCUSSION

The experimental studies of the present chapter were designed primarily to explore whether judgement in difficult mixed cue tasks could be improved. Our secondary aims were to measure explicit knowledge levels during learning phases and compare these with measures of performance and individual differences in working memory capacity (WMC). Using a slightly different measure of WMC in Experiments 7 and 8 we confirmed our earlier findings that explicit processing ordinarily contributes to learning in mixed cue (+-00) but not positive cue (++00) tasks. This was indicated by a significant positive association between WMC and performance only in mixed cue tasks. We also found that in both learning and test phases of positive cue tasks performance scores were well in excess of explicit knowledge levels, even when participants were shown the output of their explicit hypothesis-testing on learning trials. This provides strong evidence that implicit knowledge guides judgement during learning in tasks that contain only positive relevant cues. In tasks where participants were provided the output of their hypotheses selections this meant that participants performing positive cue tasks could see that their judgements more accurately predicted the criterion than their explicit hypothesis-testing. The reverse effect was shown in mixed cue tasks, with explicit knowledge levels consistently in excess (though not always significantly) of performance scores. The findings of Experiments 7 and 8 thus confirm our conclusions of earlier experiments that while judgement tasks containing only positive relevant cues can be learned well implicitly, learning in tasks containing one or more negative cues benefits from explicit processing.

When tasks contain a mixture of positive and negative cues, explicit learning is heavily demanding on working memory, such that individuals of lower WMC perform poorly. In the present experiments we attempted to improve explicit learning in mixed cue tasks by providing participants a hypotheses notepad on screen for them to note their hypotheses about each cue during learning phases. While performance did not improve in positive or mixed cue tasks with the addition of the hypotheses notepad, Experiment 7 provided some suggestion that explicit processing was elicited in positive cue tasks. This was indicated by a positive association between WMC and explicit knowledge of both positive and mixed cue tasks. However, we did not replicate this effect in Experiment 8. Furthermore, WMC was only associated with performance levels in mixed cue tasks, suggesting that even when participants are encouraged to think explicitly during positive cue tasks implicit processing continues to dominate performance.

By displaying participants' hypotheses on screen it was expected that they would use their selections to help them make their judgements. In this way, participants would not need to hold in mind their explicit beliefs about each cue, reducing the demands of explicit judgement on working memory. Hence, we expected that participants of lower WMC would benefit most from this manipulation, reducing the association between WMC and performance in mixed cue tasks. We found no such effect. Second, the hypotheses notepad was designed to improve explicit hypothesis-testing by allowing participants to test their hypotheses against feedback without having to hold their hypotheses in mind. We expected the association between WMC and explicit knowledge to be reduced in mixed cue tasks by this manipulation, and for explicit knowledge to improve. Again, our findings suggest that this was not the case. One possible explanation is that participants used the hypotheses notepad task to register their beliefs about the cues, but did not use the notepad to test hypotheses or help them make their judgements. In Experiment 8, we provided participants the output of their hypotheses selections on learning trials to encourage participants to use the notepad during learning. Again, performance was not improved in mixed cue tasks. While displaying participants the output of their hypotheses selections was designed to encourage participants to use their selections, this task may have been too demanding on participants' working memory resources. Making hypotheses selections as well as judgements on each trial, and assessing how their judgements related to outcome feedback and the output of their hypotheses selections, is likely to be very taxing on the individual's working memory resources. We know already that learning in mixed cue tasks loads heavily on working memory. For this reason, it is likely that even individuals of higher WMC were unable use the hypotheses notepad task effectively in their explicit learning of the task. In line with this view, there was some suggestion that individuals of lower WMC performed more poorly in mixed cue tasks when provided the output of their hypotheses selections (see Figure 5.5).

In sum, the present chapter confirms our conclusion of Chapters 3 and 4. While positive cue learning occurs implicitly, learning in tasks that contain one or more negative cues benefits from explicit processing effort. This was indicated by a positive association between WMC and measures of performance and explicit knowledge in mixed cue but not positive cue tasks. One exception was that providing the hypotheses notepad appeared to influence explicit knowledge of positive cue tasks during training without affecting performance. However, we did not replicate this effect. It therefore appears that learning in positive cue tasks ordinarily occurs implicitly, and even when participants are encouraged to think explicitly about the task.

Chapter 6

General Discussion

A common view in research on multiple cue judgement is that judgements are made in a conscious and explicit manner. According to this account, when provided outcome feedback in response to judgement, people test hypotheses against the feedback until a hypothesis is confirmed by the cue-outcome values they observe. Once a hypothesis is confirmed, people begin to abstract explicit beliefs about the importance of each cue in predicting the criterion, and use their beliefs to inform their judgements. In this way, learning and judgement is deliberative and cognitively demanding.

On the view that judgement is explicit, performance in judgement tasks should be fully explained by people's explicit beliefs about the importance of each cue. However, a number of findings cause problems for this account. High levels of performance in judgement tasks are often observed without the individual acquiring significant levels of explicit knowledge. Furthermore, people are repeatedly shown to demonstrate poor levels of insight into their judgement policies. In such cases, judgement appears to be influenced in part by knowledge that is implicit and not available for verbal report. Dual process theories of thinking posit that both automatic implicit and controlled explicit processes can contribute to learning and judgement in a range of cognitive tasks. Whereas explicit processing is effortful and generates knowledge that is consciously accessible, implicit knowledge is not available for verbal report, but may nevertheless influence performance on a task. The dual process framework allows us to assess the influence of both implicit and explicit knowledge on judgement from multiple cues.

In the present thesis, we introduced dual process methods not previously used in the judgement literature to measure the contribution of implicit and explicit processes to multiple cue judgement. Our studies provide substantial evidence that both kinds of processes are involved in learning in multiple cue environments. Whether learning is implicit or explicit depends on the types of cues on which participants are trained. However, our findings indicate that when explicit processing is involved in learning, it may not guide judgement after extensive practice with a task. I thus concluded that in some types of multiple cue environments explicit processing is required for learning, but that these processes only contribute to early stages of learning.

6.1. COGNITIVE PROCESSES IN LEARNING AND JUDGEMENT

One possibility is that the cognitive processes that underlie judgement from multiple cues are explicit and under conscious control. This appears to be the assumption made by some

MCPL theorists (Brehmer, 1974, 1980; Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Lagnado, Newell, Kahan, & Shanks, 2006). For example, Juslin, Olsson, and Olsson (2003) explain that "people are assumed to abstract explicit representations of the cue-criterion relations, which signify the importance of each cue by a cue weight." Cues that are considered more important are given a higher weighting and accordingly impact more on judgement than cues that receive a lower weighting. The weighted cue values are then summed in a conscious manner to produce a weighted average for judgement (Brehmer, 1994; Cooksey, 1996). Similarly, Brehmer (e.g. 1980) assumes that people's hypothesis testing strategies can be retrieved by asking them to report their strategies verbally, and Einhorn et al. (1980) use similar methods of verbal report. While an argument that these theorists assume all judgement is explicit can not be made, the possibility that judgement is guided also by implicit learning does not appear to have been taken seriously.

It is important to note that when judgement is explicit, cues are most likely attended to sequentially rather than holistically, which would involve integrating all cue values with explicit beliefs simultaneously. Analyses of verbal protocols suggests that people consider each cue one at a time, and adjust their estimate according to their explicit beliefs about the importance of each one (Einhorn & Hogarth, 1981; Einhorn et al, 1979; Kleinmuntz, 1975). Juslin, Karlsson, and Olsson (2008) recently formalised a model of explicit judgement that describes this serial process. They suggest that people rank the cues according to their explicit beliefs about each one's relevance. They then consider the cues one at a time in order of their importance, by which a running estimate is adjusted in light of the value of each cue until all are considered. At which point, a final judgement is made. Hence, while linear regression models that combine cue weights with cue values independently for each cue are shown to capture people's judgement policies (Brehmer, 1994; Cooksey, 1996; Ullman & Doherty, 1984), explicit judgement is likely to be slow, effortful, and sequential.

In contrast with explicit theories of judgement, people are often shown to have little insight into their own judgement policies (Arkes, 1981; Brehmer, 1984; Harries, Evans, Dennis, & Dean, 1996; Roose & Doherty, 1976). Poor levels of self-insight are found even when people are asked only to report the cue they believe is most important (Brehmer & Brehmer, 1987). Similarly, when learning from outcome feedback in novel environments participants are shown to perform well, even when they have insufficient explicit knowledge of a task to have performed significantly above chance (Evans, Clibbens, Cattani, Harries, & Dennis, 2003). Hence, people may also be influenced by knowledge that is unconscious and inaccessible to verbal report.

In many of the studies that have demonstrated poor levels of self-insight or explicit knowledge, participants provide a single rating of each cue's relevance following completion of a task. There are a number of potential problems with this methodology. Firstly, when explicit cue ratings made on a single occasion are compared with subjective cue weights derived from multiple trials, participants cue ratings may be more susceptible to error (e.g. Harries & Harvey, 2000). This is because whereas explicit beliefs are derived from a single rating of each cue, cue weights are averaged across many trials. Secondly, when the relevance of cues are rated sometime after learning has occurred, the conscious knowledge that guided learning may have been forgotten (Ericsson & Simon, 1980). This issue relates to the proceduralisation of explicit skills (Anderson, 1993). Through practice, conscious knowledge can be converted into implicit procedures, after which, skilled performance no longer depends on explicit beliefs (Brown & Carr, 1989). Hence, some conscious knowledge may be forgotten by the time verbal report measures are administered. Consistent with this view, when ratings are made at regular intervals during learning in multiple cue tasks, explicit beliefs are shown to correspond better with the actual cue validities (Lagnado et al, 2006). Explicit knowledge is better still when cues are rated on every trial (Harries & Harvey, 2000), indicating that these knowledge levels can be underestimated by single session post-task measures.

6.1.1. The dissociation between performance and explicit knowledge

Explicit theories of multiple cue judgement fall more generally within single process models of cognition. These models assume that there is a single learning process that is explicit and under conscious control (Lovibond & Shanks, 2002; Mitchell, De Houwer, & Lovibond, 2009; Shanks & St. John, 1994). According to these accounts, learning leads to knowledge that is conscious and available for verbal report. In Chapter 1 (see section 1.2.), I discussed a number of findings that cause problems for single process models of thinking, namely that people often demonstrate good performance of complex cognitive tasks despite being unable verbally to describe what they have learned (Gomez & Schvaneveldt, 1994; Mathews et al, 1989). Learning is also shown to occur in the absence of explicit awareness that anything was learned at all (Brooks, 1978; Dulany, Carlson, & Dewey, 1984; Reber, 1967, 1993). What these studies show is a dissociation between performance and explicit knowledge in cognitive tasks. Consistent with dual process theories of thinking, I concluded that both an implicit and an explicit component are likely to be involved in learning.

In our studies, we measured explicit knowledge of multiple cue tasks and compared this with performance levels. In each experiment we estimated how participants would have performed in a test phase had they used only their explicit beliefs about each cue's relevance. We called this measure explicit knowledge. It follows that if multiple cue judgement can be accounted for by a single process model, performance should be explained entirely by people's explicit beliefs. That is, performance levels should not exceed explicit knowledge of a task. Across all our experiments we instead found that in some task types, performance scores were substantially higher than expected based on participants' explicit knowledge alone. This was the case when tasks contained only positive or negative relevant cues (++00 and --00 tasks). Even if participants had consistently used their explicit beliefs when making judgements, they could not have reached the performance scores they achieved. This provides a striking dissociation between performance and explicit awareness, such that judgement in some types of multiple cue tasks appears to be guided by knowledge that is implicit and inaccessible to consciousness.

On the view that learning can be explained entirely by a single learning system, some theorists would likely dispute the validity of our explicit knowledge measures. A common argument against evidence of dissociation is that measures based on verbal report can be insensitive to the full extent of explicit learning (Harries & Harvey, 2000; Shanks & St. John, 1994). Hence, evidence of implicit learning could be falsely claimed. However, we found that in some types of tasks explicit knowledge levels instead exceeded performance. In Experiments 1, 2, and 3, the type of task participants performed interacted with our measures of performance and explicit knowledge. In mixed cue tasks (+-00) that contained both a positive and a negative cue, performance did not exceed what was expected based on participants' explicit beliefs. If our verbal report measures simply failed to tap into the full extent of participants' conscious knowledge, then we should have found performance levels to exceed our estimates of explicit knowledge in all task types.

Recall that a single cue rating task administered retrospectively may underestimate explicit knowledge levels. In Experiments 7 and 8 our measures were also derived from participants' trial-by-trial cue ratings during learning phases. This way, participants' explicit beliefs were recorded on the same trials used to derive their subjective cue weights. We again replicated our findings of dissociation. When participants were trained on cues that were positively related to the criterion, performance was well in excess of explicit knowledge even during training. Our findings indicate that both implicit and explicit processes can contribute to complex skill learning, but depend on the types of tasks on which participants are trained.

Dissociation between measures of performance and explicit knowledge are difficult to reconcile with models that describe a single learning process. Our findings are also in contrast with other studies of multiple cue judgement. For instance, Lagnado et al. (2008) found that participants' use of multiple cues and their explicit beliefs about each one's relevance both increasingly approximated the actual cue validities during training. They concluded from these findings that judgement is explicit. Experiments 7 and 8 of the present thesis generally confirmed their findings, by showing that performance and explicit knowledge often did improve during learning phases. However, in our experiments we also compared these measures by estimating how participants could have performed based on their explicit beliefs alone.

Our findings of dissociation are more consistent with models that describe separate learning systems. Similarly, studies of category learning demonstrate that some but not all types of tasks are learned explicitly. When simple rules determine category membership, learning appears to be explicit (Price, 2005; Ashby & Maddox, 2005). For example, a participant may learn to attend to the size and width of a square shape when deciding whether a probe belongs to category A or B. For tasks of this type, people are shown to report accurately how they made their decisions (Ashby & O'Brien, 2005). Rules that describe the rotation of a line within a circle that varies in size, on the other hand, are perhaps less easily verbalised. Accordingly, learning often occurs in absence of explicit knowledge of the category rule (Waldron & Ashby, 2002). In these studies, performance and explicit knowledge is dissociated by the type of task on which people are trained, indicating in accordance with our findings that separate cognitive systems (implicit and explicit) are involved in learning complex structures in the environment. Recently, these processes have also been mapped onto distinct neurological areas in the brain (Kolb & Wishaw, 1990; Smith & Grossman, 2008).

6.1.2. Self-insight

My discussion to this point has concerned the correspondence between performance and explicit knowledge measures in learning environments. Verbal reports measures have a long history in psychology, and are widely used within the social cognitive literature (e.g. Hofman, Gschwendner, & Schmitt, 2005). A consistent finding is that implicit tests (such as of attitudes and personality self-concept) often do not correspond well with verbal report (Blair, 2001; Dovidio, Kawakami, & Beach, 2001; Hofman, Gawronski, Geschwender, Le, & Schmitt, 2005). For some theorists this is evidence that an automatic implicit component of social cognition exists, that is inaccessible to consciousness (Nisbett & Wilson, 1977; Smith & DeCoster, 2000; Wegner & Bargh, 1998; Wilson, 2002).

Unaware of the true causes of their behaviour, people are likely to confabulate explanations for their responses in cognitive tasks (Evans & Over, 1996; Stanovich, 2004; Wilson & Dunn, 2004). A good example of this is provided by Nisbett and Schachter (1966). Before exposing participants to a series of electric shocks of increasing intensity, in one of their studies they gave a group of participants a placebo pill believed to elicit symptoms similar to those of the shocks (e.g. hand tremor). While those participants who were given the pill were able to withstand electric shocks of far greater intensity, they appeared completely unaware that the pill influenced their behaviour. Even when participants were asked if the pill could have had any effects on their improved performance, only a small minority attributed their behaviour to the true cause. Interestingly, however, participants were quite confident of the possibility that entirely unrelated factors may have led to their endurance during shock treatment. When the true causes of behaviour are inaccessible to consciousness, people appear to rationalise their behaviour. The 'selection task', used to study conditional reasoning (discussed in more detail in section 1.2.4.), provides further evidence of confabulation in verbal report. In these tasks, participants are required to choose from a set of cards (with numbers or letters printed on each) those that logically falsify a conditional rule. Rather than reason the task logically, people appear overly biased towards selecting the cards that are mentioned in the rule (Evans, 1998, 2003). Furthermore, when asked how they made their card choices participants report attempting to falsify the rule, and show no awareness of how the cards stated in the rule influenced their decisions (Evans & Wason, 1976; Wason & Evans, 1975).

Confabulation may, in part, account for experts' reports about their own judgement policies. Recall that experts are often shown to demonstrate poor levels of self-insight (e.g. Wigton, 1996), suggesting that some of expertise is not consciously accessible. Interestingly, experts often report using a large number of cues when making judgements, but actually use only a small subset of the reported cues (Evans, Harries, & Dean, 1995). Evans et al. (1995) suggest that their physicians listed the cues they believed were important based on medical training, but were unaware that they did not use all the cues they listed. Hence, experts may rationalise to some extent about their own judgement policies by drawing on general knowledge that is irrelevant to their behaviour (Harries, Evans, Dennis, & Dean, 1996). Confabulation may also account partly for participants explicit ratings in our studies. However, we did find that explicit knowledge levels were consistently above zero and significant, across task types. This was the case even in positive cue tasks that appeared to be learned implicitly. Hence, participants in our studies had acquired some explicit knowledge, even if this did not contribute to performance. One possibility is that people do engage explicitly during training, but acquire only limited conscious knowledge when a task can be learned implicitly. An alternative explanation is that explicit knowledge is "extracted" from implicit knowledge during training (Sun, Merrill, & Peterson, 1998). Sun, Merrill, and Peterson (2001) provide a computational model that describes how this process may occur. Accordingly, explicit learning often tends to lag behind implicit learning (Bowers, Regehr, Balthazard, & Parker, 1990; Reber & lewis, 1977). For instance, Stanley, Mathews, Buss, and Kotler-Cope (1989) monitored participants' verbal reports whilst performing dynamic control tasks (see section 1.2.2. for more details) and found that while good performance was achieved early in training, participants explicit knowledge of the tasks only improved towards the end of training. As in our studies, some explicit learning does appear to occur in these tasks, but does not always contribute to performance.

6.1.3. The role of working memory

Individual differences in working memory capacity (WMC), or high correlates of it such as general intelligence, are related to performance in a range of cognitive tasks, including logical reasoning (Stanovich, 1999), probability judgement (West, Toplak, & Stanovich, 2008), problem solving (Hambrick & Engle, 2003), and inductive reasoning (Feeney, 2007). Across these areas of research, effortful thinking is shown to be heavily demanding on working memory.

Studies in the reasoning and judgement literatures often pit implicit heuristic responses against analytic reasoning, and emphasise the importance of cognitive capacity in inhibiting and replacing erroneous intuitive judgements (Evans, 2007; Kahneman, 2000; Kahneman & Frederick, 2002; Stanovich & West, 2008). The *belief bias effect* in syllogistic reasoning illustrates this point (Evans, Barston, & Pollard, 1983). When told that 'All living things need water' and that 'Roses need water', one may then be asked whether it necessarily follows that 'Roses are living things'. The believability of the conclusion activates a heuristic response based on prior belief and knowledge (roses are living things) which is in conflict with the propositions, but is falsely endorsed by the majority of people

(Evans, 2003; Markovitz & Nantel, 1989). For the individual to reason logically about the propositions an intuitive response must be suppressed, but this process can load heavily on working memory (Engle, 2002; Geary, 2005). Hence, overriding heuristic processing requires a conscious mental effort and may be beyond the cognitive capacity of some individuals, such that some may not have the resources to carry out the override (De Neys, 2006; Stanovich, 2004). However, even if an individual detects the need to engage in effortful thinking, they must also have the working memory resources to sustain the inhibitory process whilst reasoning about a task (Kahneman, 2000). This process too, is demanding on working memory (Stanovich & West, 2008). If an individual has the cognitive resources to suppress an intuitive judgement, but is unable to decouple their explicit reasoning from beliefs they may be unsuccessful in overriding and replacing intuition (Evans, 2007, 2008).

In our studies, participants were required to learn how a number of cues were related to a criterion, rather than rely on prior belief or knowledge. In Chapters 4 and 5, we measured individual differences in WMC, and correlated these with performance and explicit knowledge levels in judgement tasks. Our results provided clear evidence that explicit processing is involved in learning tasks that contain one or more negative cues, but not in tasks containing only positive relevant cues. This was indicated by a positive association between our measures of WMC and performance in negative cue and mixed cue tasks, but not positive cue tasks. Moreover, when WMC predicted performance, high span individuals also acquired more accurate explicit knowledge (Experiments 4, 5, 7, and 8), confirming that these tasks benefit from explicit processing effort.

When explicit hypothesis-testing is involved in learning in multiple cue environments, this process is clearly demanding on working memory. The individual must consciously *generate* hypotheses about the cue-criterion relations, and *evaluate* hypotheses against the feedback. Both are key components of hypothesis-testing (Koehler, 1991; Sanbonmatsu, Posavac, Kardes, & Mantel, 1998; Windschitle & Wells, 1998), and are shown to load heavily on working memory (Dougherty & Hunter, 2003a,b). Our studies provide some suggestion that generating hypotheses is equally demanding for individuals of all levels of working memory capacity. In Experiment 5, we told one group of participants that a mixed cue task would contain specifically one positive, one negative, and two irrelevant cues. Hence, participants were only required to evaluate the hypotheses against the feedback, rather than generate their own hypotheses. We found that all participants benefited equally, regardless of their WMC, but that this measure was still associated with performance. This provides some suggestion that both generating and evaluating hypotheses against feedback in learning environments is demanding on cognitive resources. However, it is important to note that even when participants are told how each cue is related to the criterion in multiple cue tasks, performance is still poorer for tasks that contain negative cues, compared with when the relevant cues are positive (Evans, Clibbens, & Harris, 2005). This indicates that a third factor, learning to apply explicit knowledge, may also load heavily on working memory in learning environments.

Evaluating hypotheses against feedback in multiple cue tasks involves assessing the correlation between cues and criterion. When cues are probabilistically related to a criterion, or feedback is noisy, this means comparing cue values with outcome values over a number of trials. However, WMC will likely put an upper limit on the number of items that can be considered at one time. Low span individuals tend to consider smaller samples of events when assessing the correlation between variables (Kareev, Lieberman, & Lev, 1997). In learning environments, this means that these individuals are perhaps less likely to detect a correlation in noisy environments when larger samples of events must be considered (R. B. Anderson, Doherty, Berg, & Friedrich, 2005). Furthermore, they may also be more likely to falsely believe that a correlation exists by considering small unrepresentative samples of events (Gaissmaier, Schooler, & Rieskamp, 2006; Juslin & Olsson, 2005).

6.1.4. Working memory capacity and the dissociation between performance and explicit knowledge

I discussed earlier that when performance is in excess of explicit knowledge of a task, this indicates that implicit learning has occurred. However, our studies of multiple cue judgement suggest that such dissociation should not be taken as evidence that explicit processing is not involved in learning at any stage. We found that performance on negative cue tasks exceeded explicit knowledge levels, suggesting initially that learning occurred implicitly. However, our measures of WMC indicated that these tasks also benefit from explicit processing effort. Hence, dissociation between measures of performance and explicit knowledge alone could be used to falsely claim that learning is entirely implicit.

The findings of our studies indicate that in some environments, learning appears to benefit from both an implicit *and* an explicit component. In negative cue tasks, learning may initially require conscious hypothesis-testing, but is also boosted by implicit learning processes above what is attainable on explicit knowledge alone. One possibility is that when both relevant cues are negative, participants learn quickly to reverse the feedback values they receive. After doing so on a number of trials this process may become automated, drawing less on working memory resources. By automating the process of making judgements that are incongruent with the cue values, implicit processing may begin to contribute to learning.

Viewed in this way, learning of negative cues is similar to learning in other types of cognitive tasks that require people to provide incongruent responses to stimuli. When instructed to respond to the location of a stimulus, people are shown to respond faster when stimulus and response sets are compatible, such as when a stimulus that appears to the left corresponds to a left key-press (De Jong, Liang, & Lauber, 1994; Eimer, 1995). This effect also occurs when visual locations are replaced with verbal labels (Proctor & Wang, 1997). In these tasks, stimulus-response compatibility appears to prime the automatic response that corresponds with the stimulus. However, when responses are required that are incongruent with a stimulus, similar to when a high value on a negative cue predicts a low criterion value, controlled effortful processing is initially required to inhibit and override a congruent response (Kornblum, Hasbroucq, & Osman, 1990). Similarly, a conscious effort is also required for negating stereotypic beliefs (Kawakami, Dion, & Dovidio, 1999), but may become automated through practice (Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008; Kawakami et al, 2000).

The idea that explicit knowledge becomes automated through practice is popular in the learning literature (e.g. Jones & Vanlehn, 1994). I discuss this issue in more detail later when I consider the effects of concurrent working memory load on performance (see section 6.1.6.). When performance initially requires the application of explicit rules (e.g. a high cue value predicts a low criterion value), this process can become automated as implicit procedures through practice (Logan, 1988; Ouellet, Beauchamp, Owen, & Doyon, 2004). In this way, explicit knowledge is required less and less through practice, and may even be forgotten (Neves & Anderson, 1981). Furthermore, implicit processes can begin to adjust and refine implicit procedures during training, such that learning continues to occur implicitly (Anderson, 1986, 1993). That is, learning may occur implicitly once explicit knowledge is automated. Our studies suggest that this process may occur early in training.

6.1.5. Transfer effects in learning

When performance on a task correlates with measures of WMC this indicates that explicit processing is involved. However, two explanations can account for a lack of association between performance and WMC. First, it may be that implicit learning explains performance. Second, the task may be easily performed explicitly by participants of all levels of cognitive capacity. In such cases, an association would not be observed because the explicit processing used to perform the task is not sufficiently demanding. Heuristic strategies that are explicit for instance, would take account of only limited amounts of information compared with more complex strategies, and are likely to load less on working memory resources. For instance, Newell and Shanks (2003) found that participants were able to report verbally their use of a single cue strategy when performing a two alternative forced choice task, suggesting that this strategy is applied explicitly. Moreover, Bröder (2003) found that individuals who apply this simple strategy tend to be of higher in intelligence but not WMC. Hence, people appear to apply simple single cue strategies explicitly without taxing working memory resources. Recall however, that other types of heuristic processing are shown to be implicit (e.g. Stanovich, 2004).

In our studies, we found no association between individual differences in WMC and performance (Experiments 4, 6, 7, and 8) or explicit knowledge (Experiments 7, and 8) in positive cue tasks. Furthermore, while these measures were ordinarily predicted by WMC in mixed cue tasks, they were not if participants first performed the positive cue task (Experiments 4 and 6). In this way, explicit processing that ordinarily contributed to mixed cue learning was apparently switched off. Hence, prior training on positive cues appears to induce an implicit mode of thinking that transfers to learning of mixed cue tasks.

An interesting implication of our findings is that high cognitive capacity does not guarantee better performance on tasks that benefit from explicit reasoning. If an individual does not recognise the need to make a conscious effort then automatic implicit processes appear to guide performance. Our transfer effects indicate that one way this can occur is if a prior task is performed implicitly. Consistent with this view, a number of studies have shown thinking dispositions (e.g. need for cognition) to predict performance in cognitive tasks even when individual differences in cognitive ability are controlled (Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Stanovich, 1999, 2008). What these studies imply is that performance can depend both on the individual's tendency to engage in effortful thinking, and their cognitive ability (Evans, 2007; West et al, 2008; Stanovich, 2008). Hence, higher WMC may not be sufficient for good performance; the individual must also be motivated to engage in explicit thinking. A wealth of research suggests that people are in fact prone to rely too heavily on intuition, which has led a number of theorists to describe people as 'cognitive misers' (Evans, 2006; Hull, 2001; Krueger & Funder, 2004; Tversky & Kahneman, 1974).

According to the cognitive miser hypothesis, inferences are often made intuitively unless the individual believes that effortful thinking is required of them (e.g. Stanovich & West, 2008). For this reason, task instructions that emphasis logical thinking tend to reduce the influence of intuition on reasoning by encouraging people to think analytically (Evans, 2003; Stevenson & Over, 1995; Vadeboncoeur & Markovits, 1999). A study by Tversky & Kahneman (1983) illustrates this point. After reading a personality description of an individual called 'Linda', in their studies the vast majority of people falsely assigned a higher probability to the conjunction of two events "Linda is a bank teller and is active in the feminist movement" than to a single event "Linda is a bank teller" (see section 1.2.5. for more details). When the two statements were rated one after the other, statisticians were far more likely to apply the correct conjunction rule than untrained participants. However, if the two statements were not rated in succession (placed within a longer list of items) or participants rated only one of the statements (between-subjects design), the statisticians showed no advantage. What this suggests is that even if an individual has the cognitive capacity and statistical knowledge to reason logically, they may not detect the need to apply a rule without sufficient cues (Kahneman, 2000).

A common finding in the problem solving literature is that with experience people tend to become fixed in their behaviour, and less sensitive to changes in the environment (Reder, 1987, 1988). In problem solving tasks, people show a strong tendency to continue to apply previously successful complex rules on transfer problems even when these can be solved using a simpler rule (Chen & Mo, 2004; Luchins & Luchins, 1959). In this way, previous success appears to induce an inflexible mode of thinking that is insensitive to changes in the task. This is the case even when a complex rule leads to errors on transfer problems that can only be solved by a simpler rule (Luchins, 1942). In this way, prior experience appears to have similar effects on both problem solving and multiple cue judgement. We found that participants became fixed in an implicit mode of thinking when a prior task could be performed implicitly.

In our studies explicit processing was switched off by prior training. It appears that participants did not detect that an explicit effort was required to perform a transfer task when a previous task could be learned implicitly. Similarly, in studies of problem solving people appear to be unaware of the errors they make on a task when applying previously successful rules (Woltz, Bell, Kyllonen, & Gardner, 1996). In this study, participants did not detect the errors they made on transfer problems, and did not recognise a need to engage in effortful thinking about the task. In our studies we attempted to undo the effects of prior experience on learning. In Experiment 6 we provided one group of participants explicit instruction, explaining that the mixed cue task would contain at least one negative cue. Despite this, learning continued to occur implicitly following positive cue training, indicating that prior experience can have a strong effect on the cognitive processes involved in learning.

Transfer effects may have important implications for learning in real-life situations. Acquisition of expertise likely involves a process in which additional cues become available at different stages of learning. For example, a student doctor may receive training on symptoms associated with one type of ailment, before learning a separate set of cues. If learning of the first task can be achieved implicitly, this may influence the way they learn a second set of cues.

6.1.6. The effects of working memory load

On the view that multiple cue judgement is explicit and demanding on cognitive resources, loading people's working memory should strongly disrupt their performance on multiple cue tasks. In Experiments 1 and 2, we introduced a concurrent load to the test trials of judgement tasks. It is important to note that we did not load working memory during training, as a great deal of research indicates that attention is required for both implicit and explicit learning of complex tasks (Cohen, Ivry, & Keele, 1990; Frensch, Buchner, & Lin, 1994). However, we expected that a concurrent load would selectively interfere with judgement that is mediated by explicit processing (DeCaro, Thomas, & Beilock, 2008; Stanovich & West, 1998a,b,c). Instead, we found no evidence that performance was affected in any of our task types by the addition of a visual load in Experiment 1, nor a verbal load in Experiment 2.

Recall that our measures of WMC indicate that learning in some types of multiple cue tasks is explicit and demanding on working memory (see section 6.1.3.). Hence, we should have expected some effect of load. Studies of reasoning for instance show that logical thinking is both correlated with individual differences in WMC, and disrupted by working memory loads designed to interfere with explicit reasoning (De Neys, 2006; Stanovich, 1999; 2004; Stannovich & West, 2000). Moreover, our load tasks were similar to those that have been applied to reasoning tasks, and should have strongly affected performance on tasks that were performed explicitly (De Neys, 2006). The lack of effect of load on performance in our studies indicates that judgement is not demanding on working memory when performance is assessed after training.

Our findings are not surprising, however, when we consider other types of complex skill learning. It is commonly shown that even when learning is explicit, conscious knowledge used to perform a task can become automated through practice (Kramer, Strayer, & Buckley, 1990; Logan, 1988, 1992; Logan & Klapp, 1991; Ouellet et al, 2004). When an explicit skill is acquired performance is characterised by slow effortful application of knowledge (Anderson et al. 2004). With practice, this process can become automated as implicit procedures, drawing less on working memory dependent explicit processing. Accordingly, performance is shown to be unaffected by the addition of concurrent load designed to disrupt explicit processing (Brown & Carr, 1989; Logan, 1979). However, for this process to occur the individual must have extensive practice with a task. In the multiple cue tasks used in the present thesis participants completed 80 learning trials before a further 40 test trials without outcome feedback. It is highly likely that these conditions provide sufficient practice for explicit knowledge to become proceduralised. This explains why performance was not affected by the addition of concurrent load to the test phases of our tasks, even for those that are initially learned explicitly.

Most theories of automaticity describe the proceduralisation of explicit knowledge as involving a process of knowledge compilation (Anderson, 1982, 1993; Neves & Anderson, 1981). During this phase, explicit rules are converted directly into procedural knowledge. Another way that explicit knowledge may become proceduralised is by observing one's performance on a task. In our experiments, participants may have trained implicit learning processes by observing their own judgements and attending more to the cues they believed were important. By doing so, participants may have acquired implicit knowledge indirectly from their explicit beliefs about the cues, which then begin to guide judgements (Willingham & Goedert-Eschmann, 1999).

Much expert judgement may involve a similar process of automation of skills. A stockbroker for instance, will make numerous predictions about share price during a financial year. Learning how market indicators are related to share price may initially involve a process of explicit hypothesis-testing, during which, judgements would be guided by the stockbroker's explicit beliefs. In an environment where fast judgements must be made from multiple pieces of information, the slow and cognitively demanding process of explicit judgement likely becomes proceduralised through practice. In the process, fast intuitive judgements generated by implicit processes would begin to compete with explicit reasoning to control behaviour (Neves & Anderson, 1981; Logan, 1988, 1992).

Theories of automation imply that judgement can become more intuitive with experience, and draw less on explicit beliefs. Furthermore, explicit knowledge acquired during learning may even be forgotten (Anderson, 1993). This issue has important implications for studies of expert judgement. Recall that experienced experts often demonstrate only limited insight into how they make their judgements (e.g. Wigton, 1996). For these individuals, judgement may be intuitive. Accordingly, we would expect more experienced experts to have less insight into their judgement policies. Consistent with this view, Slovic, Fleissner, & Bauman (1972) found that the more experience a stock broker had making financial decisions, the less they were able verbally to report how they made their judgements.

6.2. MULTIPLE CUE JUDGEMENT AND DUAL PROCESS THEORIES OF THINKING

Dual process theories are popular in cognitive psychology (Ashby & Maddox, 2005; Evans, 2008; Evans & Over, 1996; Kahneman & Frederick, 2002; Sloman, 1996; Sun, Slusarz, & Terry, 2005). That both implicit and explicit processing can contribute to performance on a task is widely recognised in cognitive research (Berry & Broadbent, 1988; Cleeremans, Destrebecqz, & Boyer, 1998; Pothos, 2007; Reber, 1993; Smith & Grossman, 2008). Dual process theories describe how implicit and explicit processes interact to control behaviour, and make predictions about the contribution of each type of process to performance on cognitive tasks. In our studies of multiple cue judgement we found that both implicit and explicit processes are involved in learning. I concluded earlier that explicit processing is ordinarily involved in learning multiple cue tasks that contain negative cues (see section 6.1.3.), but that explicit knowledge becomes automated through practice (see section 6.1.6.). Positive cue learning on the other hand, appears to occur implicitly, and induces an implicit mode of thinking. In this section I relate our findings to dual process models of other cognitive tasks, and discuss multiple cue judgement in more detail within a dual process framework.

6.2.1. Dual process theories of reasoning and judgement

Dual process theories in the reasoning and judgement literatures often make a distinction between heuristic and analytic processes, which are both believed to influence performance on cognitive tasks (De Neys, 2006; Evans, 1989, 2006, 2008; Evans & Over, 1996;

Kahneman & Frederick, 2005, 2005; Stanovich, 2004). These processes generally map on to implicit (heuristic) and explicit (analytic) modes of thought. I discussed in Chapter 1 (see section 1.2.) how heuristic processes direct attention to relevant information based on prior beliefs or knowledge, such as when reasoning about logical statements or making judgements of probability (Evans, 1984, 1996; Kahneman, 2003). However, heuristic responses can lead to errors on cognitive tasks. This is because heuristic processes can both direct attention towards information that is irrelevant and direct attention away from relevant information (Evans, 2006). Hence, reasoning can be biased by heuristic processing when relevant information is neglected and irrelevant information is included (Kahneman & Tversky, 1996). In such cases, effortful analytic processing may be required to inhibit and override an intuitive response for logical reasoning. However, analytic processing is slow, effortful, and demanding on working memory.

Most heuristic-analytic theories describe a sequential process of reasoning and judgement (Evans, 2006; 2008; Kahneman, 2000). Automatic heuristic processes generate fast intuitive responses to information provided in cognitive tasks. An analytic mode of thinking may then inhibit and override an intuitive response, but this process is slow and effortful. Hence, heuristic processing is the default in sequential models. Our findings are generally consistent with this view. Learning in multiple cue tasks appears to occur implicitly, unless an individual believes an explicit effort is required of them, such as when a task contains negative cues. I discussed earlier how our findings of transfer effects in learning provide further support for this class of models (see section 6.1.5.). Whether an individual engages in analytic reasoning depends in part, on whether they detect the need for conscious effort (West et al, 2008), and whether they are sufficiently cued to apply rule based knowledge (Tversky & Kahneman, 1983). Similarly, our transfer effects indicate that implicit processing is not overridden in tasks that ordinarily benefit from explicit reasoning if an implicit mode was previous successful.

Dual process models that describe a competition between intuitive and effortful thinking provide a good account of expert judgement. I discussed earlier the possibility that experienced experts make fast intuitive judgements based on proceduralised knowledge (see section 6.1.6.). Judgements are likely to be made intuitively, unless the individual engages in slow effortful weighing up of available information. As with expert judgement, sequential models are specifically designed to account for reasoning and judgement when prior beliefs and knowledge are available. In our studies, we instead trained participants to make judgements from multiple pieces of information. While intuitive responses based on prior beliefs and knowledge may be generated quickly by implicit processes, implicit learning is likely to be slow and gradual. Following experiential learning, implicit processes may over take effortful explicit thinking by generating fast intuitive responses. However, implicit knowledge must first be acquired through a process of incremental learning.

Another class of dual process models describes a competition between parallel implicit and explicit processing. (e.g. Sloman, 1996, 2002; Smith & DeCoster, 2000). Parallel-competitive models propose that conscious and unconscious modes of thinking proceed independently, and compete to control behaviour. Cognitive processes of the implicit system are described as associative and are responsible for learning statistical regularities in the environment such as correlations (Smith & DeCoster, 2000). These are contrasted with explicit rule based processing that is involved in conscious hypothesis-testing and formalising verbal rules. Both these processes are believed to be active when performing a task, and can generate conflicting responses. Describing implicit processing as associative and explicit processing as rule-based fits well with our studies of multiple cue judgement. Implicit learning in multiple cue tasks likely involves a process of trial-by-trial learning, in which cue-criterion associations are learned gradually in response to feedback. Explicit processing on the other hand appears to involve a process of deliberative hypothesis-testing. In this way, the individual consciously tests hypotheses against the feedback they receive in a controlled explicit manner.

While parallel-competitive models do not specify the exact nature of implicit associative processes, it is possible to speculate how this kind of learning may occur in MCPL. In recent years, learning theorists have begun to formalise associative learning mechanisms in terms of neural network (or connectionist) models (Cleeremans & McClelland, 1991; Dienes, 1992; Gibson, Fichman, & Plaut, 1997). In a seminal paper by Gluck and Bower (1988; and extended by Cobos, Almaraz, & García-Madruga, 2003), the authors describe a simple one-layer neural network capable of learning to predict a binary outcome from binary cues. The neural network adjusted the weight given to each cue in response to outcome feedback and learned successfully to predict an outcome in a learning phase. Moreover, the pattern of responses generated by the model was remarkably similar to that of people trained on the same task.

A compelling aspect of this type of model is its simplicity. Each cue is represented in the network by an input unit which are connected directly to a single output unit corresponding to the predicted outcome (or judgement). During training, the weighted node connections are then adjusted in response to feedback by means of an error correcting rule (e.g. least mean squares rule, also known as delta rule). Over a number of trials, error in judgement (the difference between judgment and actual outcome) is reduced as the weighted node connections are adjusted. It is easy to imagine how such a simple network could form part of the implicit learning system in humans.

An assumption made by some theorists is that two contradictory responses can be held in mind whilst performing a task (Sloman, 2002), one generated by implicit and another by explicit processes. Sloman (1996) illustrates how this can occur by reference to the Müller-Lyer illusion displayed in Figure 6.1. When asked whether the two lines are of equal length, knowledge that the two lines are indeed similar does not influence our perception that the above line is longer. Hence, two contradictory beliefs generated by separate cognitive systems can be held in mind. However, this idea is perhaps less consistent with our findings of transfer effects in learning. Recall that when a task could be performed implicitly, this mode of thinking transferred to learning of a second task. In this way, an explicit component in learning was reduced, indicating that the two modes of processing do not occur independently.

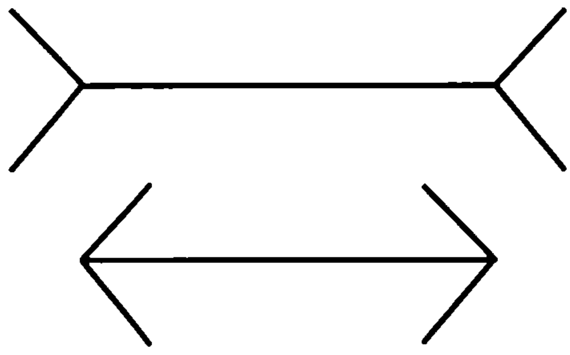

Figure 6.1. Müller-Lyer illusion

Our studies of multiple cue judgement appear most consistent with dual process theories of reasoning and judgement that describe a sequential process. Accordingly, learning is implicit unless individuals engage in effortful analytic thinking. Models that propose a competition between parallel systems, on the hand, do not appear to account for multiple cue judgement. This type of model proposes that cognitive systems proceed in parallel. However, our findings indicate that explicit processing is reduced when a task can be performed implicitly, indicating that these modes of thinking are not independent.

6.2.2. Dual process theories of learning

It is commonly assumed that learning of complex cognitive tasks occurs either implicitly or explicitly. When tasks are highly complex performance is usually dominated by implicit learning processes (Cohen et al, 1990; Hayes & Broadbent, 1988; Nissen & Bullemer, 1987; Reber, 1967; Sun et al, 2001). However, when the underlying rules of a task are simple and easily verbalised (Ashby & Maddox, 2005), or when few variables must be considered (Broadbent, FitzGerald, Broadbent, 1986; Geddes & Stevenson, 1997), learning is more likely to be explicit. The contribution of an explicit component to learning, therefore, appears to depend on whether the information provided in a task can be managed in working memory, and whether the structure of a task can be formalised as simple rules (Baddeley, 1986; Barrett, Tugade, & Engle, 2004).

I discussed in the previous section how our studies of multiple cue judgement can be understood in terms of a competition between multiple cognitive systems. Dual process models have a long history in the learning literature (e.g. Brooks, 1978; Reber, 1967), and theorists have recently begun to propose how cognitive processes may compete during learning (Ashby & Maddox, 2005; Price, 2005; Smith & Grossman, 2008; Sun et al, 1998, 2001). Ashby, Alfonso-Reese, Turken, and Waldron's (1998) COVIS (competition between verbal and implicit systems) model describes this kind of competition (Ashby & Maddox, 2005; Maddox, Filoteo, Hejl, & Ing, 2004; Waldron & Ashby, 2001). They explain that implicit and explicit processes initially compete when a new task is learned, both generating separate and often conflicting responses. The mode of thinking that generates more accurate responses then begins to control the inferences an individual makes and dominate learning.

Consistent with competitive models of learning, when accurate explicit knowledge is acquired an explicit mode of processing is shown to control behaviour. For example, when a single input must be manipulated to control the output of a system in dynamic control tasks, participants acquire accurate conscious knowledge of how to control the system (Dienes & Fahey, 1995; Geddes & Stevenson, 1997). Confirming that an explicit mode dominates learning in these tasks, performance is strongly disrupted by the addition of concurrent working memory load during training (Hayes & Broadbent, 1988). Similarly, when learning to categorise stimuli based on multiple cues, performance is more affected by concurrent load tasks when category rules are simple and easily verbalised (Waldron & Ashby, 2001). Interestingly, when category rules are harder to verbalise performance is instead disrupted by delaying the onset of feedback (Maddox, Bohil, & Ing, 2003; Maddox et al, 2004), indicating that performance in these tasks is dominated by some implicit reward system (Maddox & Ashby, 2004).

Neuropsychological studies of learning suggest that brain areas associated with implicit and explicit learning systems are activated exclusively during learning, rather than in parallel (Hazeltine, Grafton, & Ivry, 1997; Ivan, Krams, Turner, & Passingham, 1998; Pascual-Leone, Grafman, & Hallett, 1994). Consistent with competitive models, these studies indicate that learning occurs via an implicit or an explicit route, in which brain activation is associated with one or the other system. Studies of sequence learning have shown that when learning occurs in absence of awareness, decreased activation is observed in brain areas associated with explicit processing (namely, within the temporal cortex; Grafton, Hazeltine, & Ivry, 1995; Poldrack et al, 1997). Hence, when one learning system dominates performance, activation of the other is reduced. Our findings provide additional support for models that describe this kind of competition between cognitive systems. When multiple cue learning was implicit, explicit reasoning was reduced when a second task was learned.

The findings of the present thesis, therefore, appear most consistent with dual process models that describe a competition between implicit and explicit modes of processing early in training. When learning occurs via one route, activation of the other is reduced. We found that learning in negative cue tasks appeared to benefit from both an implicit and an explicit component. However, this finding does not cause problems for our conclusions. When brain activity is studied during sequence learning, increased activation is observed in areas associated with explicit reasoning for individuals who are explicitly aware of the repeating sequence (e.g. Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994). However, activation in these areas soon reduces when performance becomes automated (Ivan et al, 1998). It is likely that implicit processing begins to contribute to learning at this stage in negative cue tasks.

6.2.3. Dual process theories: Quality and quantity

I discussed earlier how individuals of high cognitive ability are more likely to engage in explicit reasoning (see section 6.1.3.). When deciding whether a conclusion necessarily follows from a set of propositions, people are heavily influenced by the believability of the conclusion, an effect referred to as *belief bias* (Evans et al, 1983). However, some individuals are shown to inhibit and override prior belief and engage in explicit reasoning about these tasks. These individuals tend to be of higher cognitive ability (Stanovich, 2004). Hence, individual differences in ability (or working memory capacity) appear to predict the *quantity* of explicit reasoning on the part of the individual, such that more able participants are expected to engage in more explicit reasoning (Evans, 2007; Stanovich, 1999). This conclusion has been made by a number of theorists in the reasoning and judgement field, namely Stanovich and his colleagues (Kokis et al, 2002; Stanovich, 2009).

However, recall that independent of cognitive ability thinking dispositions (e.g. open-mindedness) also predict performance in reasoning tasks (Kokis et al, 2002). An alternative hypothesis is that of those individuals who engage in explicit reasoning, the *quality* of their reasoning is predicted by cognitive ability (Evans, 2007b). That is, those of higher ability are more likely to reason correctly given that they engage in analytic thinking. Evans (2007b) argues that a number of findings are consistent with a quality hypothesis. For example, participants of all levels of ability are shown to be equally influenced by the believability of conclusions in syllogistic tasks, indicating that individuals of higher ability are simply more effective when they reason analytically. This hypothesis is also more consistent with our findings of multiple cue judgement. Experiment 5 addresses this issue. When individual differences predicted performance in multiple cue tasks, all participants benefited equally from task information designed to improve explicit learning. This was demonstrated by a main effect of task information on performance. We concluded that task information improved explicit learning by directing participants' attention towards relevant cues. A quantity hypothesis would predict that only high span individuals would benefit from task information. This is because low span individuals would be less likely to engage in explicit reasoning. In line with the predictions of a quality hypothesis, we found that participants of all levels of WMC equally benefitted from task instructions, indicating that all participants were explicitly engaged in learning. The association between WMC and performance in our tasks demonstrates that individuals of higher ability are more effective in explicit learning.

6.3. CONCLUSIONS

In the present thesis we explored the cognitive processes of multiple cue judgement. It is commonly argued that this kind of judgement is explicit. These theories fall within single process models of thinking that describe a single explicit learning system. Consistent with this view, we found that performance in some tasks does benefit from an explicit component during learning. However, other tasks were learned well implicitly. Furthermore, when learning is explicit, judgement appears to become more intuitive (automated) through practice. Our findings imply that multiple cue judgement is better understood within a dual process framework.

A robust finding in research on multiple cue judgement is that performance is poorer on tasks that contain negative cues. This trend makes sense when we think of negative cue learning as effortful and demanding on working memory. Whereas positive cue tasks are easily learned implicitly, negative cue learning loads heavily on working memory dependent explicit processing.

Our findings appear most consistent with dual process models of thinking that describe a competition between implicit and explicit modes of processing early in training. According to these models, cognitive processes initially compete to control behaviour. During learning one system begins to dominate learning, which reduces activation of the other. Our findings provide further support for this class of dual process models.

A puzzling finding in the judgement literature is that experts are often unable verbally to report how they make their judgements. One possibility is that some expertise is acquired unconsciously and is not available for verbal report. Our findings suggest that while both implicit and explicit processes likely contribute to the acquisition of expertise, judgement becomes more intuitive through practice. In the process, explicit knowledge becomes automated as implicit procedures, during which, explicit beliefs contribute less and less to judgement. This explains, in part, why self-insight is often poor.

Appendix

7.1. Additional Analyses: Performance in Learning Phases

Experiment 1

Performance scores in the learning phase were calculated by correlating judgments with the criterion for each participant separately for the 1st 40 learning trials and 2nd 40 learning trials. These are displayed in Table 7.1. All group performance scores were significantly above zero. Significant group performance scores in the 1st 40 learning trials indicate that learning had occurred even in the early stages of the task.

Table 7.1

Experiment 1: Mean group performance scores in learning phases

	1st 40 learning trials		2nd 40 learning trials	
	M	t	M	t
++00 task	.50*	17.329	.56*	21.916
+-00 task	.11*	3.370	.19*	3.806

* $p < .05$

To explore how performance progressed during training a two-way mixed ANOVA was conducted on performance scores in the learning trials, using task-type (++00 or +-00) as an independent factor, block (1st 40 and 2nd 40 learning trials) as a within-subjects factor, and performance correlations as the dependent variable. We collapsed across load and no load groups for this analysis, as the load task was not administered until the test phase.

There was a significant effect of block ($F_{(1,78)}$ = 8.296, MSE = .481, p = .005, partial η^2 = .096), indicating that performance levels improved from the 1st 40 (.30) to the 2nd 40 (.38) learning trials. There was also a significant effect of task-type ($F_{(1,78)}$ = 74.440, MSE = 8.283, p <.001, partial η^2 = .488), with participants performing better in the ++00 task (.53) than the +-00 task (.15). There were no significant interactions. These findings confirm that performance improved during training, and was generally better in ++00 than +-00 tasks even during training.

Experiment 2

Mean group performance scores in the 1st 40 and 2nd 40 learning trials are displayed in Table 7.2. All group performance scores were significant and above zero, except for those performing the --00 task in 1st 40 learning trials.

Table 7.2

Experiment 2: Mean group performance scores in learning phases

	1st 40 learning trials		2nd 40 learning trials	
	M	t	M	t
++00 task	.47*	16.578	.51*	21.724
+-00 task	.08*	2.598	.11*	2.579
--00 task	.05	1.214	.31*	6.833

* p <.05

A two-way mixed ANOVA was conducted on performance scores in the learning phase, using task-type (++00, +-00, or --00) as an independent factor, and block (1st 40 and

2nd 40 learning trials) as a within-subjects factor. As in Experiment 1, we collapsed across load and no load groups. There was a significant effect of task-type ($F_{(2,117)} = 45.622$, $MSE = 4.732$, $p < .001$, partial $\eta^2 = .438$). Two-way mixed ANOVAs including block as a within-subjects factor confirmed that this was because participants trained on the ++00 task (.49) performed significantly better than those performing the --00 task (.18; $F_{(1,78)} = 44.773$, $MSE = 5.166$, $p < .001$, partial $\eta^2 = .365$), and that participants performed marginally significantly better in the --00 task than the +-00 task (.09; $F_{(1,78)} = 3.720$, $MSE = .434$, $p = .057$, partial $\eta^2 = .046$). There was a significant effect of block ($F_{(1,117)} = 24.234$, $MSE = 1.030$, $p < .001$, partial $\eta^2 = .172$), and a significant interaction between task-type and block ($F_{(2,117)} = 8.895$, $MSE = .378$, $p < .001$, partial $\eta^2 = .132$). Pairwise comparisons confirmed that while performance levels improved significantly from the 1st 40 (.05) to the 2nd 40 (.31) learning trials in the --00 task ($t = 5.561$, $df = 39$, $p < .001$), performance levels did not significantly improve in the ++00 ($t = 1.605$, $df = 39$, $p = .117$) or the +-00 ($t = .886$, $df = 39$, $p = .381$) tasks. This indicates that while the majority of learning in the +-00 and ++00 tasks occurred early in the training phase, a substantial amount of learning continued between the 1st 40 and 2nd 40 learning trials in the --00 task.

Experiment 3

Mean group performance scores in the 1st 40 and 2nd 40 learning trials are displayed in Table 7.3. All groups achieved significant performance scores in the learning phase, except participants performing the +-00 and --00 tasks in the 1st 40 learning trials under 'low' instructions. While our instructional manipulation had no effect on performance in the test phase (see section 3.3.2.), directing participants towards the presence of negative cues may have given them an initial advantage in early stages of training. A three-way mixed ANOVA was conducted on performance scores in the learning trials, using task-type (++00, +-00, or --00) and instruction type ('high' or 'low') as independent factors, and block (1^{st} 40 and 2^{nd} 40 learning trials) as a within-subjects factor. There was a significant effect of task-type ($F_{(2,114)} = 44.666$, $MSE = 5.520$, $p < .001$, partial $\eta^2 = .439$). Three-way mixed ANOVAs, which included instruction type as an independent factor and block as a within-subjects factor, confirmed that performance was significantly better in the ++00 task (.55) than the --00 task (.23; $F_{(1,76)} = 50.265$, $MSE = 6.724$, $p < .001$, partial $\eta^2 = .398$). Although participants appeared to perform better in the --00 task than the +-00 task (.16), this effect did not reach significance ($F_{(1,76)} = 2.168$, $MSE = .253$, $p = .145$, partial $\eta^2 = .028$). There was also a significant effect of block ($F_{(1,114)} = 11.053$, $MSE = .562$, $p = .001$, partial $\eta^2 = .088$), with performance scores generally improving from the 1^{st} 40 (.33) to the 2^{nd} 40 (.35) learning trials. There was some suggestion of an effect of instructions, with participants performing generally better in the learning phase if provided 'high' instructions (.35) than 'low' instructions (.28). However, this effect did not reach significance ($F_{(1,114)} = 2.021$, $MSE = .250$, $p = .158$, partial $\eta^2 = .017$). In contrast with our analyses of performance scores in learning phases of Experiment 2, there was no significant interaction between task-type and block ($F_{(2,114)} = 1.580$, $MSE = .080$, $p = .210$, partial $\eta^2 = .027$), indicating that performance levels generally improved from the 1^{st} 40 to the 2^{nd} 40 learning trials regardless of the type of task participants performed. There were no other significant interactions.

Table 7.3.

Experiment 3: Mean group performance scores in learning phases

	1st 40 learning trials		2nd 40 learning trials	
	M	t	M	t
++00 task				
'High' instructions	.53*	22.259	.56*	16.336
'Low' instructions	.52*	8.350	.59*	13.239
+-00 task				
'High' instructions	.21*	4.779	.23*	3.570
'Low' instructions	.08	1.528	.13*	2.148
--00 task				
'High' instructions	.22*	3.616	.33*	6.090
'Low' instructions	.09	1.315	.27*	4.772

* $p < .05$

Experiment 4

Task 1 performance in the learning phase

Mean group performance scores and significance are displayed in Table 7.4. All group performance scores were significant, except for those performing the --00 task in the 1st 40 learning trials.

Table 7.4.

Experiment 4: Mean group performance scores in learning phases

	1st 40 learning trials		2nd 40 learning trial	
	M	t	M	t
++00 task	.52*	15.052	.54*	15.851
--00 task	.08	1.584	.31*	6.663
+-00 task (++00 transfer)	.16*	3.794	.23*	5.168
+-00 task (--00 transfer)	.17*	4.428	.22*	5.209

* $p < .05$

In our analysis of performance scores in learning phases we explored the moderating effects of learning block, as a within-subjects factor, on an association between working memory capacity (WMC) and performance. For this reason we followed the multilevel procedure to analyses our data (see section 7.2. for more details). We entered task-type (++00 or --00), block (1^{st} 40 and 2^{nd} 40 learning trials), and WMC as predictors, and a random intercept (χ^2 difference (1) = 30, p <.01). Consistent with analyses of performance in test phases, there was a significant effect of task-type (β = -.400, SE = .066, $t(80)$ = -6.097, p <.001), indicating that performance levels were better in the ++00 task (.53) than the --00 task (.19). There was also a significant effect of block (β = .161, SE = .035, $t(80)$ = 4.599, p <.001), with performance scores improving from the 1^{st} 40 (.30) to the 2^{nd} 40 (.42) learning trials. There was some suggestion that WMC was positively associated with performance, but this effect did not reach significance (β = .005, SE = .003, $t(80)$ = 1.741, p = .085). There was however, a significant interaction between task-type and WMC (β = .012, SE = .006, $t(80)$ = 2.168, p = .033). Simple slope analyses confirmed that higher WMC was significantly associated with better performance in the --00 task (β = .011, SE = .004, $t(40)$ = 2.781, p = .008), but not the ++00 task (β = -.001, SE = .004, $t(40)$ = -.279, p = .782). There was also a significant interaction between task-type and block (β = .238, SE = .065, $t(80)$ = 3.680, p <.001). Simple slope analysis showed that performance scores improved significantly in the --00 task from the 1^{st} 40 (.08) to the 2^{nd} 40 (.31) learning trials (β = .279, SE = .053, $t(40)$ = 5.267, p <.001), but not in the ++00 task (β = .042, SE = .037, $t(40)$ = 1.133, p = .264), suggesting that the majority of learning in these tasks occurs early in training. There were no other significant interactions.

Mixed-cue transfer task performance in the learning phase

Both groups achieved significant performance scores in the 1st 40 and 2nd 40 learning trials (see Table 7.4.), indicating that a significant degree of learning had occurred in the early stages of the task. We followed the same multilevel procedure used to analyse performance in the learning phase of the first task in our analysis of learning in the mixed cue transfer task (see section 7.2. for more details). For this, we entered previous task-type (++00 or --00), block (1st 40 and 2nd 40 learning trials), and WMC as predictors, and a random intercept (χ^2 difference (1) = 36, p <.01). This analysis yielded only one significant main effect, which was for block (β = .075, SE = .031, $t(80)$ = 2.421, p = .018), indicating that performance scores improved from the 1st 40 (.16) to the 2nd 40 (.23) learning trials. There were no significant effects of previous task-type (β = .002, SE = .062, $t(80)$ = .031, p = .975), or WMC (β = .004, SE = .003, $t(80)$ = 1.306, p = .195). However, confirming our analyses of performance scores in test phases, there was a significant interaction between the type of task participants had previously performed and WMC (β = .015, SE = .005, $t(80)$ = 2.930, p = .004). Simple slope analysis showed that higher WMC was associated with better performance for those who had previously completed the --00 task (β = .011, SE = .003, $t(40)$ = 3.572, p = .001), but not for those who had previously performed the ++00 task (β = -.004, SE = .004, $t(40)$ = -.984, p = .331).

Experiment 5

Mixed cue task performance in the learning phase

Performance scores in the 1st 40 and 2nd 40 learning trials are displayed in Table 7.5. These were all significant and above zero for both groups.

Table 7.5.

Experiment 5: Mean group performance scores in learning phases

	1st 40 learning trial		2nd 40 learning trial	
	M	t	M	t
+-00 task	.21*	5.592	.30*	6.490
+-00 task (with task information)	.22*	5.051	.39*	6.653

* p <.05

Following the multilevel analyses procedure (see section 7.2.) we entered task information (no task information or with task information), WMC, and block (1st 40 and 2nd 40 learning trials) as predictors, and a random intercept (χ^2 difference (1) = 17, p <.01). Consistent with our analyses of performance scores in test phases (see section 4.2.2.), there was a significant effect of WMC (β = .009, SE = .003, $t(72)$ = 3.078, p = .003), with higher WMC associated with better performance. There was also a significant effect of block (β = .206, SE = .043, $t(72)$ = 4.760, p <.001), indicating that performance scores improved from the 1st 40 (.21) to the 2nd 40 (.35) learning trials. There was some suggestion that task information improved performance in the learning trials, indicated by better performance for participants provided task information (.31) than for those who were not provided task information (.26). However, this effect did not reach significance (β = .100, SE = .070, $t(72)$

= 1.421, p = .160).There was also a significant interaction between WMC and block (β = .007, SE = .003, $t(72)$ = 2.072, p = .042). Simple slope analysis confirmed that WMC was more strongly associated with performance in the 2nd 40 (β = .012, SE = .004, $t(70)$ = 2.988, p = .004) than the 1st 40 (β = .005, SE = .003, $t(70)$ = 2.046, p = .045) learning trials.

Experiment 6

Task 1 performance in the learning phase

Mean group performance scores during learning phases are displayed in Table 7.6. These were both significant and above zero.

Table 7.6

Experiment6: Mean group performance scores in learning phases

	1st 40 learning trials		2nd 40 learning trials	
	M	t	M	t
++00 task	.41*	13.890	.54*	21.609
+-00 task (++00 transfer)	.15*	3.848	.28*	5.515
+-00 task (++00 transfer with explicit instruction)	.17*	3.502	.21*	4.075

*p <.05

Since the instructional manipulation was not introduced until the mixed cue transfer task (see section 4.3.), we collapsed across all 72 participants' performance scores in the learning phase for this analysis. Following the multilevel procedure, we entered block (1st 40 and 2nd 40 learning trials) and WMC as predictors, and a random intercept (χ^2 difference (1) = 9, p <.01). This yielded a significant effect of block (β = .185, SE = .042,

$t(72) = 4.347$, $p < .001$), indicating that performance levels improved from the 1st 40 (.41) to the 2nd 40 (.54) learning trials. Confirming our analyses of performance scores in test phases (see section 4.3.2), there was no association between WMC and performance (β = .003, SE = .002, $t(72) = 1.339$, $p = .185$). There were no significant interactions.

Mixed cue transfer task performance in the learning phase

Mean group performance scores are displayed in Table 7.6. These were all above zero and significant. For this analysis, we entered explicit instruction (no explicit instruction or with explicit instruction), block (1st 40 and 2nd 40 learning trials), and WMC as predictors, and a random intercept (χ^2 difference (1) = 43, $p < .01$), along with a random slope for block (χ^2 difference (1) = 8, $p = .01$), in to a multilevel model. There was a significant effect of block (β = .114, SE = .034, $t(72) = 3.364$, $p = .001$), confirming that performance scores improved from the 1st 40 (.16) to the 2nd 40 (.25) learning trials. In our analyses of performance scores in the test phases we found that explicit instruction had no effect on performance (see section 4.3.2.). Confirming these findings, there was no effect of explicit instruction (β = .015, SE = .073, $t(72) = .205$, $p = .839$), or WMC (β = .001, SE = .003, $t(72) = .216$, $p = .830$) on performance in learning phases. There were no significant interactions.

7.2. Details of Analyses

Moderated Regression

A moderated regression model is a multiple regression with one or more interaction terms. Using moderated regression in the present thesis allowed us to explore the moderating effects of one or more independent predictors on an association between a continuous predictor and a dependent variable. We prepared our data in SPSS by coding categorical variables with zeros and ones. We then followed the two-step process advised by Aiken and West (1991) to assess the main effects of our predictors and whether there were any significant interactions. In the first block, we entered our predictors into a multiple regression model to assess the effects of each predictor on the dependent variable. All continuous variables were mean centred to reduce colinearity. This analysis provides regression coefficients and significance tests for each predictor. In a second block, we then entered all possible interactions. We specified the interaction terms by multiplying each predictor by each other predictor. For this, we used the 'compute' function provided in SPSS. In this final stage, we reported the R^2 value of the regression model, provided in the SPSS analyses output. R^2 provides a measure of the goodness-of-fit of a regression model.

Multilevel Analyses

Moderated regression assumes that all observations are independent, such as when participants either perform a positive cue (++00) or mixed cue (+-00) task. In the present thesis we also wanted to assess the moderating effects of within-subjects factors, such as when participants performed a judgement task under two conditions. Unlike moderated regression, multilevel regression models allowed us to explore the effects of within-subjects factors. This was achieved by modelling our data on two levels. The macro (upper) level units were participants and the micro (lower) level units were observations within participants. Hence, between-subjects effects and WMC act to differentiate macro level units and within-subjects factors differentiate micro level units. We followed the procedure advised by Hoffman and Rovine (2007) for our multilevel analyses. Their procedure is specifically designed for analyses using experimental designs. We performed our analyses using the SPSS Mixed Procedure. As well as assessing main effects and interactions, we also wanted to improve our model by entering additional parameters. For this reason we opted for the *maximum likelihood* statistic. This allowed us to make comparisons between successive models to assess whether entering additional parameters improved the goodness-of-fit of the model.

In our first step, we entered our main effects as predictors into a multilevel model. At this stage we also noted the -2 log-likelihood value (χ^2), which provides a measure of the goodness-of-fit of the model. A number of adjusted measures of this statistic are provided in the SPSS analysis output. We opted for the *Akaike's information criterion* (AIC), which is most commonly used. We then assessed whether a number of additional parameters improved the model. This was achieved by adding each parameter one-by-one and assessing whether each had a significant effect on the model. For this, we computed the chi-square statistic on the difference in -2 log-likelihood value after adding a parameter. These analyses had one degree of freedom because only one parameter was added at a time. If a parameter did not significantly improve the goodness-of-fit of the model it was removed. We first tested whether adding a random intercept significantly improved the model. Adding random intercepts allows the intercepts of the regression lines to vary between individuals. We then introduced random slopes for each predictor to assess whether each improved the model. This involved adding a random slope for each predictor one at a time and testing on each occasion whether the -2 log-likelihood was significantly different. Adding random slopes allows the regression slopes to vary across people. We then assessed the coefficients and significance tests of each predictor.

In a final stage, we added all possible interaction terms to the model. In some cases, this meant entering a large number of terms. For this reason, in line with the advice of Hoffman and Rovine (2007), we removed insignificant interactions one-by-one in successive models, starting with those of the highest order. We then assessed the coefficients and significance of each remaining interaction.

7.3. Task Instructions

Experiments 1, 2, 4, 5, and 6.

Four personality tests are being investigated by a business for their ability to predict the suitability of job applicants. However, these tests could predict suitability both positively and negatively. For example, extraversion could predict good applicant suitability for some occupations, but introversion could be a good predictor of suitability for others. Others tests could be irrelevant and fail to predict applicant suitability either way.

You will be required to examine how good each test is by rating the suitability of each applicant based on their results at the time of applying for the job. You will then be told how suitable each applicant turned out to be six months into the job. Use this information to try to improve your prediction of the job success of later applicants. You may find this task difficult at first, but as the trials go on you may find your judgments improve.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

When you begin the experiment you will see the four tests labelled from A to D at the top of the screen. Along side each test you will see the results of the applicant rated as either; **Very Low, Low, Average, High**, or **Very High**.

Below the tests you will need to select the rating that you think best describes the suitability of the applicant by clicking on the rating using the mouse.

Once you have made your selection the actual suitability of the applicant will appear below. By then pressing the spacebar you will be presented the next applicant.

Remember, each test can predict applicant suitability either **positive, negatively**, or be **irrelevant**:

Positive: If a **high** result on a test predicts **good** applicant suitability, and a **low** result predicts **poor** suitability. For example, high IQ may predict good applicants and low IQ predict poor applicants.

Irrelevant: If a test is **not predictive** of applicant suitability. For example, the height of the applicant may not predict applicant suitability.

Negative: If a **low** result on a test predicts **good** applicant suitability, and a **high** result predicts **poor** suitability. For example, low anxiety may predict good applicants and high anxiety predict poor applicants.

Experiment 3.

Instructions for predicting "high" house prices

You have recently started working for a luxury estate agent that sells only expensive houses so you are interested in finding houses with high house prices. You would like to know which features of houses, such as their size and age, and of the local area, such as crime rate and amenities, predict house prices.

You have chosen four features of houses and the local areas, **Feature A, Feature B, Feature C**, and **Feature D.** These can vary from **Very low, Low, Medium, High, to Very High**. For instance, the size of a house could be small so it would be rated as **Low**, or the amount of amenities in the local area could be **high**.

Each feature could predict house price positively, negatively, or could be irrelevant to house price:

If a feature is **positively predictive** of house prices, then if its value is high the house price may be high, and when it is low the house price may be low. For instance, large houses may be more expensive than small houses.

If a feature is **negatively predictive** of house prices, then if its value is high the house price may be low and when its value is low the house price may be high. For instance, houses with local areas that have low crime rates may be more expensive than houses with local areas that have high crime rates

If a feature is **irrelevant** to house prices, then it is not predictive of house prices. For instance, the size of nearby roads may not affect house prices.

On each trial you will be presented a different house in a different area. Using the levels (from very low to very high) you will be required to judge the price of the house by selecting a label from very low to very high. Once you have made each judgment you will receive feedback as to the actual price level of the house. You may find this task difficult at first, but use the feedback to improve your judgments.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

Instructions for predicting "low" house prices

You have chosen to buy a house, however, you do not have much money so you are only interested in finding houses with low house prices. You would like to know which features of houses, such as their size and age, and of the local area, such as crime rate and amenities, predict house prices.

You have chosen four features of houses and the local areas, **Feature A**, **Feature B**, **Feature C**, and **Feature D**. These can vary from **Very low**, **Low**, **Medium**, **High**, to **Very High**. For instance, the size of a house could be small so it would be rated as **Low**, or the amount of amenities in the local area could be **high**.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

Each feature could predict house price positively, negatively, or could be irrelevant to house price:

If a feature is **positively predictive** of house prices, then if its value is high the house price may be high, and when it is low the house price may be low. For instance, large houses may be more expensive than small houses.

If a feature is **negatively predictive** of house prices, then if its value is high the house price may be low and when its value is low the house price may be high. For instance, houses with local areas that have low crime rates may be more expensive than houses with local areas that have high crime rates

If a feature is **irrelevant** to house prices, then it is not predictive of house prices. For instance, the size of nearby roads may not affect house prices.

On each trial you will be presented a different house in a different area. Using the levels (from very low to very high) you will be required to judge the price of the house by selecting a label from very low to very high. Once you have made each judgment you will receive feedback as to the actual price level of the house. You may find this task difficult at first, but use the feedback to improve your judgments.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

Experiment 7.

Instructions for participants performing the ''standard' task first

"JUDGMENT TASK ONE"

Four personality tests are being investigated by a business for their ability to predict the suitability of job applicants. However, these tests could predict suitability both positively and negatively. For example, extraversion could predict good applicant suitability for some occupations, but introversion could be a good predictor of suitability for others. Other tests could be irrelevant and fail to predict applicant suitability either way.

You will be required to examine how good each test is by rating the suitability of each applicant based on their results at the time of applying for the job. You will then be told how suitable each applicant turned out to be six months into the job. Use this information to try to improve your prediction of the job success of later applicants. You may find this task difficult at first, but as the trials go on you may find your judgments improve.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

When you begin the experiment you will see the four tests labelled from A to D at the top of the screen (See *Figure 1*). Along side each test you will see the results of the applicant rated as either; **Very Low, Low, Average, High**, or **Very High.**

Below the tests you will need to select the rating that you think best describes the suitability of the applicant by clicking on the rating using the mouse.

Once you have made your selection the actual suitability of the applicant will appear below (see *Figure 2*). By then clicking on "click here to view next applicant" you will be presented the next applicant.

Remember, each test can predict applicant suitability either **positive, negatively**, or be **irrelevant**:

Positive: If a **high** result on a test predicts **good** applicant suitability, and a **low** result predicts **poor** suitability. For example, high IQ may predict good applicants and low IQ predict poor applicants.

Irrelevant: If a test is **not predictive** of applicant suitability. For example, the height of the applicant may not predict applicant suitability.

Negative: If a **low** result on a test predicts **good** applicant suitability, and a **high** result predicts **poor** suitability. For example, low anxiety may predict good applicants and high anxiety predict poor applicants.

"JUDGMENT TASK TWO"

To help you learn whether each test predicts applicant suitability negatively, positively, or not at all, in this task you will be able to highlight your hypothesis for each test. As you can see in *Figure 7.1.*, when you begin the task the hypotheses for each test will be selected as irrelevant. At this point you will not be able to change your hypotheses selections.

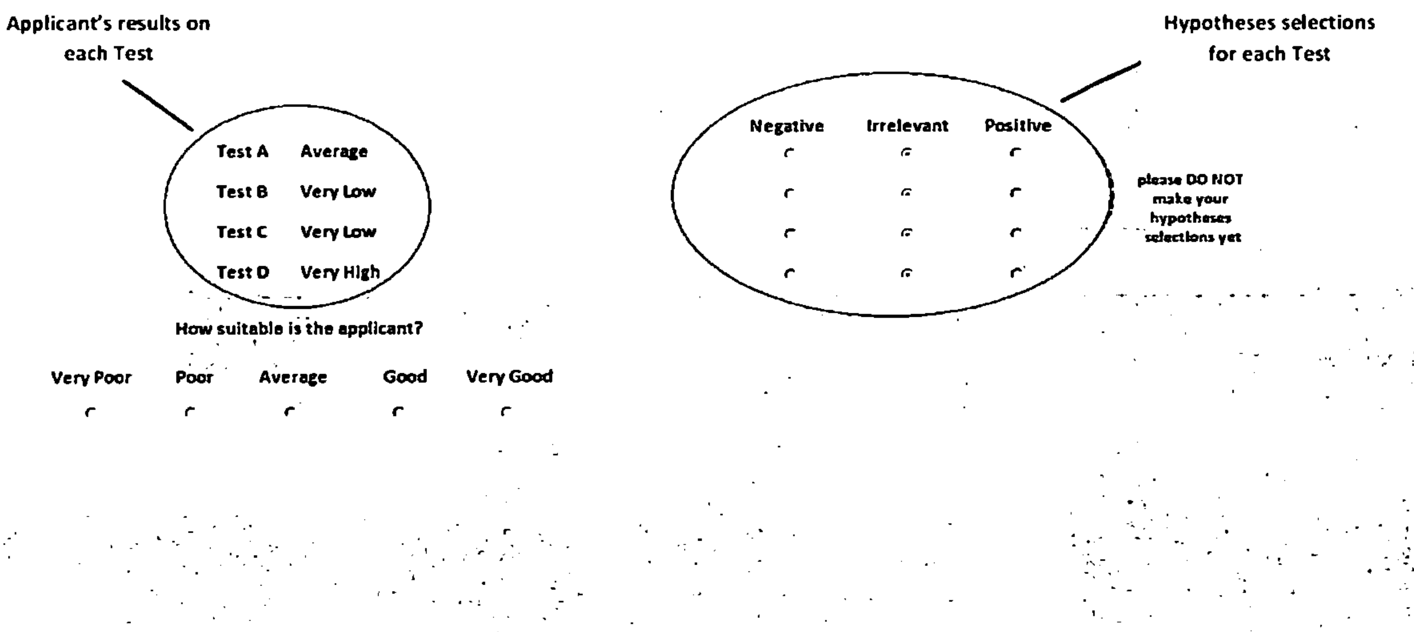

Figure 7.1.

Once you have made your judgment of the applicant's suitability, you will then be presented with the actual suitability of the applicant. You will also be able to change your hypotheses for each test if you wish (see *Figure 7.2.*). Use the feedback about how suitable the applicant turned out to be to help you choose your hypotheses selections.

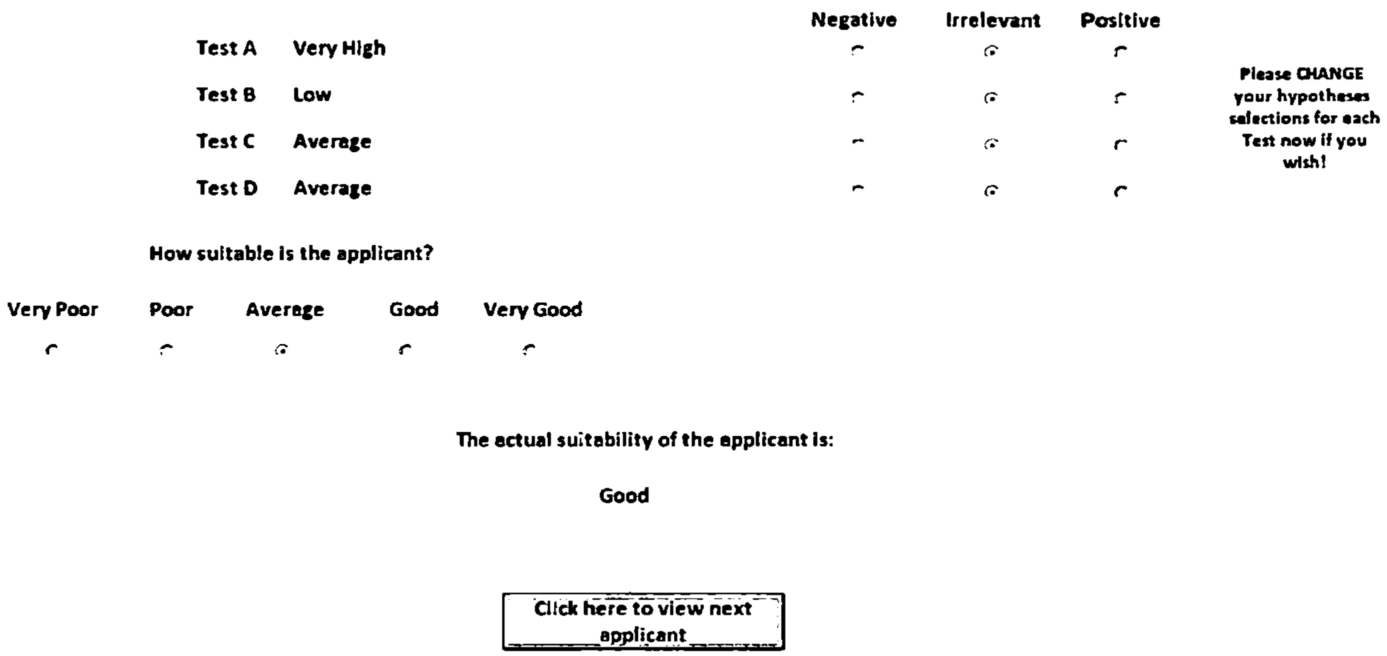

Figure 7.2.

Once you have changed your hypotheses for each test, if you wished to do so, then click on "click here to view next applicant" and the next applicant will appear (see *Figure 7.3.*). This time, use your hypotheses selections for each test to help you make your judgment. You will be able to change your hypotheses selections again once you have made your judgment. This procedure then continues for each trial in the learning phase.

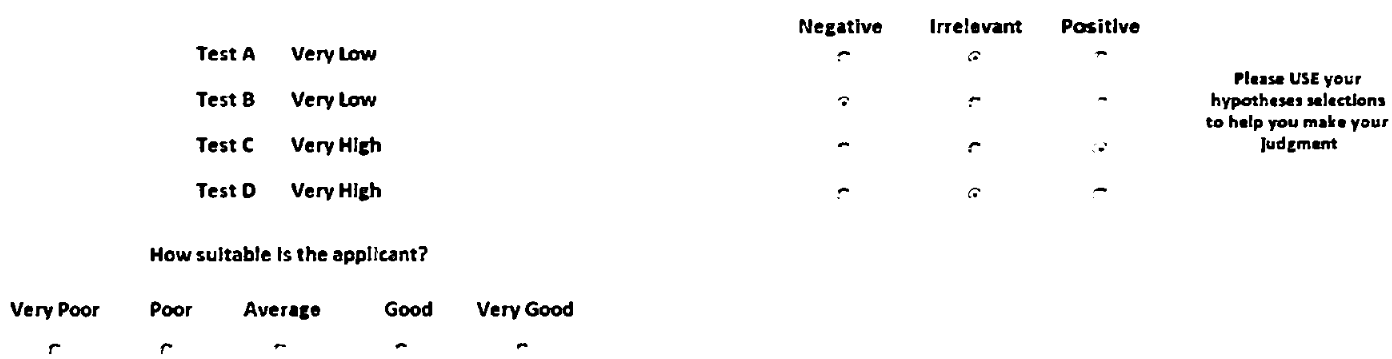

Figure 7.3.

Instructions for participants performing the "hypotheses notepad" task first

"JUDGMENT TASK ONE"

Four personality tests are being investigated by a business for their ability to predict the suitability of job applicants. However, these tests could predict suitability both positively and negatively. For example, extraversion could predict good applicant suitability for some occupations, but introversion could be a good predictor of suitability for others. Other tests could be irrelevant and fail to predict applicant suitability either way.

You will be required to examine how good each test is by rating the suitability of each applicant based on their results at the time of applying for the job. You will then be told how suitable each applicant turned out to be six months into the job. Use this information to try to improve your prediction of the job success of later applicants. You may find this task difficult at first, but as the trials go on you may find your judgments improve.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

When you begin the experiment you will see the four tests labelled from A to D at the top of the screen (See *Figure 1*). Along side each test you will see the results of the applicant rated as either; **Very Low, Low, Average, High**, or **Very High**.

Below the tests you will need to select the rating that you think best describes the suitability of the applicant by clicking on the rating using the mouse.

Once you have made your selection the actual suitability of the applicant will appear below (see *Figure 2*). By then clicking on "click here to view next applicant" you will be presented the next applicant.

Remember, each test can predict applicant suitability either **positive**, **negatively**, or be **irrelevant**:

Positive: If a **high** result on a test predicts **good** applicant suitability, and a **low** result predicts **poor** suitability. For example, high IQ may predict good applicants and low IQ predict poor applicants.

Irrelevant: If a test is **not predictive** of applicant suitability. For example, the height of the applicant may not predict applicant suitability.

Negative: If a **low** result on a test predicts **good** applicant suitability, and a **high** result predicts **poor** suitability. For example, low anxiety may predict good applicants and high anxiety predict poor applicants.

To help you learn whether each test predicts applicant suitability negatively, positively, or not at all, you will be able to highlight your hypothesis for each test. As you can see in *Figure 7.4.*, when you begin the task the hypotheses for each test will be selected as irrelevant. At this point you will not be able to change your hypotheses selections.

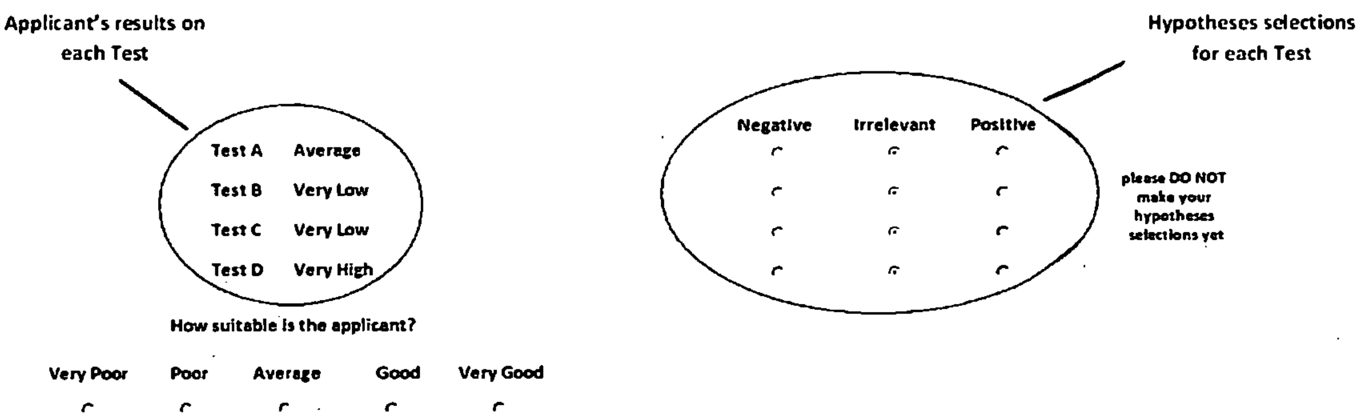

Figure 7.4.

Once you have made your judgment of the applicant's suitability, you will then be presented with the actual suitability of the applicant. You will also be able to change your hypotheses for each test if you wish (see *Figure 7.5.*). Use the feedback about how suitable the applicant turned out to be to help you choose your hypotheses selections.

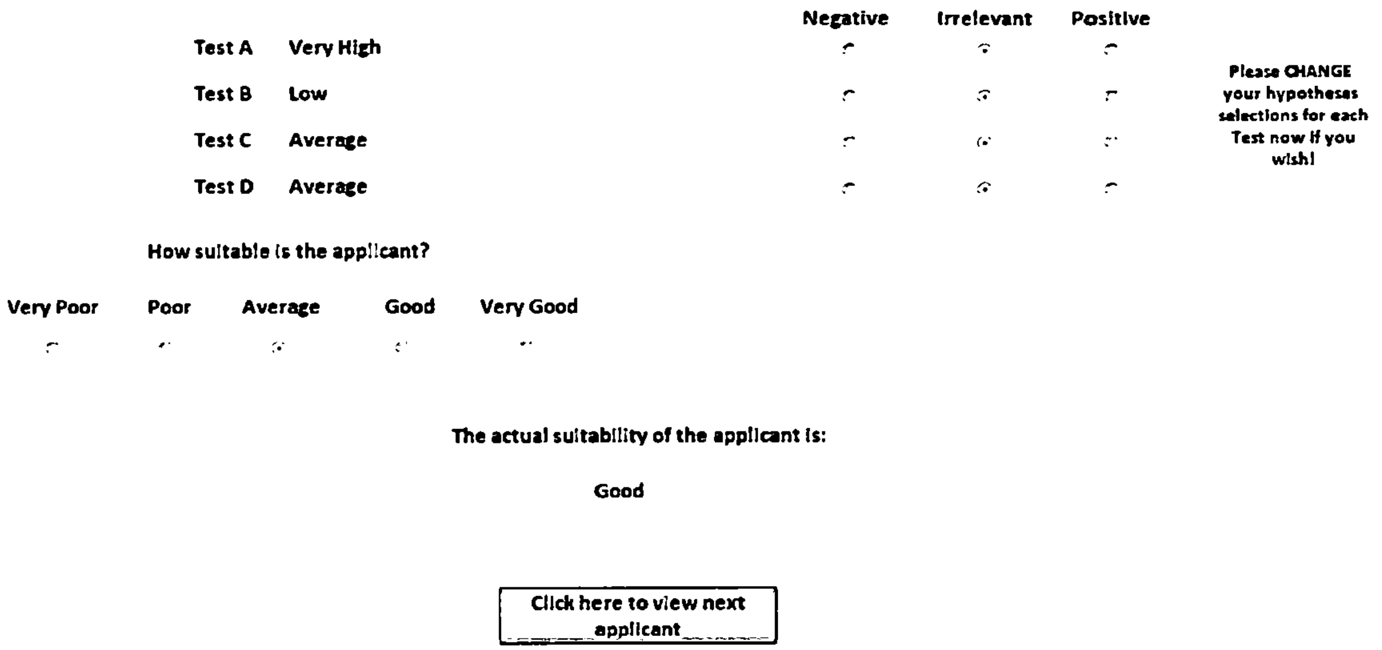

Figure 7.5.

Once you have changed your hypotheses for each test, if you wished to do so, then click on "click here to view next applicant" and the next applicant will appear (see *Figure 7.6.*). This time, use your hypotheses selections for each test to help you make your judgment. You will be able to change your hypotheses selections again once you have made your judgment. This procedure then continues for each trial in the learning phase.

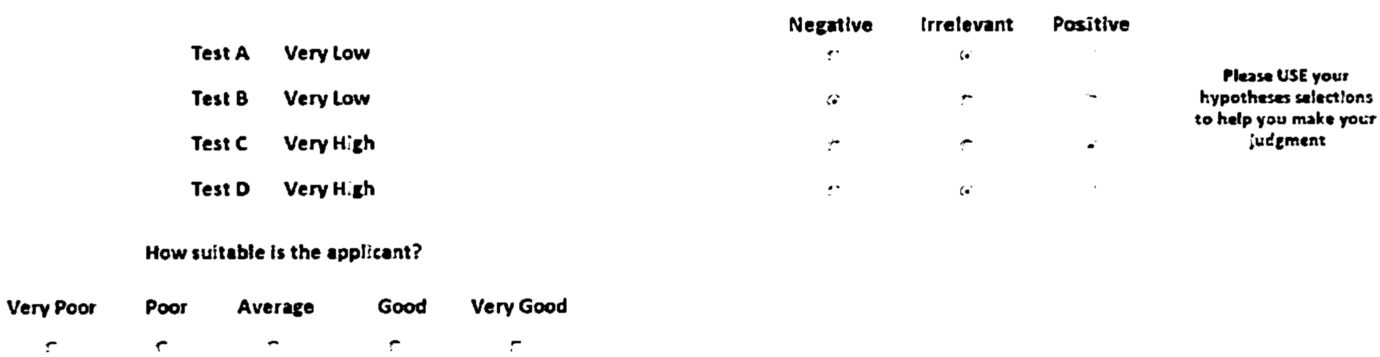

Figure 7.6.

"JUDGMENT TASK TWO"

In this judgement task you will not be able to highlight your hypotheses about each test.

Experiment 8.

Instructions for participants performing the 'hypotheses notepad' task first

"JUDGMENT TASK ONE"

Four personality tests are being investigated by a business for their ability to predict the suitability of job applicants. However, these tests could predict suitability both positively and negatively. For example, extraversion could predict good applicant suitability for some occupations, but introversion could be a good predictor of suitability for others. Other tests could be irrelevant and fail to predict applicant suitability either way.

You will be required to examine how good each test is by rating the suitability of each applicant based on their results at the time of applying for the job. You will then be told how suitable each applicant turned out to be six months into the job. Use this information to try to improve your prediction of the job success of later applicants. You may find this task difficult at first, but as the trials go on you may find your judgments improve.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

When you begin the experiment you will see the four tests labelled from A to D at the top of the screen (See *Figure 1*). Along side each test you will see the results of the applicant rated as either; **Very Low, Low, Average, High,** or **Very High.**

Below the tests you will need to select the rating that you think best describes the suitability of the applicant by clicking on the rating using the mouse.

Once you have made your selection the actual suitability of the applicant will appear below (see *Figure 2*). By then clicking on "click here to view next applicant" you will be presented the next applicant.

Remember, each test can predict applicant suitability either **positive, negatively**, or be **irrelevant**:

Positive: If a **high** result on a test predicts **good** applicant suitability, and a **low** result predicts **poor** suitability. For example, high IQ may predict good applicants and low IQ predict poor applicants.

Irrelevant: If a test is **not predictive** of applicant suitability. For example, the height of the applicant may not predict applicant suitability.

Negative: If a **low** result on a test predicts **good** applicant suitability, and a **high** result predicts **poor** suitability. For example, low anxiety may predict good applicants and high anxiety predict poor applicants.

To help you learn whether each test predicts applicant suitability negatively, positively, or not at all, you will be able to highlight your hypothesis for each test. As you can see in *Figure 7.7.*, when you begin the task the hypotheses for each test will be selected as irrelevant. At this point you will not be able to change your hypotheses selections.

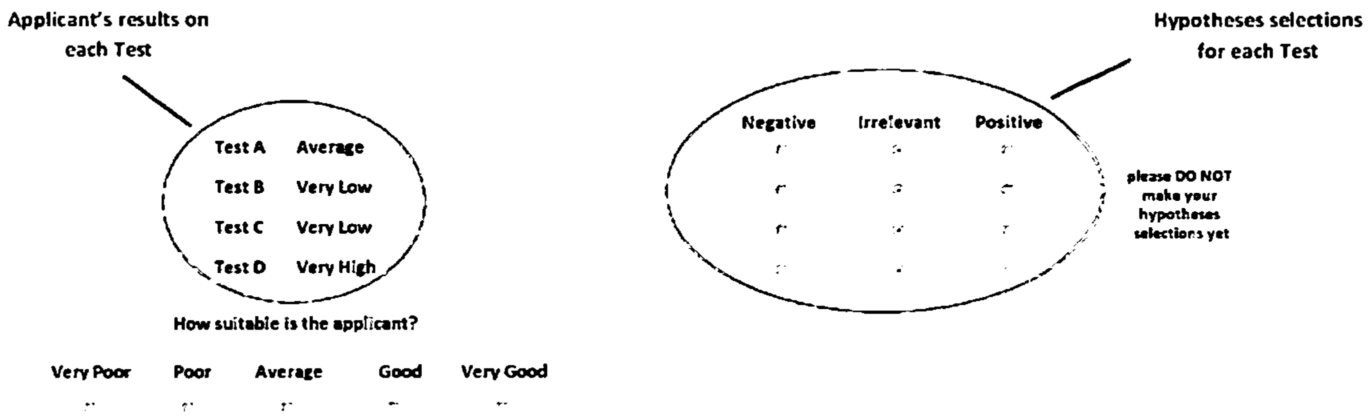

Figure 7.7.

Once you have made your judgment of the applicant's suitability, you will then be presented with the actual suitability of the applicant. You will also be able to change your hypotheses for each test if you wish (see *Figure 7.8.*). Use the feedback about how suitable the applicant turned out to be to help you choose your hypotheses selections.

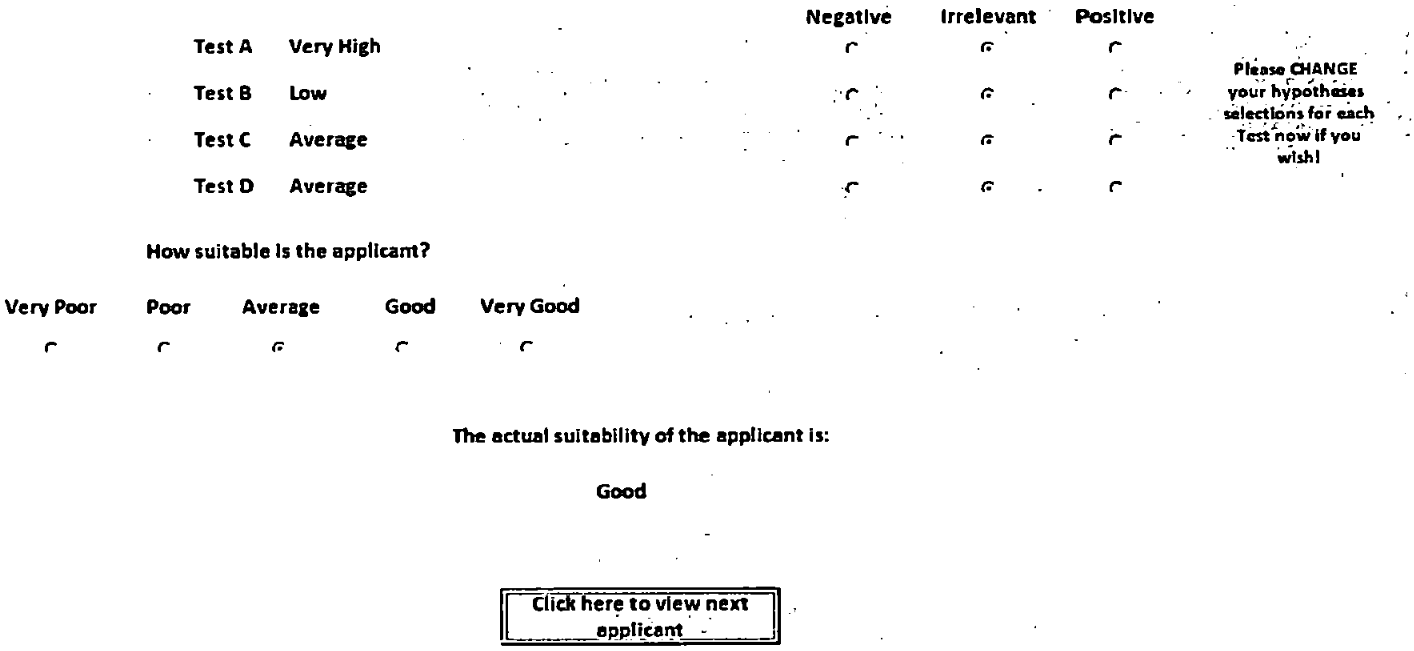

Figure 7.8.

Once you have changed your hypotheses for each test, if you wished to do so, then click on "click here to view next applicant" and the next applicant will appear (see *Figure 7.9.*). This time, use your hypotheses selections for each test to help you make your judgment. You will be able to change your hypotheses selections again once you have made your judgment. This procedure then continues for each trial in the learning phase.

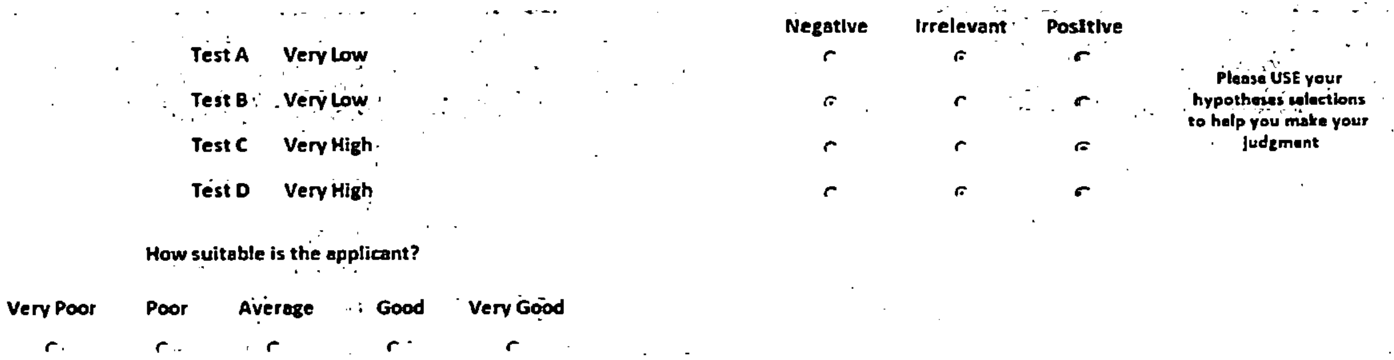

Figure 7.9.

"JUDGMENT TASK TWO"

In this second judgment task you will also be presented the results of your hypotheses selections on each of the learning trials. As you can see in Figure 7.10. below, after making your judgment on each trial you will be presented the results of your hypotheses selections. If you find that the results of your hypotheses selections are incorrect (do not predict applicant suitability) then you can change your hypotheses selections to better predict applicant suitability.

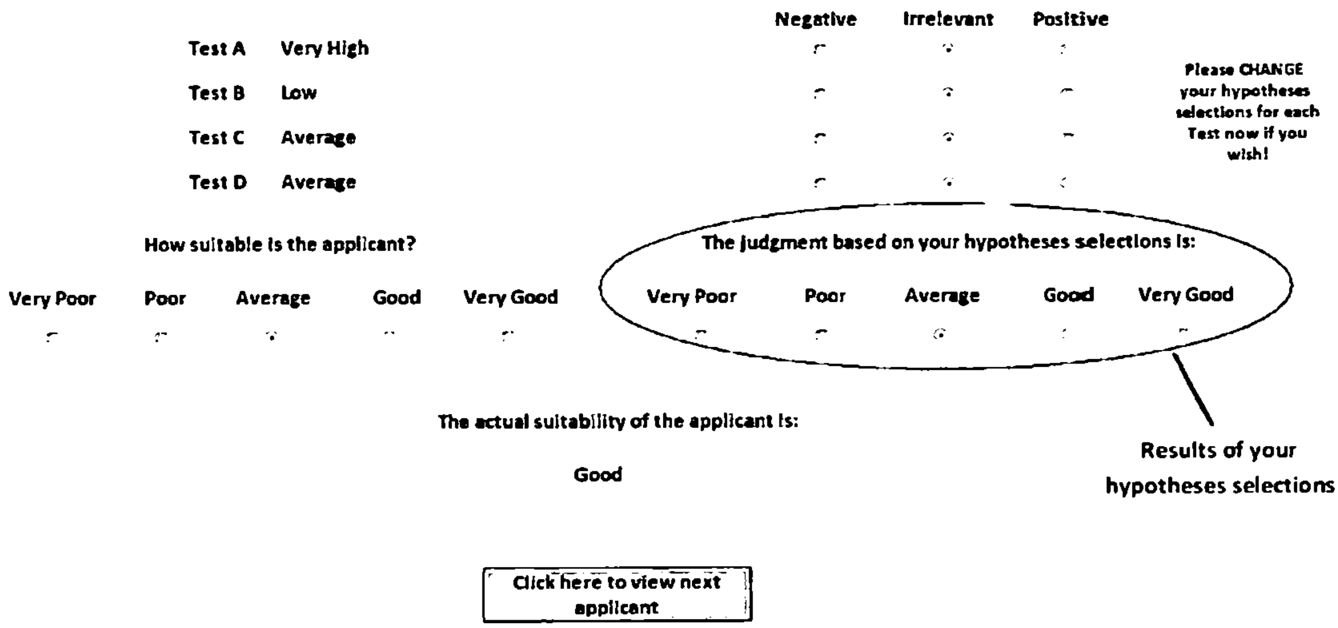

Figure 7.10.

Instructions for participants performing the "hypotheses notepad output" task first

"JUDGMENT TASK ONE"

Four personality tests are being investigated by a business for their ability to predict the suitability of job applicants. However, these tests could predict suitability both positively and negatively. For example, extraversion could predict good applicant suitability for some occupations, but introversion could be a good predictor of suitability for others. Other tests could be irrelevant and fail to predict applicant suitability either way.

You will be required to examine how good each test is by rating the suitability of each applicant based on their results at the time of applying for the job. You will then be told how suitable each applicant turned out to be six months into the job. Use this information to try to improve your prediction of the job success of later applicants. You may find this task difficult at first, but as the trials go on you may find your judgments improve.

Importantly, to represent real-life the feedback about how accurate each applicant turned out to be will not always be accurate. Therefore, even if you work out which tests are predictive and in which way, the feedback you receive in response to your judgments may be inaccurate on some trials.

When you begin the experiment you will see the four tests labelled from A to D at the top of the screen (See *Figure 1*). Along side each test you will see the results of the applicant rated as either; **Very Low**, **Low**, **Average**, **High**, or **Very High**.

Below the tests you will need to select the rating that you think best describes the suitability of the applicant by clicking on the rating using the mouse.

Once you have made your selection the actual suitability of the applicant will appear below (see *Figure 2*). By then clicking on "click here to view next applicant" you will be presented the next applicant.

Remember, each test can predict applicant suitability either **positive**, **negatively**, or be **irrelevant**:

Positive: If a **high** result on a test predicts **good** applicant suitability, and a **low** result predicts **poor** suitability. For example, high IQ may predict good applicants and low IQ predict poor applicants.

Irrelevant: If a test is **not predictive** of applicant suitability. For example, the height of the applicant may not predict applicant suitability.

Negative: If a **low** result on a test predicts **good** applicant suitability, and a **high** result predicts **poor** suitability. For example, low anxiety may predict good applicants and high anxiety predict poor applicants.

To help you learn whether each test predicts applicant suitability negatively, positively, or not at all, you will be able to highlight your hypothesis for each test. As you can see in *Figure 7.11*, when you begin the task the hypotheses for each test will be selected as irrelevant. At this point you will not be able to change your hypotheses selections.

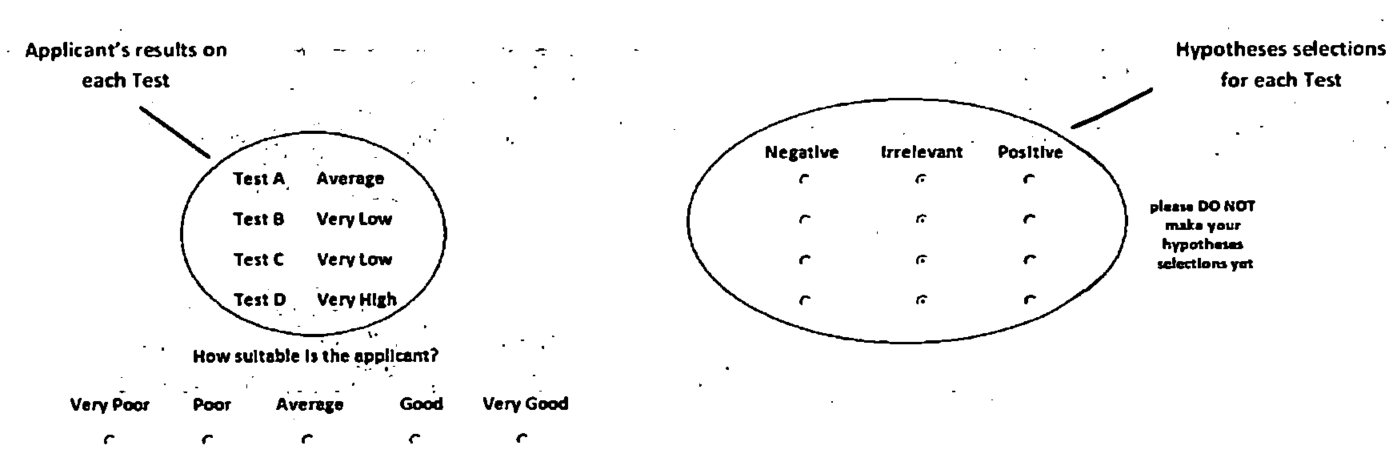

Figure 7.11.

Once you have made your judgment of the applicant's suitability, you will then be presented with the actual suitability of the applicant and the results of your hypotheses selections. You will also be able to change your hypotheses for each test if you wish (see *Figure 7.12*). Use the feedback about how suitable the applicant turned out to be to help you choose your hypotheses selections. If you find that the results of your hypotheses selections are incorrect (do not predict applicant suitability) then you can change your hypotheses selections to better predict applicant suitability.

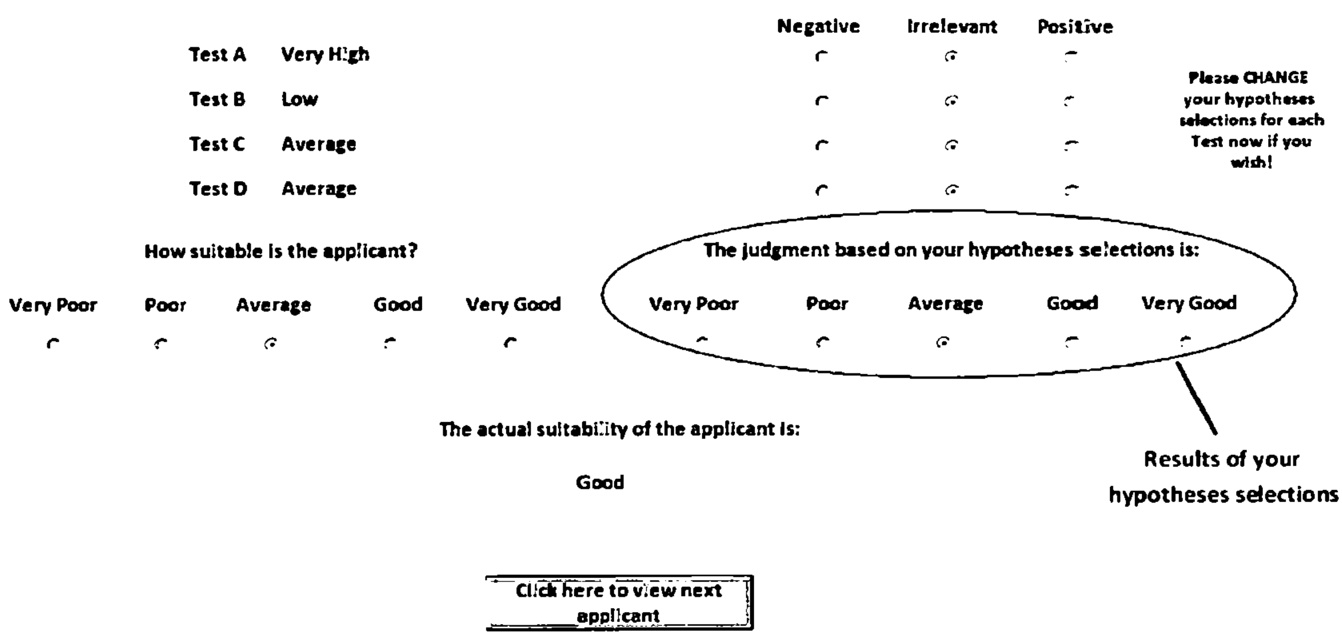

Figure 7.12.

Once you have changed your hypotheses for each test, if you wished to do so, then click on "click here to view next applicant" and the next applicant will appear (see *Figure 7.13.*). This time, use your hypotheses selections for each test to help you make your judgment. You will be able to change your hypotheses selections again once you have made your judgment. This procedure then continues for each trial in the learning phase.

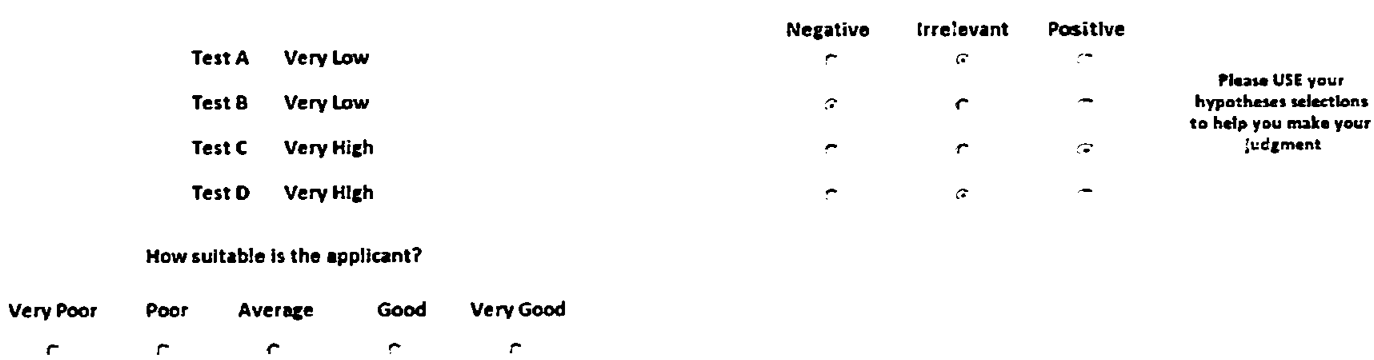

Figure 7.13.

“JUDGMENT TASK ONE”

In this second judgment task you will again be presented your hypotheses selections on each of the learning trials. However, this time, your will not be provided the results of your hypotheses selections.

Reference List

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

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

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (1986). Knowledge compilation: The general learning mechanism. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach Volume II* (pp. 289-328). Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *The Rules of the Mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036-1060.

Anderson, R. B., Doherty, M. E., Berg, N. D., & Friedrich, J. C. (2005). Sample size and the detection of correlation: A signal detection account. *Psychological Review, 112,* 268-279.

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology, 49,* 323-330.

Ashby, F. Q., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105,* 442-481.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56,* 149-178.

Ashby, F. G., & O'Brian, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences, 9*, 83-89.

Athanasou, J. A., & Cooksey, R. W. (2001). Judgment of factors influencing interest: An Australian study. *Journal of Vocational Education Research, 26*, 1-13.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Clarendon Press.

Balke, W. M., Hammond, K. R., Meyer, G. D. (1973). An alternative approach to labor-management relations. *Administrative Science Quarterly, 18*, 311-327.

Balzer, W. K., Doherty, M. E., & O'Conner, R. Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*, 410-433.

Bargh, J. A. (2006). *Social judgment and the unconscious.* New York: Psychology Press.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553-573.

Berry, D. C. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology, 43*, 881-906.

Berry, D. C. (1994). A step too far? *Behavioural and Brain Science, 17*, 397-398.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology, 36A*, 209-231.

Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology, 79*, 251-272.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning.* Hove, UK: Erlbaum.

Blair, I. V. (2001). Implicit stenotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology.* Mahwah, NJ: Lawrence Erlbaum.

Brannen, A. L., Godfrey, L. J., & Goetter, W. E. (1989). Prediction of outcome from critical illness: A comparison of clinical judgment with a prediction rule. *Archives of Internal Medicine, 149,* 1083-1086.

Brehmer, A. (1988). Grading as a quasi-rational judgment process. In J. Lowyck, C. Clark & R. Halkes (Eds.), *Teaching thinking and professional action.* Cambridge, MA: Harvard University Press.

Brehmer, A., & Brehmer, B. (1987). Policy capturing with experimenter defined and subject defined cues. In B. Brehmer, & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 75-114). Amsterdam, Netherlands: Elsevier Science Publishers.

Brehmer, A., & Brehmer, B. (1988). What have we learned about human judgment from thirty years of policy capturing? In B. Brehmer, & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 75-114). Amsterdam, Netherlands: Elsevier Science Publishers.

Brehmer, B. (1973). Single-cue probability learning as a function of the sign and magnitude of the correlation between cue and criterion. *Organizational Behavior and Human Performance, 9,* 377-95.

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance, 11,* 1-27.

Brehmer, B. (1979). Effect of cue validity on learning of complex rules in probabilistic inference tasks. *Acta Psychologica, 44,* 201-210.

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45,* 233-241.

Brehmer, B. (1984). The role of judgment in small-group conflict and decision making. In G. M. Stephenson & J. H. Davis (Eds.), *Progress in Applied Social Psychology* (Vol. 2). New York: Wiley.

Brehmer, B. (1994). The psychology of linear judgment models. *Acta Psychologica, 87,* 137-154.

Brehmer, B., & Kuylenstierna, J. (1978). Task information and performance in probabilistic inference tasks. *Organizational Behavior and Human Performance, 22,* 445-464.

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

Broadbent, D. E., & Aston, B. (1978). Human control of a simulated economic system. *Ergonomics, 21,* 1035-1043.

Broadbent, D. E., FitzGerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology, 77,*33-50.

Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making, 19,* 361-380.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In. E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hilsdale, NJ: Erlbaum.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews *et al.* (1989). *Journal of Experimental Psychology: General, 120,* 316-323.

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686-700.

Brunswik, E. (1944). Distal focussing of perception: Size constancy in a representative sample of situations. *Psychological Monographs, 56,* 1-49.

Brunswik, E. (1955). In defence of probabilistic functionalism: A reply. *Psychological Review, 62,* 236-242.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology.* New York: Guilford.

Chen, Z., & Mo, L. (2004). Schema induction in problem solving: A multidimensional analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 583-600.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2,* 406-416.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235-253.

Cobos, P. L., Almaraz, J., & García-Madruga, J. A. (2003). An associative framework for probability judgment: An application to biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 80-96.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17-30.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review, 12,* 769-786.

Cooksey, R. W. (1996). The methodology of social judgment theory. *Thinking and Reasoning, 2,* 141-173.

Cooper, R. P., & Werner, P. D. (1990). Predicting violence in newly admitted inmates: A lens model analysis of staff decision making. *Criminal Justice and Behaviour, 17,* 431-447.

Curren, T., & Keele, S.W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189-202.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, and the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 97-118). New York: Oxford University Press.

Dawes, R. M. (1975). The mind, the model, and the task. In F. Restle, R. M. Shiffrin, N. J. Catellan, H. R. Lindman, & D. P. Pisoni (Eds.), *Cognitive theory* (Vol. 1). Hilldale, NJ: Erlbaum.

Deane, D. H., Hammond, K. R., & Summers, D. A. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology, 92,* 20-26.

De Caro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107,* 284-294.

De Klerk, L. F. W., De Leeuw, J., & Oppe, S. (1966). Funktioneel leren. *Hypothese, 11,* 10-19.

De Jong, R., Liang, C-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 731-750.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoned. *Psychological Science, 17,* 428-433.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning, 11,* 349-381.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91,* 385-405.

Dienes, Z. (1992). Connectionist and memory-array models of artificial grammar learning. *Cognitive Science, 16,* 41-79.

Dienes, Z., Altmann, G., Gao, S-J., & Goode. (1995). The transfer of implicit knowledge across domains. *Language and Cognitive Processes, 10,* 363-367.

Dienes, Z., & Fahey, R. (1995). The role of specific instances in controlling a dynamic system. *Journal of* Experimental Psychology: Learning, Memory, and Cognition, 21, 848-862.

Dienes, Z., & Perner, J. (1994). Dissociable definitions of consciousness. *Behavioural and Brain Science, 17,* 403-404.

Doherty, M . E., & Kurz, E. M. (1996). Social judgment theory. *Thinking and Reasoning, 2,* 109-140.

Dougherty, M. R. P., & Hunter, J. E. (2003a). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica, 113,* 263-282.

Dougherty, M. R. P., & Hunter, J. E. (2003b). Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval. *Memory and Cognition, 31,* 968-982.

Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 175-197). Malden, MA: Blackwell.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 113,* 541-555.

Dunwoody, P. T., Haarbauer, E., Mahan, R. P., Marino, C., & Tang, C-C. (2000). Cognitive adaptation and its consequences: A test of cognitive continuum theory. *Journal of Behavioral Decision Making, 13,* 35-54.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment.* New York: Wiley.

Eimer, M. (1995). Stimulus-response compatibility and automatic response activation: Evidence from psychophysiological studies. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 837-854.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Performance, 7*, 86-106.

Einhorn, H. J., & Hogarth, R. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*, 395-416.

Einhorn, H. J., & Hogarth, R. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32*, 53-88.

Einhorn, H. J., Kleinmuntz, D, N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86*, 465-485.

Engle, R. W. (2002). Working memory capacity and executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Engle, R. W., Cantor, J., & Carullo, J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992.

Enkvist, T., Newell, B., Juslin, P., & Olsson, H. (2006). On the role of causal intervention in multiple-cue judgment: Positive and negative effects on learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 163-179.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive-experiential theory. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462-482). New York: Guilford Press.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107-140.

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

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75,* 451-468.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* London: Erlbaum.

Evans, J. St. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87,* 223-240.

Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: Do we understand it after 25 years? *Thinking and Reasoning, 4,* 45-82.

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review, 13,* 378-395.

Evans, J. St. B. T. (2007a). *Hypothetical thinking: Dual processes in reasoning and judgement.* Hove, UK: Psychology Press.

Evans, J. St. B. T. (2007b). On the resolution of conflict in dual-process theories of reasoning. *Thinking & Reasoning, 13,* 321-329.

Evans, J. St. B. T. (2008). Dual-process accounts of reasoning, judgment, and social cognition, *Annual Review of Psychology, 59,* 255-278.

Evans, J. St. B.T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition, 11,* 295-306.

Evans, J. St. B.T., Clibbens, J., Cattani, A., Harris, A., & Dennis, I. (2003). Explicit and implicit processes in multicue judgment. *Memory & Cognition, 31,* 608-618.

Evans, J. St. B.T., Clibbens, J., & Harris, A. (2005). Prior belief and polarity in multicue learning. *The Quarterly Journal of Experimental Psychology, 58A,* 651-665.

Evans, J. St. B. T., Harries, C. H., & Dean J. (1995). Tacit and explicit policies in general practitioners' perception of lipid lowering agents. *British Journal of General Practice, 45,* 15-18

Evans, J. St. B.T., & Over, D. E. (1996). *Rationality and Reasoning.* Hove, U.K: Psychology Press.

Evans, J. St. B. T., & Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology, 67,* 479-486.

Faust, D. (1986). Research on human judgment and its applications to clinical practice. *Professional Psychology, 17,* 420-430.

Feeney, A. (2007). How many processes underlie category-based induction? Effects of conclusion specificity and cognitive ability. *Memory and Cognition, 35,* 1830-1839.

Feldman, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*,553-573.

Fodor, J. A. (1983). *The modularity of mind.* Scranton, PA: Crowell.

Folk, R. (1992). A closer look at the probabilities of the notorious three prisoners. *Cognition, 43,* 197-223.

Frensch, P. A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and ambiguous serial transitions in the presence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 567-584.

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fuelled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 966-982.

Ganis, G., & Schendan, H. (1992). Hebbian learning of artificial grammars. In *Proceedings of the 14th annual meeting of the Cognitive Science Society* (pp. 838-843). Hillsdale, HJ: Erlbaum.

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When 'just say no' is not enough: Affirmation verses negation training and the reduction of automatic stereotypic activation. *Journal of Experimental Social Psychology, 44,* 370-377.

Geary, D. C. (2005). *The origins of the mind: Evolution of brain, cognition, and general intelligence.* Washington, DC: American Psychological Association.

Geddes, B. W., & Stevenson, R. J. (1997). Explicit learning of a dynamic system with a non-salient rule. *Quarterly Journal of Experimental Psychology, 50A,* 742-765.

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

Gigerenzer, G., Todd, P., & the ABC Research Group. (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Gilinsky, A. S., & Judd, B. B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging, 9,* 356-371.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language, 27,* 166-195.

Gluck, M. A., Oliver, L. M., & Myers, C. E. (1996). Late-training amnesic deficits in probabilistic category learning: A neurocomputational analysis. *Learning and Memory, 3,* 326-240.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning and Memory, 9,* 408-418.

Goel, V., & Dolan, R. J. (2003). Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *Neuroimage, 20,* 2314-2321.

Goldberg, L. R. (1965). Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI. *Psychological Monographs, 79* (Whole no. 602).

Goldberg, L. R. (1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representations of the human inference process. *Organizational Behavior and Human Performance, 6,* 458-479.

Gomez, R. L., & Schvaneveldt, R. W. (1994). What is learned from artificial grammars? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 396-410.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7,* 497-510.

Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The Psychology of Problem Solving* (pp. 176-206). London: Cambridge University Press.

Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62,* 255-262.

Hammond, K. R. (1966). *The psychology of Egon Brunswik.* New York: Holt, Reinhart & Winston.

Hammond, K. R. (1971). Computer graphics as an aid to learning. *Science, 172,* 903-908.

Hammond, K. R. (1990). Intuitive and analytic cognition: Information models. In A. P. Sage (Ed.), *Concise encyclopedia of information processing in systems and organizations.* Oxford: Pergamon Press.

Hammond, K. R. (1996). Coping with uncertainty: The rivalry between intuition and analysis. In K. R. Hammond, *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice* (pp. 60-93). New York: Oxford University Press.

Hammond, K. R. (2001). Expansions of Egon Brunswik's psychology, 1955-1995. In K. R. Hammond & T.R. Stewart (Eds.), *The essential Brunswik: Beginnings, explications, applications* (pp. 464-479). Oxford, England: Oxford University Press.

Hammond, K. R., & Boyle, J. R. (1971). Quasi-rationality, quarrels, and new conceptions of feedback. *Bulletin of the British Psychological Society, 24,* 103-113.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytic cognition in expert judgment. *IEEE Transactions on Systems, Man, and Cybernetics, 17,* 753-770.

Hammond, K. R., & Stewart, T. R. (2001). Introduction. In K. R. Hammond & T.R. Stewart (Eds.), *The essential Brunswik: Beginnings, explications, applications* (pp. 3-11). Oxford, England: Oxford University Press.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 271-317). New York: Academic Press.

Hammond, K. R., & Summers, D. A. (1965). Cognitive dependence on linear and nonlinear cues. *Psychological Review, 72,* 215-224.

Hammond, K. R., & Summers, D. A. (1972). Cognitive control. *Psychological Review, 79,* 58-67.

Harries, C., Evans, J. St. B. T., Dennis, I., & Dean, J. (1996). A clinical judgment analysis of prescribing decisions in general practice. *Le Travail Humain, 59,* 87-111.

Harries, C., & Harvey, N. (2000). Taking advice, using information and knowing what you are doing. *Acta Psychologica, 104,* 399-416.

Harvey, N. (1995). Why are judgments less consistent in less predictable task situations? *Organizational Behavior and Human Decision Processes, 63,* 247-263.

Hayes, N. A., & Broadbent, D. E. (1988). Two modes of learning in interactive tasks. *Cognition, 28,* 249-276.

Hazeltine, R. E., Grafton, S. T., Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor sequence encoding: A PET study. *Brain, 120,* 123-140.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psycgological Bulletin, 47,* 116-131

Hoffman, P. J., Earle, T. C., & Slovic, P. (1981). Multidimensional functional learning (MFL) and some new conceptions of feedback. *Organizational Behavior and Human Performance, 27,* 75-102.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods, 39,* 101-117.

Hofman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31,* 1369-1385.

Hofman, W., Gschwendner, T., & Schmitt, M. (2005). On implicit-explicit consistency: The moderating role of individual differences in awareness and adjustment. *European Journal of Personality, 19,* 25-49.

Holyoak, K. J., & Gattis, M. (1994). *Behavioural and Brain Science, 17,* 406-407.

Holzworth, J. (2001). Multiple cue probability learning. In K. R. Hammond, & T. R. Stewart. *The essential Brunswik: Beginnings, explications, applications.* New York: Oxford University Press.

Houdé, O., & Moutier, S. (1996). Deductive reasoning and experiential inhibition training: The case of the matching bias. *Current Psychology of Cognition, 15,* 409-434.

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training. *Journal of Cognitive Neuroscience, 12,* 721-728.

Hull, D. L. (2001). *Science and selection: Essays on biological evolution and the philosophy of science.* Cambridge, UK: Cambridge University Press.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-probability studies. *Psychological Review, 71,* 42-60.

Ivan, T., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: A whole-brain fMRI study. *Neuroimage, 8,* 50-61.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, S. J., & Passingham, R. E. (1994). *Journal of Neuroscience, 14,* 3775-3790.

Jones, R. M., & VanLehn, K. (1994). Acquisition of children's addition strategies: A model of impasse-free, knowledge-level learning. *Machine Learning, 16,* 11-36.

Juslin, P., Jones, S., Olsson, H., & Winman, A. (2003). Cue abstraction and exemplar memory in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 924-941.

Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple-cue judgment: A division of labor hypothesis. *Cognition, 106,* 259-298.

Juslin, P., & Olsson, H. (2005). Capacity limitations and the detection of correlation: Comment on Kareev (2000). *Psychological Review, 112,* 256-267.

Julsin, P., Olsson, H., & Olsson, A-C. (2003). Exemplar effects in multiple-cue judgment. *Journal of Experimental Psychology: General, 132,* 133-156.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. *Behavioral and Brain Sciences, 23,* 681-3.

Kahneman, D. (2003). Maps of bounded rationality: Psychology of behavioral economics. *The American Economic Review, 93,* 1449-1475.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge, UK: Cambridge University Press.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. Holyoak, R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-294). Cambridge, UK: Cambridge University Press.

Kahneman, D. Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Kahneman, D. Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kahneman, D. Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582-591.

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General, 126,* 278-287.

Kaufmann, E., & Athanasou, J. A. (2009). A meta-analysis of judgment and achievement as defined by the lens model equation. *Swiss Journal of Psychology, 68,* 99-112.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1999). Implicit stereotyping and prejudice and the primed Stroop task. *Swiss Journal of Psychology, 58,* 241-250.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): The effects of training in the negation of stereotypic associations an stereotypic activation. *Journal of Personality and Social Psychology, 78,* 871-888.

Klauer, C. K., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107,* 852-884.

Klayman, J. (1984). Learning from feedback in probabilistic environments. *Acta Psychologica, 56,* 81-92.

Klayman, J. (1988). Cues discovery in probabilistic environments: Uncertainty and experimentation. *Learning, Memory, and Cognition, 14,* 317-330.

Kleinmuntz, B. (1963). MMPI decision rules for the identification of college maladjustment: The digital computer approach. *Psychological Monographs, 77,* (14, Whole No. 577).

Kleinmuntz, B. (1975). The computer as clinician. *American Psychologist, 30,* 379-387.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3,* 172-179.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 79-91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 169-181.

Koehler, D. K. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110,* 449-519.

Koehler, J. (1996). The base-rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19,* 1-53.

Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83,* 26-52.

Kolb, B., & Wishaw, I. Q. (1990). *Fundamentals of human neuropsychology* (3rd edition). Freeman, New York.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility – A model and taxonomy. *Psychological Review, 97,* 253-270.

Kramer, A. F., Strayer, D. L., & Buckley, J. (1990). Development and transfer of automatic processing. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 305-322.

Krueger, J., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences and cures for the problem-seeking approach to social cognition and behaviour. *Behavioral and Brain Sciences, 27,* 313-376.

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 3-26.

LaDuca, A., Engel, J. D., & Chovan, J. D. (1988). An exploratory study of physicians' clinicial judgment: An application of social judgment theory. *Evaluation and Health Professions, 11,* 178-200.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General, 135,* 162-183.

Lee, J. W., & Yates, J. F. (1992). How quantity judgment changes as the number of cues increases: An analytic framework and review. *Psychological Bulletin, 112,* 363-377.

Lindell, M. K. (1976). Cognitive and outcome feedback in multiple-cue probability learning tasks. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 739-745.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 189-207.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 883-914.

Logan, G. D. & Klapp, S. T. (1991). Automatizing alphabet arithmetic: Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 179-195.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 3-26.

Luchins, A. S. (1942). Mechanization in problem solving – the Einstellung effect. *Psychological Monographs, 54,* 95.

Luchins, A. S., & Luchins, E. H. (1959). *Rigidity of behavior: A variational approach to the effect of Einstellung.* Oxford, England: University of Origan Press.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics, 53,* 49-70.

Maddox, W. T., & Ashby, F. G. (2004). Human category learning. *Annual Review of Psychology, 56,* 149-178.

Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning based system in perceptual category learning. *Psychological Bulletin and Review, 11,* 945-952.

Maddox, W. T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category number impacts rule-based but not information-integration category learning: Further evidence for dissociable category learning systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 227-235.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory and Cognition, 17,* 11-17.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1083-1100.

Mear, R., & Firth, M. (1987). Assessing the accuracy of financial analyst security return predictions. *Accounting, Organizations and Society, 12,* 331-340.

Millimet, C. R., & Greenberg, R. P. (1973). Use of an analysis of variance technique for investigating the differential diagnosis of organic versus functional involvement of symptoms. *Journal of Consulting and Clinical Psychology, 40,* 188-195.

Mitchell, C. J., De Houwer, J. D., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32,* 183-198.

Moutier, S., Houdé, O. (2003). Judgment under uncertainty and conjunction fallacy inhibition training. *Thinking and Reasoning, 9,* 185-201.

Naylor, J. C., & Clark, R. D. (1968). Intuitive inference strategies in interval learning tasks as a function of validity magnitude and sign. *Organizational Behavior and Human Performance, 3,* 378-399.

Naylor, J. C., & Domine, K. D. (1981). Inferences based on uncertain data: Some experiments on the role of slope magnitude, instructions, and stimulus distribution shape on the learning of contingency relationships. *Organizational Behavior and Human Performance, 27,* 1-31.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson, *Cognitive skills and their acquisition,* (pp. 57-84). Hillsdale, NJ: Lawrence Erlbaum Associates.

Newell, A., & Simon, H. A. (1961). Computer simulation of human thinking. *Science, 134,* 2011-2017.

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

Newton, J. R. (1965). Judgment and feedback in a quasi-clinical situation. *Journal of Personality and Social Psychology, 1,* 336-342.

Nisbett, R. E., & Schachter, S. (1966). Cognitive manipulation of pain. *Journal of Experimental Social Psychology, 2,* 227-236.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Nissen, M. J., & Bullemer, P. (1987). Attention requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1-32.

Nosofsky, R. M., & Johanson, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin and Review, 7,* 375-402.

Ouellet, M-C., Beauchamp, M. H., Owen, A. M., & Doyon, J. (2004). Acquiring a cognitive skill with a new repeating version of the tower of london task. *Canadian Journal of Experimental Psychology, 58,* 272-288.

Pascual-Leone, A., Grafman, J., & Hallett, M. (1994). Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science, 263,* 1287-1289.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Perruchet, P., Gallego, J., & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology, 22,* 493-516.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge. *Journal of Experimental Psychology: General, 119,* 264-275.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997 November). *Learning to classify probabilistically: Changes in cortical activity revealed by functional MRI.* Paper presented at the 1997 Society for Neuroscience annual meeting, New Orleans, LA.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin, 133,* 227-244.

Price, A. L. (2005). Cortico-striatal contributions to category learning: Dissociating the verbal and implicit systems. *Behavioral Neuroscience, 119,* 1438-1447.

Price, A. L. (2006). Explicit category learning in Parkinson's disease: Deficits related to impaired rule generation and selection processes. *Neuropsychology, 20,* 249-257.

Proctor, R. W., & Wang, H. (1997). Differentiating types of set-level compatibility. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus-response compatibility* (pp. 11-31). Amsterdam: North-Holland.

Reber, A. S. (1967). Implicit learning and artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855-863.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 88-94.

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

Reber, A. S. (1993). *Implicit learning and tacit knowledge.* Oxford, UK: Oxford University Press.

Reber, A. S., & Lewis, S. (1977). Toward a theory of implicit learning: The analysis of the form and structure of a body of tacit knowledge. *Cognition, 5,* 333-361.

Reder, L. M. (1987). Strategy selections in questions answering. *Cognitive Psychology, 19,* 90-138.

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 227-259). New York: Academic Press.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A re-evaluation. *Journal of Experimental Psychology: General, 125,* 123-138.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 585-594.

Reilly, B. (1996). Self-insight, other-insight, and their relation to interpersonal conflict. *Thinking and Reasoning, 2,* 213-223.

Reilly, B., & Doherty, M. E. (1992). The assessment of self-insight in judgment policies. *Organizational Behavior and Human Decision Processes, 53,* 285-309.

Roose, J. E., & Doherty, M. E. (1976). Judgment theory applied to the selection of life insurance salesmen. *Organizational Behavior and Human Performance, 16,* 231-249.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* Cambridge, USA: The MIT Press.

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin and Review, 5,* 197-220.

Schmitt, N. (1977). Interrater agreement in dimensionality and combination of assessment centre judgments. *Journal of Applied Psychology, 62,* 171-176.

Schmitt, N., Coyle, B. W., & King, L. (1976). Feedback and task predictability as determinants of performance in multiple-cue probability learning tasks. *Organizational Behavior and Human Performance, 16,* 388-402.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioural and Brain Science, 17,* 367-447.

Sheets, C. A., & Miller, M. J. (1974). The effect of cue-criterion function form on multiple-cue probability learning. *The American Journal of Psychology, 87,* 629-641.

Simon, H. A. (1976). Discussion: Cognition and social behaviour. In J. S. Carroll, & J. W. Payne (Eds.), *Cognition and social behavior.* Hillsdale, NJ: Erlbaum.

Singh, H. (1990). Relative evaluation of subjective and objective measures of expectations formation. *Quarterly Review of Economics and Business, 30,* 64-74.

Sloman, S. A. (1996). The empirical case for two reasoning systems. *Psychological Bulletin, 119,* 3-22.

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379-396). New York, US: Cambridge University Press.

Slovic, P. (1969). Analyzing the expert judge: A descriptive study of a stockbroker's decision process. *Journal of Applied Psychology, 53,* 255-263.

Slovic, P. (1974). Hypothesis testing in the learning of positive and negative linear functions. *Organizational Behavior And Human Performance, 11,* 368-376.

Slovic, P., Fleissner, D., & Bauman, W. S. (1972). Analyzing the use of information in investment decision making: A methodological proposal. *Journal of Business, 45,* 283-301.

Slovic, P., & Lichtenstein, S. (1971). Statistical and subjective weights: Some problems and proposals. *Organizational Behavior and Human Performance, 6,* 649-744.

Smedslund, J. (1955). *Multiple probability learning.* Oslo, Norway: Oslo University Press.

Smith, E. R., & DeCoster, J. (2000). Dual process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology, 4,* 108-131.

Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. *Neuroscience and Biobehavioral Reviews, 32,* 249-264.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies in categorization. *Cognition, 65,* 167-196.

Smith, E. E., Patalano, A. L., Jonides, J., & Koeppe, R. A. (1996). *PET evidence for difference categorization mechanisms.* Paper presented at the 37th Annual Meeting of the Psychonomic Society, Chicago.

Stanley, W. B., Mathews, R. C., Buss, R. R., & Kotler-Cope, S. (1989). Insight without awareness: On the interaction of verbalization, instruction, and practice in a simulated process control task. *Quarterly Journal of Experimental Psychology, 41,* 553-577.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Elrbaum.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin.* Chicago: Chicago University Press.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomic minds: Is it time for a tri-process theory? In J. Evans, & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford: England: Oxford University Press.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*, 342-357.

Stanovich, K. E., & West, R. F. (1998a). Cognitive ability and variation in selection task performance. *Thinking and Reasoning, 4*, 193-230.

Stanovich, K. E., & West, R. F. (1998b). Individual differences in framing and conjunction effects. *Thinking and Reasoning, 4*, 289-317.

Stanovich, K. E., & West, R. F. (1998c). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161-188.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioural and Brain Sciences, 23*, 645-726.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*, 672-695.

Stevenson, R. J., & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology, 48*, 613-643.

Stewart, T. (1988). Judgment analysis: Procedures. In B. Brehmer & C. R. B. Joyce (Eds.), *Human Judgment: The SJT view* (pp. 41-74). New York: North-Holland.

Stewart, T. R. (1990). Notes and correspondence: A decomposition of the correlation coefficient and its use in analyzing forecasting skill. *American Meteorological Society, 5*, 661-666.

Stewart, T. R., Roebber, P. J., & Bosart, L. F. (1997). The importance of the task in analyzing expert judgment. *Organizational Behavior and Human Decision Processes, 69*, 205-219.

Summers, D. A., & Hammond, K. R. (1966). Inference behavior in multiple cue tasks involving both linear and nonlinear relations. *Journal of Experimental Psychology, 71*, 751-757.

Sun, R., Merrill, E., & Peterson, T. (1998). A bottom-up model of skill learning. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 1037-1042). Mahwah, NJ: Erlbaum.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science, 25*, 203-244.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review, 112*, 159-192.

Todd, P. M., & Gigerenzer, G. (2003). Bounding rationality to the world. *Journal of Economic Psychology, 24*, 143-165.

Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch and by Hammond, Hursch, and Todd. *Psychological Review, 71*, 528-530.

Turner, C. W., & Fischler, I. S. (1993). Speeded tests of implicit knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1165-1177.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-32.

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

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293-315.

Uhl, C. N. (1960). *An analysis of dimensional judgment concepts.* Unpublished doctoral dissertation, University of Oregan.

Ullman, D. G., & Doherty, M. E. (1984). Two determinants of the diagnosis of hyperactivity: The child and the clinician. *Advances in Developmental and Behavioral Pediatrics, 5,* 167-219.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order ability. *Psychological Bulletin, 133,* 1038-1066.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498-505.

Vadeboncoeur, I., & Markovits, H. (1999). The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task. *Thinking and Reasoning, 5,* 97-113.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 18,* 328-344.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin and Review, 8,* 168-176.

Waller, W. S. (1988). Brunswikian research in accounting and auditing. In B. Brehmer & C. R. B. Joyce (Eds.), *Human Judgment: The SJT view* (pp. 247-272). New York: North-Holland.

Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition, 3,* 141-154.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content.* London, UK: Batsford.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Kiske, G. Linzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 446-496). New York: McGraw-Hill.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology, 100,* 930-941.

Wigton, R,W. (1996). Social judgment theory and medical judgment. *Thinking and Reasoning, 2,* 175-190.

Wigton, R. W., Hoellerich, V. L., & Patil, K. D. (1986). How physicians use clinical information in diagnosing pulmonary embolism: An application of conjoint analysis. *Medical Decision Making, 2,* 2-11.

Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science, 10,* 531-534.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1047-1060.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Harvard University Press.

Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology, 55,* 493-518.

Wilson, T. D., Lindsey, S., Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101-126.

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology, 75,* 1423-1441.

Wittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122,* 227-248.

Wittlesea, B. W. A., Write, R. L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 181-200.

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

Woodworth, R. (1938). *Experimental psychology.* New York: Holt.