The effects of success on task enjoyment and persistence

Richard Remedios

Thesis submitted to the University of Stirling for the degree of Doctor of Philosophy.

March 2000

Contents

Abstract

This thesis explored two issues: Firstly, how participants would respond, in terms of task persistence and task enjoyment, to differing levels of success, when a task was presented to them with a mastery-focus (Experiments 1-5). Secondly, whether improving at task caused participants to enjoy tasks more than achieving a constant level of success (Experiments 6-10).

Experiments 1-3 provided evidence that when participants were given the opportunity to persist with a task for as long as they wanted, they persisted longer after performing poorly. However, despite persisting longer, they did not enjoy the task. Experiments 4-5 adopted the same paradigm as Experiments 1-3, but included a second free-choice persistence phase where participants were unaware their behaviour was being monitored. In Experiments 4 and 5, participants who performed poorly persisted longer initially, but less during the subsequent free-choice phase. Again, those who performed poorly during the initial phase reported that they did not enjoy the task. It was suggested that neither the achievement-goal theories of Nicholls (1984) and Dweck (1986) nor Deci's (1975) theory of intrinsic motivation could adequately account for the persistence behaviours observed in the second persistence phase in Experiments 4 and 5. Instead, it was suggested that participants persisted because of the pleasure derived from solving the problems.

Experiments 6-10 examined the role of improvement in task enjoyment. Experiments 6 and 7 were control studies intended to establish whether the paradigm was appropriate

to examine improvement. Experiments 8-9 showed that relative to achieving a consistent level of performance, improvement increased task enjoyment. However, this result was found only when participants did well; when they did poorly at a task, improvement produced less enjoyment (Experiment 10). Both results can be explained if participants' expectations are taken into account as well as their rate of success.

The final conclusions chapter discusses the types of achievement targets individuals might set themselves when what constitutes good performance at a task is ambiguous, and relates this analysis to the findings from all ten experiments.

Acknowledgements

There are many people I need to thank for providing valuable support, advice and inspiration. Firstly, to my supervisor David Lieberman who not only agreed to take on a project which was not his specialised field, but guided me through the various drafts with considerable patience and precision. I suspect that I have still to realise the full impact of his tutelage.

Thanks are also due to the Department of Psychology here at Stirling University, especially Lindsay Wilson, for constantly putting teaching my way to allow me to self-fund the first two years of this thesis, and more latterly, for the studentship which funded the last two years of my studies. Thanks also to Jill McCormick for helping to run Experiments 6-10 in this thesis, and to Sarah for proof-reading the final draft.

Thanks to the poker crew for those heady nights of sublime bluffing and psychological torment, and to Derek for providing various golfing challenges, and particularly for allowing me to experience a variance in golf shots that redefine the term "outliers". More recently, thanks to Paul and Tim for Friday afternoon surreal tea-breaks, I'm not sure I could have got through the resubmission without this important nonsense.

I must also acknowledge those very special people in my life whose remarkable insight and wisdom go beyond the ordinary. In this regard, I thank Corinne, and most especially, Robert. Thanks also to Jamie for being a good friend and soul-mate. Lastly, I offer thanks to my family whom I do, and always will, love very much.

1 Introduction

The term "success" is ubiquitous; teachers talk of successful students; sports journalists talk of successful sportsman, indeed, we talk of success as if there is an agreed understanding of what it is. In education, the American grade-point-average system purports to be an index of educational success; in golf, success is measured in terms of how few shots it takes to play a series of 18 holes; in marathon running, success in measured in how quick an athlete can run the 26 miles. The implicit claim within all these examples is that greater performance is equivalent to greater success.

Psychologists are also interested in the concept of success but in a different way. For them, success is not about the score that an individual attains, but how that individual responds to objectively measured success (or failure). For Psychologists, absolute scores may be a normative indicator of success, but they tell us little about individuals' experience of success. Normative scores are merely stimuli that individuals respond to; for psychologists, the interest lies in why individuals respond differently to the same normative scores. Built into a lay understanding of success seems to be an implicit recognition that success is largely relative. For example, we would not expect an A grade student to be as satisfied with a B+ as a B grade student, and we would not expect a golfer who is used to going round in par to be as satisfied with an eight over par score as a golfer who usually goes round in ten over par. It seems that we implicitly recognise that different individuals set themselves different goals and that success is as much about achieving an ambition as it is about achieving some absolute standard.

This thesis is concerned with addressing issues surrounding individuals' experiences of success and in particular, investigating questions such as: how do individuals define success? How do individuals respond to differing levels of success in achievement situations? What factors influence their behaviour in these situations? What theories best account for their behaviours? This introductory chapter reviews the literature regarding individuals' experiences in achievement settings with particular emphasis on achievement motivation theory. Chapter two introduces the five experiments that subsequently appear in Chapter three. These five experiments investigate how individuals respond, in terms of task persistence and task enjoyment, to differing levels of success. Chapter four goes on to investigate a potential feature of success, namely, improvement. Five experiments investigate whether or not improving at a task relative to attaining a consistent score significantly affects individuals' experiences of that task. Chapter five offers several alternative explanations for the findings from chapters three and four. This chapter culminates in a discussion of how individuals might construct standards for themselves and how these standards are crucial in determining their subsequent experiences of success.

1.1. Success as a motive

Implicit in an investigation of how individuals respond to success in achievement situations is a taken-for-granted premise that they want to be successful in the first place. But to what extent is it true that being successful figures high in a motivational hierarchy? Also, if being successful is important, what plausible reasons are there for this? To address these questions, I will begin by briefly reviewing motivation theory, explaining how the drive to be successful became established in the motivational literature.

Early motivational research was based on the premise that all organisms are motivated to fulfil primary needs. For example, Hull (1943, p. 17) stated that " ... *when any of the commodities or conditions necessary for individual or species survival are lacking ... a state of primary need is said to exist*". Hull (1943, p. 59-60) suggested these primary needs were restricted to states such as hunger, thirst, air, avoidance of tissue injury, maintenance of optimal temperature, and sex. Satisfaction of these needs was thought to end the behaviour which brought about the satiation of the need, for example, searching-for-food behaviour would be terminated by the finding and eating of food. Motivation was thus characterised as that which energised behaviour, and the motives for behaviour were restricted to those related to satisfying primary needs. A second feature of Hull's theory was *habit* which Hull defined as " *a well-worn mode of action*" (p.102). Hull's (classic) theory was that performance = drive x habit, so that the greater the need to obtain a reinforcer and the more the behaviour had been reinforced in the past, the greater the tendency to behave in that particular way.

On the face of it, Hull's theory afforded little room for the concept of a need for success, as it was based so rigidly on the concept of biological need, and certainly none of Hull's empirical research (largely with animals) attempted to cater for such a concept. However, Hull's theory did not deny the possibility that individuals mental states may also play a role in determining behaviour. He believed that *intervening* variables, if defined precisely enough, could also be included as hypothetical variables as long as they were functionally related to antecedent and consequent events (see Hull, 1943, pp. 22-23). Hull was extremely wary of inventing hypothetical entities (e.g., ".. *the grossest fallacies may be committed"* p.22) but nevertheless, internal states were at least becoming acknowledged as a relevant feature of motivated behaviour.

One problem with Hull's theory was that clearly there are some human behaviours that occur when primary drives are lacking. For example, Miller (1948; summarised in Weiner, 1992) provided evidence that animals learned that some situations were dangerous and a priori attempted to avoid them. This suggested to Miller that the rats had learned to expect fear and therefore to anticipate it. The finding that "anticipated fear" seemed to be a reason for behaviour had two important implications. Firstly, it indicated that motivation to behave did not necessarily require a deficit in tissue needs to drive that behaviour. Secondly, it also suggested that anticipations of future outcomes were relevant, that is, that expectations of future outcomes determined the direction and intensity of behaviour.

However, Hull was still uncomfortable employing mentalistic explanations (e.g. expectations) as the ultimate causes of behaviour. He suggested that just because organisms looked as if they were anticipating a future event, this did not mean they were. He proposed that when an animal was placed in the start-box of a maze and was allowed to run to a goal-box containing food, the responses in the goal box would be connected to the cues that were present, and these responses could then generalise to other situations which signalled the beginning of the response chain, such as the start-box. He labelled these behaviours *anticipatory goal-reactions.* Thus when the animal was returned to the start-box, it would display a *fractional anticipatory goal-response* which would appear to an observer to be an expectation (Hull, 1943; pp. 99-100; see also Weiner, 1992). Fowler and Miller (1963) provided support for Hull's theory. In their study, when rats reached a goal-box containing food, they received a shock to their hind legs before being allowed to eat the food. A control group received no shock. If subjects were expecting to be shocked, then the shocked rats, when placed in the start-box, should have run to the goal more slowly than the non-shocked rats. If, on the hand, the rats generalised behaviour from the goal-box to the start box, then because a shock to the hind paws results in a rat surging forward, then they should exhibit at least a part of this behaviour. This would lead to the prediction that shocked rats should run faster towards the goal-box than the non-shocked rats, and this was the result reported by Fowler and Miller.

Cognitive theorists could not accept that internal states could be accounted for simply in terms of muscle contractions. A study by McFarlane (1930) showed that rats trained to run a maze that was subsequently flooded still swam down the correct alleys to the goal box. Critics of Hullian theory suggested that if the rats had simply learned the muscular responses of running down the alley, then when placed in the start box, they should have exhibited running behaviour, not swimming behaviour. The fact that the rats now swam to the goal-box suggested to cognitive theorists that the rats had an a priori purpose, and were executing the appropriate behaviours to achieve that purpose.

Whether the cognitive retort is a tenable one is still debatable (one could argue that swimming and running behaviour were similar and all the rats were doing were generalising their running behaviour), but what is clear is that such criticisms contributed to a move away from behavioristic approaches to understanding motivation with an increasing emphasis on cognition. Whilst Thorndike (cited in Bolles, 1969, p.

435) was prepared to make the claim that an idea has no power to initiate action, cognitivists fundamentally disagreed. For them, thoughts and especially thoughts about the future and the anticipated consequences of their future behaviour, were important causes of behaviour.

In amongst the criticisms of behaviorism was a series of experiments reported and conducted by Harlow (1953a, 1953b) that suggested that the opportunity to explore novel environments seemed to be at least as rewarding as stimuli such as food. In one example, Davis, Settlage and Harlow (1950) found that rhesus monkeys who were not rewarded with food for successfully completing a manipulation task persisted with that task as long as monkeys who were. Further, it was found that when the food rewards were stopped, the monkeys persisted with the task at the same rate as when they were rewarded. Harlow (1953a) also reported a series of studies where monkeys worked in order to be rewarded by visual stimulation (e.g., Davis, Settlage and Harlow, 1950). In a similar vein to the "anticipated fear" example cited earlier, Harlow emphasised that the causes of behaviour should not be interpreted simply in terms of primary drives such as hunger and thirst that relied on tissue deficits to energise them. Harlow (1953b) found it strange that behaviourists should place such a heavy emphasis on the environment without recognising that features other than food and water were influencing behaviour. Exploration and curiosity became an important explanation for behaviour, especially for intrinsic motivation theorists (see Deci, 1975) and their work will be covered in more detail later in this chapter. The significance of Harlow's work was that it seemed to identify behaviours that appeared to be driven by needs other than tissue deficits or the desire to avoid pain. There seemed to be classes of behaviour that were caused by a variety of other needs.

One such need was identified in probably the most influential of the cognitively-orientated theories of behaviour in performance situations, namely, Achievement Motivation theory (McClelland, Atkinson, Clark and Lowell, 1953). McClelland et al believed that individual motives were best analysed from responses to fantasy-based questions. Using the Thematic Apperception Test (TAT) projective technique previously utilised by Murray (1938) to devise a taxonomy of human needs, McClelland and coworkers gave participants a series of photos portraying everyday scenes in life - e.g. two men in a machine shop; two women who appear to be chemists. Participants were asked to describe what might be going on in the photos. Among several other categories, statements such as "The men are the Wright brothers" or "The women want to do a perfect job" were taken as indicating the presence of a desire for achievement. Coupled with the findings of the earlier work of Murray, McClelland et al interpreted answers such as these to mean that individuals had a fundamental and drive-like need to achieve, and that certain situations elicited this desire to achieve. Atkinson (1982) thus suggested that striving for achievement (or motive-disposition) was *".. presumed to be latent until aroused by situation cues which indicate that some performance will be instrumental to achievement"* (p. 12). Note that such an interpretation is now a far cry from Hullian theory which asserted that organisms sought to resolve primary needs. Atkinson's quotation emphasises that the need-to-achieve is latent, and that is only under specific situations that the need is aroused or called into play as a motive. The situation-specific nature of achievement motivation has been taken up by several contemporary theorists (e.g. Nicholls, 1984; Butler, 1992; Thorkildsen and Nicholls, 1998) and will be discussed later in this chapter. What is relevant at this point is to recognise that achievement-motivation theory represented the first cognitive goaldirected theory of behaviour in achievement settings. Contrary to Thorndike's claim that ideas did not initiate action, what individuals thought about in task situations were now directly relevant to their behaviour. Thoughts were not just intervening variables, they were important causes of action. The need to achieve, or, as it is translated in this thesis, the need to succeed, was established as a fundamental motive for behaviour in achievement settings.

Having outlined how the need to be successful emerged as an explanation for behaviour in the motivation research, the rest of this chapter reviews the literature germane to the issues relevant to this thesis. Firstly, there follows a review of what is meant by "task orientation" and how different schools of thought - e.g. intrinsic motivation theorists; achievement motivation theorists – have dealt with this concept. Secondly, there follows a review of McClelland et al's (1953), Nicholls (1984) and Dweck and Leggett's (1988) versions of achievement-goal theory. Included in this review are Weiner's (1985) highly influential attribution theory and Bandura's (1990) self-efficacy theory. Thirdly, there is a review of how task outcomes, namely positive and negative feedback, affect behaviour in task situations. In this review, I re-introduce the literature on task orientation to examine how individuals respond to success and failure when tasks are presented with either a mastery or a performance focus. This latter review suggests that it is unclear how individuals should respond to success and failure when tasks are presented to them with a mastery-focus. It is this issue that is examined empirically in chapter 2. The final review section of this chapter examines the literature regarding how individuals perceive themselves in experimental situations. This section examines the work of Tedeschi and Reis (1981) on impression management, and Page

(1981) on participant behaviour in psychology experiments, and emphasises the need for a critical view of experimental methodology.

1.2 Task orientation

Task orientation refers to how the way a task is presented to individuals can cause them to respond to that task in (crucially) different ways. When the instructions for a task indicate to a person they are going to be evaluated on their performance of that task - for example, they are going to be graded or that there is some indication of how well others have done the task, then this information is thought to create a pressure for the individual to perform well (see Nicholls, 1984). When the instructions for a task de-emphasise the importance of doing well - for example, advising individuals that the experimenter is only interested in their views about the task, - then this information is thought to minimise concerns about doing well and allows individuals to try and master the task (e.g. Harackiewicz, 1979; Grolnick and Ryan, 1987; Koestner, Zuckerman and Olsson, 1990; Ryan, Koestner and Deci, 1991).

These different orientations to a task have been labelled on various (and generally dichotomous) dimensions. For example, intrinsic motivation theorists have anchored these scales by the terms intrinsic-extrinsic or internal-external (e.g. Deci, 1975; Rigby, Deci, Patrick, and Ryan, 1992). According to intrinsic motivation theorists, individuals have a drive-like need to feel competent and self-determined when dealing with their environment (Deci, 1975). Self-determination, in turn, has been defined by DeCharms (1968, p. 328) as *"whenever a person experiences himself to be the locus of causality for his own behaviour"*, that is, whenever individuals feel they are engaging with an activity because they have to rather than because they want to, then this is thought to undermine their feelings of self-determination. According to Deci, whenever perceptions of competence are undermined, individuals are thought to be extrinsically motivated (also referred to as externally regulated). When faced with an interesting task, intrinsic motivation theorists suggest that individuals have a pre-existing (intrinsic) motivation to perform them but that features of the task may undermine this pre-existing motivation. An example is when an individual is specifically instructed to perform a task rather than engaging with the task because they personally choose to. According to DeCharms (1968; see also Deci, 1975), the instruction takes away the individual's perception that he or she has a choice whether or not to engage with the task, thus eliciting an external perspective (e.g., the only reason I am engaging with this task is because someone else wants me to, not because I want to).

Nicholls (1984) labels external orientation as ego-involvement. This is because his position is that the way that some tasks are presented will cause individuals to become concerned about how their ability will be perceived by others. He suggests that *"announcements that skill tasks are being used to test subjects should induce concerns about personal competence, especially if important or valued skills are being used"* (Nicholls, 1984, p. 330). At the other end of his continuum lies task-involvement, labelled as such because Nicholls believes that when individuals become immersed in a task, they become focused on solving the task regardless of how they appear to others. Dweck (1986) prefers to use the terms performance-focused (external orientation) and mastery-focused (internal orientation) to emphasise that individuals are concerned either about performing well or mastering the task. Because, as noted above, these terms have been used in variety of forms by different researchers, they will be hereafter referred to as mastery-focus and performance-focus for the remainder of this thesis.

The common theme of these theories is that the ways in which tasks are presented to individuals are likely to engender different orientations towards that particular task. If this is true, then it would seem plausible that they should also respond to the tasks differently. There is ample evidence to support this hypothesis. For example, Ryan (1982) found that when a task was introduced to participants in either a controlling way, (they were told that the task was reflective of creative intelligence) or in a non-controlling way (they were not told anything about the nature of the task), participants in the controlling condition reported experiencing significantly more tension whilst completing the task. Similarly, Harackiewicz, Manderlink and Sansone (1984) found that participants who expected to be evaluated for their performance on a pinball game spent less time playing the game than a group who were not expecting evaluation and they were also more concerned about how well they did at the task (see also Harackiewicz, Abrahams and Wageman, 1987). In a series of studies, Butler (1989, 1990, 1992, 1993; Butler and Neuman, 1995) has shown various positive consequences of approaching tasks with a mastery-orientation – e.g., increased task persistence; more accurate estimation of quality of personal achievement; less engagement in comparison of work relative to others; and increase in requests for additional information about tasks. Elliot and Harackiewicz (1996, see also Elliot and Harackiewicz, 1994) have expanded the mastery vs. performance orientation distinction to suggest that individuals have three different approaches to tasks: a mastery-approach where the emphasis is on trying to understand and master the task, a performance-approach orientation where the emphasis is on attaining favourable judgements of competence, and performance-avoidance orientations where the emphasis was on avoiding unfavourable judgements of oneself. In these studies, the evidence has shown again that when tasks are approached with a mastery-focus, they are more likely to lead to positive experiences. In short, it seems that whenever task instructions de-emphasise evaluation, then participants seem to enjoy these tasks.

1.3 Achievement-goal theory

It was mentioned earlier in this chapter, that McClelland et al (1953) identified that individuals had a fundamental need to achieve. This theory began as a theory of achievement-motivation but has also been regarded as the first real attempt to construct a theory that focused on the types of cognised goals that individuals pursued. Although never called an achievement-goal theory by McClelland et al, the theory became synonymous with the later achievement-goal theorists.

McCleland et al (1953), suggested that when individuals are faced with tasks in achievement situations, they either approach the task with a view to doing well or they avoid the task altogether because they fear failure. They measured levels of need for achievement (a term they labelled *n* achievement or *n* ach) and found that individuals who were high in *n* ach tended to relish achievement tasks whilst individuals who were low in *n* ach looked for opportunities to avoid such tasks. Throughout the 1960's, need-for-achievement became a prominent underlying explanatory motive for why individuals approached or avoided tasks. However, need for achievement was only half the story. Need for achievement may have been the push mechanism that impelled individuals towards or away from tasks, but researchers also recognised the importance of the goal that individuals were pursuing. This role of incentives, that is, the value of the goal to the organism, had earlier been recognised by Hull (1943). In a similar fashion, achievement-motivation theorists also recognised that the value of a goal interacted with need for achievement to pull individuals towards (or away from) tasks. Additionally, coupled with the recognition that outcomes had values, researchers also recognised that individuals also had expectations of their chances of attaining those outcomes. Atkinson (1957, 1964) incorporated these three components into his model of behaviour to claim that an individual's tendency to approach a task (T_s) was determined by their need for achievement or motive for success (M_s); the probability of success of achievement (P_s) and the incentive value of success (I_s). Thus the attitudes that individuals were likely to have towards a task could be calculated by the formula $T_s = M_s \times P_s \times I_s$. Together with Feather, Atkinson constructed a series of experiments that established the relationship between the three components (see Atkinson and Feather, 1966). For example, they found that individuals who were high on M_s were more likely to choose tasks of intermediate difficulty because tasks that had low probabilities of success were unlikely to be achieved (very difficult tasks), whilst tasks with high probabilities of success (very easy tasks), did not offer the opportunity for them to attain a standard of excellence. Individuals who were low on M_s tended to choose tasks that had high probabilities of success (very easy tasks) that were likely to be achieved or tasks with low probabilities of success (very difficult tasks) because failure at these difficult tasks would not indicate low ability (because even individuals with relatively high ability would fail at these tasks).

Nicholls (1984) offered a slightly different view of the types of motives individuals had in task situations. He suggested that rather than individuals striving to attain standards of excellence as McClelland et al had suggested (McClelland et al, 1953, pp. 78-80), individuals were actually trying to optimise perceptions of competence. That is, individuals were primarily concerned with how this performance would be perceived by others, rather than with their own standards of excellence. This line of reasoning led Nicholls to suggest that when faced with different task situations, individuals would amend their goals relative to the degree to which they felt they were going to be evaluated by others. According to Nicholls, whilst individuals aspire to attain competence and display ability, when situations indicate that they are about to be evaluated, then they would no longer just try to prove competence to themselves, but would be concerned about how this competence appeared to others. For example, if an individual was about to engage with an easy task, then accomplishment of the task could only prove that the person was not incompetent – accomplishment of the task did not offer the opportunity to prove competence (or high ability). Similarly, very difficult tasks also offered little opportunity to exhibit high ability because high ability was unlikely to be demonstrated. If individuals want to prove high ability, their best option was to choose tasks of intermediate difficulty. (Nicholls' version of achievement-goal theory actually includes another component, namely, perceived competence. The relationship of this component to task difficulty and task orientation will be discussed more fully in a later section of this thesis).

So although Nicholls arrives at the same conclusion as McClelland et al, the route is somewhat different. For Nicholls, it is the task and probability of success that determines the orientation individuals will adopt, whilst for McClelland et al, it is individuals' pre-existing levels of n *ach* combined with the incentive value and probability of success that determine the direction and intensity of behaviour.

In partial agreement with McClelland et al, Dweck and Leggett (1988; see also Dweck, 1986; 1996) also believe that task orientation is better understood as a personality variable. According to Dweck and Leggett (1988), the way in which the attribute of intelligence is perceived by an individual can predict their approach behaviour towards and subsequent responses to task situations. They suggest individuals may hold two types of theories of intelligence: *entity* and *incremental* theories. According to Dweck (1986; 1996), entity theorists are those who believe that intelligence is fixed, and such a belief *".. orients an individual towards the goal of measuring, judging, or evaluating the trait"* (p. 69). This is because in trying to make sense of themselves and others, individuals with an entity theory of intelligence will tend towards judgement and evaluation not only of their own intelligence, but that of others. In contrast, incremental theorists believe that intelligence is a dynamic and malleable attribute and such a belief orients individuals towards trying to master or understand a task. This is because individuals with an incremental theory of intelligence are not concerned with evaluating their performance, but with trying to improve their performance. Dweck and Leggett (1988) further include perceived ability as an important component. They suggest that entity theorists who perceive their ability to be high will seek out challenges and exhibit high persistence with tasks. This is because they are confident of solving the problems they are set. However, entity theorists who have low perceived ability are likely to exhibit what Diener and Dweck (1978) refer to as a "helpless" response to tasks, that is, in the face of failure, they will give up with a task. On the other hand, incremental theorists regardless of their perceived ability will persist with tasks because their goal is to improve or master the task.

Elliott and Dweck (1988) tested these predictions. In their study, fifth-grade children were given either a task where the instructions emphasised that the experimenter was interested in their performance or a task where the emphasis was on learning. Prior to completing the experimental task, participants were given a short task and, based on their performance in this task, were then (bogusly) advised that they had either a good chance of solving the next task (high perceived ability conditions) or a low chance (low perceived ability conditions). In line with their hypotheses, in the conditions where both performance and high ability was emphasised, participants exhibited high persistence at the task and made positive comments as they went about the task. A similar pattern of behaviour occurred in the condition where learning was emphasised regardless of perceived ability. However, in the condition where performance was emphasised but participants were told they had low ability, participants evidenced low task persistence and gave negative evaluations of their performance and the task. Also included within their set of dependent measures were participants' problem-solving effectiveness. Interestingly, Elliott and Dweck report that although participants in the performance condition with high perceived ability responded in a mastery-orientated fashion to obstacles, that is, they exhibited persistence in the face of failure, unlike participants in the learning condition, they passed up the opportunities to increase their skills. The results from this study lead Elliott and Dweck to suggest that different goals *"run off a different "program" with different commands, decision rules, and inference rules, and hence with, with different cognitive, affective, and behavioural consequences"* (p. 11). In short, individuals who approach tasks with an incremental theory of intelligence are likely to engage with and respond to tasks in a positive manner, relishing the challenge the task offers and persisting in the face of failure. As Dweck and Leggett (1988) neatly summarise their position, individuals with an entity theory attempt to *prove* their ability whereas individuals with an incremental theory attempt to *improve* their ability. It is these personality characteristics that determine the goal an individual will adopt towards a task. At the risk of oversimplifying their argument, it may be useful to see their position as one that suggests that it is not the task that creates the goal, it is the individual who brings the goal to the task.

1.4 Weiner's (1985) Attribution theory

Two theories that also contribute to the analysis of goal-adoption are Weiner's (1985) Attribution theory of achievement motivation and Bandura's (1986) Self-efficacy theory. As Bandura's analysis alludes to Weiner's analysis, Weiner's analysis will be dealt with first.

The attributional approach to psychology gained its impetus from the work of Heider (1958). In Heider's classic experiment, participants were shown drawings in which geometric shapes (a large triangle, a small triangle and a small circle) were displayed in various positions. These drawings were placed in a sequence and run as a film so that the shapes appeared to be moving (Heider and Simmel, 1944). Participants were asked to "write down what happens in the picture". Of the twenty participants, only one participant described events in terms of geometrical shapes moving in space, the other nineteen assigned characters (e.g. jealous boyfriend) and meaningful actions to the movement of the shapes (e.g. moving aggressively towards; chasing; following). Heider suggested that were trying to fill in the gaps in their knowledge by assigning reasons for apparent behaviour, that is, individuals were trying to 'attribute reasons for why the shapes moved as they did. Kelley (1967) suggested that humans are motivated to *"attain a cognitive mastery of the causal structure of the environment"* (p. 193).

According to Weiner (1985), the perceived causality of an event can be mapped on three main dimensions: locus of causality, stability of the object of causality and the controllability of the object of causality. Locus of causality refers to whether the perceived cause of an event is attributed to an internal cause (e.g., it was my personal fault) or an external cause (e.g., something else made me do it). Stability of an event refers to how changeable that event is over time. For example, the outcome of a coin toss is unstable because it could go either way, heads or tails. The outcome of a coin toss on a one-sided coin would be stable because no matter what you did, the outcome would be the same. To explain his view, Weiner (1971) gave the example of four possible attributions that could be made to explain performance (either of oneself or others) on a task. The four possible attributions for an individual's performance were ability, effort, task difficulty or luck. Framing these attributions within a 2 (internal-external) x 2 (stable-unstable) matrix, Weiner suggested that ability could be seen as internal-stable - that is, it is a product of the individual and is relatively unchanging. Effort, because it is a product of the individual but can be changed - e.g., one could try harder - was classified as internal-unstable. Task difficulty was classified as external-stable because one cannot do anything about a task that is too difficult, and the difficulty of the task does not change. Luck involves an external-unstable attribution because one cannot do anything about luck and it is just as likely to work in one's favour as against.

The important point to note about Weiner's analysis is that all of the above forms of attribution are possible for the same event. If an individual failed at a task, they could make any one of the four types of attributions. Considering that the range of attributions includes luck and ability, then it becomes easy to see why the types of attribution should have such an important effect on the types of responses individuals have to events such as success and failure. Attributing failure to one's ability is likely to have a very different effect on one's perceptions of oneself than attributing that failure to luck.

Weiner (1979) added a third dimension to his analysis. He realised that ability could be perceived as unstable if learning were possible, or that effort could be seen as stable if one perceived oneself or others as having an industrious or lazy character (e.g., there is not much I can do about the amount of effort I put into the task, it is just part of my character). So although effort was presented in Weiner's original model to be a factor that was entirely volitional, Weiner (1979) suggested that this rather depended on whether something like effort was perceived to be controllable or not. The third dimension that was introduced to the model was therefore *controllability*.

In addition, Weiner also included the dimension of globality. Globality referred to the degree to which a trait was a general or specific one. For example, an individual may be poor at statistics whilst showing considerable ability in other academic domains. Thus, if ability is perceived to be specific in nature, one could attribute the cause of failure at an exam to a simply being poor at that particular subject whilst still maintaining a positive overall attitude towards academic tasks. However, if one saw ability as general, then failure in one domain may also be perceived to indicative of poor ability in all related domains.

1.5 Bandura's (1986) theory of self-efficacy

Moving onto Bandura's theory, he suggested that there are three forms of cognitive motivators: causal attributions, outcome expectancies and cognized goals. These concepts are represented in the theories already discussed in this thesis, namely, attribution theory (Weiner, 1985), expectancy-value theory (e.g., Atkinson, 1953; 1957) and achievement-goal theory (e.g. McClelland et al, 1953; Nicholls, 1984; Dweck and Leggett, 1988). Bandura (1990, 1997) contends that all the above theories can be explained in terms of perceived self-efficacy, which he suggests "*.. refers to beliefs in one's capabilities to organise and execute the courses of action required to produce given attainments*" (Bandura, 1997, p. 3). For Bandura, perceived self-efficacy (or the feeling of self-efficacy) was the appropriate unit of analysis when attempting to understand and predict human behaviour. For example, an attribution analysis suggests that behaviour depends on the causal ascriptions made about, for example, performance at a task. However, Bandura (1990) argues that attributions for performance can be seen as excuses rather than as distinct motivators for subsequent actions. That is, the attribution that one makes may not necessarily be the reason for one's subsequent behaviour. Covington and Omelich (1979) have provided evidence that causal attributions do not change performance but simply function as a "self-serving excuse", that is, the attribution just makes the individual feel better about themselves. Moreover, Bandura (1990) cites several studies that have shown that when attributional causation is arbitrarily varied, it is changes in self-efficacy that best determine subsequent engagement with a task (e.g. Schunk and Gunn, 1986; Schunk and Rice, 1986). Bandura (1990) also suggested that the spectrum of types of attributions that individuals make about their performance (e.g. effort, ability, task difficulty, chance) is far too restrictive. He suggests that individuals incorporate a much wider set of factors into their attributional appraisals such as whether the task was performed under favourable or unfavourable conditions or the amount of extra help the person was given. Thus, when individuals make effort attributions, these are likely to vary depending on the efficacy-relevant information that is available. Moreover, several studies have shown that regardless of whether effort attributions correlate positively or negatively with perceived self-efficacy, the best predictor of performance is the strength of self-efficacy belief. Bandura (1990) summarises the position thus: *"The overall evidence reveals that causal attributions, whether in the form of ability, effort, or task difficulty, generally have weak or no independent effect on achievement motivation"* (p.73).

With respect to goal-theory, Bandura suggested that it is not the goals that mediate motivational effects but rather the fact that individuals respond evaluatively to their own behaviour that is, simply adopting a goal has no lasting motivational impact, it is how their perceptions of efficacy changes over time that determines how they will respond in task situations. Thus, self-efficacy determines the challenges that individuals take, how much effort they employ and how long they persist because it is feelings of self-efficacy that determine people's beliefs that they can attain the goals they set for themselves.

At the risk of over-simplifying Bandura's detailed theory, his position is that self-efficacy lies at the root of behaviour in task situations. Understand an individual's self-efficacy beliefs and you can predict their goals, their expectations, their choice behaviour, how long they will persist with a task and how much they will enjoy that task. However, at this point in the thesis, the purpose is merely to introduce the various positions (e.g., attribution theory; expectancy-value theory; achievement-goal theory; self-efficacy theory) and how they relate to one another. A closer analysis of how each theory may help to understand and interpret the behaviour observed in chapters 3 and 4 will be discussed at various points throughout this thesis.

Having outlined the theories that form the crux of the analysis in this thesis, the next section in this chapter outlines the rationale for the experiments in chapter 3. These experiments concern how individuals respond to success and failure when tasks are either presented with a mastery or a performance focus.

1.6 Task orientation and responses to success and failure

Earlier in this chapter, evidence was cited that suggested that task orientation determined individuals' experiences of tasks, that is, mastery-type orientations resulted in positive experiences and performance-type orientations resulted in negative experiences (e.g., Harackiewicz, 1979; Ryan, 1982; Harackiewicz, Manderlink and Sansone, 1984; Plant and Ryan, 1985). It has also been shown that success generally results in more positive experiences of tasks relative to failure (e.g., Deci, 1971; Blank, Reis and Jackson, 1984; Vallerand and Reid, 1984). However, a less examined phenomenon has been the potential interactive relationship between task orientation and performance outcome (i.e., success and failure). For example, when a task is presented with a mastery or performance focus, how long will individuals persist with tasks after doing well compared to when they do poorly? What types of experiences will they report?

In a series of experiments, Ryan, Koestner and Deci (1991) investigated such interactive effects by crossing task orientation (performance-focus vs. mastery-focus) with performance feedback (positive feedback vs. no feedback). The performance-focus orientation was induced by telling participants that the task was a test of creative intelligence. In the mastery-focus conditions, participants were simply introduced to the task but not told it was a test of creative intelligence. In experiment 1, Ryan et al examined the length of time participants who did well at a task would persist with a task when it was presented to them with either a mastery or performance-focus. Using a between-subjects design, participants were given three cartoon drawings and asked to find as many instances as they could of a word (NINA) which was embedded in these drawings. Participants were given two minutes to complete each puzzle after which they were allowed to compare the number of NINAs they found against a sheet that reported the average number of NINAs found by similar age students. These averages were bogusly low to ensure that all participants recognised that they had done well at the task (i.e., the positive feedback manipulation). They were then told that the experiment was over and that they were to stay in the room whilst the experimenter visited another participant. In the meantime, the participant was free to try a few more tasks if they wished. During this time, participants were secretly filmed and the amount of time they spent on additional NINAs was taken as their level of (intrinsic) motivation. Intrinsic motivation theorists refer to this measure as "free-choice persistence" because during this time, participants are ostensibly under no external pressure to perform the task (e.g., Deci, 1975). When the experimenter returned to the room, participants were given a questionnaire and asked to rate their levels of interest and enjoyment of the task. Results showed that participants in the mastery condition persisted significantly longer during the free-choice period relative to those in the performance-focus condition. These results suggested that when participants were given positive feedback, they persisted for a greater amount of time with the task when it was presented to them with a mastery-focus. Experiment 2 investigated the effects of not giving participants any feedback, again, under mastery and performance conditions. The results revealed that it was now the participants in performance-focus group who persisted longer with the task ($p < .10$). This finding from experiment 2 suggested that when the task was presented with a performance-focus, receiving no feedback and positive feedback had opposite effects on the length of time participants persisted with the task.

To confirm the suggestive results of experiments 1 and 2, in experiment 3, Ryan, Koestner and Deci fully crossed task orientation (performance-focus vs mastery-focus) with performance feedback (positive feedback vs no feedback). Using a similar procedure to experiments 1 and 2, but this time using a different task and administering feedback verbally (e.g., positive feedback presented as " .. it looks like you've done better than most subjects I've seen so far"), the results from the fully crossed study broadly reflected the findings from experiments 1 and 2. Table 1 below presents a summary of the means for task persistence; they indicate that the different feedback and orientations had different effects on task persistence. Participants persisted with the task for longer after no feedback (i.e. no success), but this pattern of persistence only occurred when the task was presented with a *performance-focus.* When the task was presented in a *mastery* fashion, the persistence patterns were in the opposite direction.

Table 1: Mean number of seconds of free-choice persistence (range = 0 – 360; s.d. in italics and parentheses) in Ryan, Koestner and Deci (1991), Experiment 3.

	No feedback (No success)	**Positive feedback (Success)**
Mastery-focus	128 *(135)*	201 *(151)*
Performance-focus	236 *(111)*	167 *(136)*

[Taken from Ryan, Koestner and Deci, 1991, p. 199]

On the face of it, it would appear that for the dependent measure of free-choice persistence, the relationship between task orientation and task outcomes (i.e., positive feedback vs. no feedback) is well established. However, a closer examination of the data from the Ryan et al study reveals a less conclusive picture. For example, in Experiment 3, although the means reported in Table 1 above suggest that the results were significant, none of the analyses were significant at the $p < .05$ level. The only significant result reported for persistence was a "marginally significant interaction" ($p < .10$); for all other results, $p > .20$. Ryan et al suggested that the lack of significant results may have been due to the relatively small cell sizes and that the large differences in means suggested that with a larger sample, the results would have been more conclusive. This suggestion is speculative because it is not entirely clear whether or not it is reasonable for Ryan et al to interpret their persistence data in terms of a Type I error.

In particular, the position with regards to the amount of time participants will persist with a task after success, compared to failure under mastery-focus conditions seems to be equivocal. To make the argument clearer, it is probably useful to outline the logic given why after failure, participants persisted for so long in performance-focus conditions. According to Ryan (1982; see also Deci and Ryan, 1985; 1986; Ryan et al, 1991), individuals are constantly trying to prove competence to themselves and in doing so, they put themselves under an *intrapsychic* pressure to perform well. Thus, when participants fail at a task, they still want to prove to themselves that they are competent at that task. However, because the task is now over, there is no immediate opportunity for them to resolve their feelings of (in)competence. The free-choice period therefore becomes an arena where individuals can try to prove to themselves that they could solve the task, leading them to persist with the task during this period. According to Ryan (1982), even though the free-choice period is ostensibly free of external pressure, the individual still strives to surpass an internal standard of excellence they set themselves whilst they were completing (and failing) at the experimental task.

So according to Ryan, the performance-focus manipulation induces a concern in participants to do well and that after performing poorly at the task, this concern transfers to the free-choice period resulting in participants persisting at the task during this period in an effort to recover lost self-esteem. However, and importantly for the experiments in this thesis, this hypothesis leaves it unclear with respect to the amount of time individuals should persist when they do poorly at tasks presented with a mastery-focus. For example, Dweck and Leggett (1988) suggest that a mastery-orientation should lead to "*... the maintenance of effective striving under failure*" and "*... the generation of effective strategies in the face of obstacles*". In an earlier study by Diener and Dweck

(1978) where children were asked to verbalise their experiences of a task as they progressed through that task, the authors reported that mastery-orientated children appeared to maintain an *"unflagging optimism"* in the face of failure. For example, mastery-orientated children made positive verbalisations about the task such as "I did it before, I can do it again" or "I'm sure I have it now". In an analysis of what they labelled as adaptive and maladaptive responses to failure, Dweck and Leggett (1988) suggested that when tasks were approached with a mastery-orientation, individuals were able to adopt an adaptive response to failure by persisting, and eventually, resolving the task. Those individuals who adopted a maladaptive (or helpless) approach to the task sought to desist from the task at the earliest opportunity. Elliott and Dweck (1988) suggest that individuals who approach tasks with a mastery-orientation focus their attention on improving their ability as opposed to proving it. The case seems to be that theoretically, approaching a task with the intention of trying to master it should engender a motivation to persist, even in the face of failure.

A counter argument is provided by intrinsic motivation theorists who claim that feelings of competence are important predictors of task persistence. For example, Deci's (1975) cognitive evaluation theory (proposition II) states that *"If a person's feelings of competence ... are enhanced, his intrinsic motivation will increase"*. By this, Deci is suggesting that when individuals feel competent at a task, they should continue to persist with it. Thus, in terms of making predictions as to whether participants should persist longer after doing poorly compared to when they do well, it would seem that participants should persist longer after doing well because they will their feelings of competence should be higher. This position is well supported (e.g. Deci, 1972; Blank, Reis and Jackson, 1984; Vallerand and Reid, 1984). However, a close examination of the supporting evidence leaves it unclear whether or not some of the tasks in these experiments were presented with a mastery-focus. For example, in the Blanck, Reis and Jackson (1984) study, participants were set a target performance criteria to reach and were additionally told they would be observed from behind a one-way mirror. In the Vallerand and Reid (1984) study, the task was a stabilometer motor-task where participants were told a target performance level they were expected to try to attain (e.g. to balance on the stabilometer for 20 seconds over a certain number of trials). Indeed, in studies where different feedback levels have been the independent variable, the orientation has (probably unwittingly) been performance-focused (e.g. Feather, 1959,1963, 1967). When task orientation has been controlled for, as it was in the Ryan et al study (see also Anderson and Rodin, 1989 later), the results have been less conclusive.

Indeed, the theoretical point that needs to be emphasised at this point is whether or not performing poorly at a task that has a mastery-focus should undermine feelings of competence at all. If the suggestion is that mastery orientations cause individuals to focus on mastering the task, or, as Nicholls (1984) more neatly summarises *"mastery tends to be experienced as a means to an end rather than an end itself"*, then poor performance at a task should merely be diagnostic that the task is either a very difficult one, or one that requires more effort, or one that is just impossible. If individuals are not putting themselves under pressure to do well, then poor performance should just serve as information for evaluation of the task. For example, if an individual was given a task and asked to comment on what they thought were the key characteristics of it, then to conclude that the task was a difficult one does not seem to imply anything negative about that persons ability to perform it. After all, their reason for engaging with the task in the first place was not to perform well, but to provide information to someone else about the task. Looked at in this way, Deci's suggestion that competence plays a role in determining positive or negative experiences of tasks may still be correct, but if the goal of the individual is simply to provide information to another person (the experimenter), competence may be defined in terms of how well they are able to perform this task, not the actual experimental task.

A further examination of the data for the Ryan et al study reveals some additional interesting behaviour in the free-choice period. For example, although there appears to be a large difference between the means for the no feedback and positive feedback groups in the mastery-focus conditions, the standard deviations were large. For example, in the mastery-focus-no-feedback condition, the mean persistence time was 128 seconds but the standard deviation was 135 seconds. Clearly some participants in this condition persisted with the task for a considerable length of time during the free-choice period. Indeed, the size of the standard deviations suggests an "all or nothing" effect whereby, during the free choice period, participants either persisted with the task a lot or not at all. Recall that Ryan et al suggested that by increasing the sample size, this might results in clearer difference emerging between the two groups. Whilst this may be true, it is not clear why increasing the sample size should result in a population with a larger proportion of participants who do not persist with the task for a long time. If the effect is an all or nothing one, then the question is not the amount of time they persist, but whether they persist at all. Indeed, it may be that it does not matter whether an individual performs particularly well or poorly at a task that has a mastery-focus; maybe a mastery-focus negates the potentially negative effects of not performing well.

As well as persistence, participants' ratings of interest and enjoyment have also been taken to be indicative of their motivation to engage with a task. According to Deci (1975), free-choice persistence and self-report ratings of interest and enjoyment are thought to operationalise the concept of intrinsic motivation, that is, both the amount of time an individual persists at a task and their reported interest and enjoyment are thought to reflect their motivation to engage with the task. Following this preposition, Ryan et al examined the correlations between free-choice persistence and participants' reports of their interest and enjoyment. In Experiment 3, for the factor of task orientation (i.e., performance vs mastery-focus), although the correlations were not significant in themselves ($r = -.35$ and $r = .22$ respectively), when the co-efficients were compared, there was a significantly lower correspondence between these co-efficients for the performance-focused groups relative to the mastery-focused groups. In a similar study, Anderson and Rodin (1989) crossed task orientation (performance vs mastery) with feedback (positive vs. mildly negative). In their study, participants were asked to solve a series of short tasks (e.g., 2 sample multiple-choice brain teasers) and were given feedback on their performance. Participants in the positive feedback conditions were told that they had performed in the 95^{th} percentile of students whilst participants in the mildly negative feedback conditions were told that they had performed in the 55^{th} percentile. A control group was given no feedback at all. A free-choice period then followed after which participants were asked to complete a series of questionnaires that assessed their mood and self-efficacy beliefs. The results revealed that in the mastery-focus conditions, when mood and free-choice persistence times were standardised and the difference between the two scores calculated, there were no significant differences between the mild negative and positive feedback groups. Post-hoc comparisons revealed that the scores for the performance-focus negative feedback group were significantly higher relative to all the other groups. This analysis therefore suggests that in the performance-focus conditions, whilst participants may have been persisting with the task, they were not necessarily enjoying the experience.

However, it is both theoretically and empirically unclear whether or not the experiences of tasks undertaken with a mastery-focus would be subject to the same type of intra-psychic pressure outlined by Ryan (1982). If the suggestions of Dweck and Leggett (1988) and Nicholls (1984) are correct, then there may be a case that participants should enjoy tasks just as well regardless of whether they perform well or not. In both the Ryan et al and the Anderson and Rodin studies, when participants received either no feedback or mild negative feedback (i.e., whenever feedback was NOT positive) in the mastery-focus conditions, there were either a) no significant differences in reported interest and enjoyment or b) no significant relationships between free-choice persistence and interest and enjoyment ratings. This suggests that participants in the no feedback/mild negative feedback conditions were persisting and enjoying the task to same degree as those who received positive feedback.

So, whilst in the Ryan et al and Anderson and Rodin studies, non-positive feedback under performance-focus conditions resulted in a difference in the amount of time participants persisted with the task, and a disassociation between free-choice persistence and interest and enjoyment ratings, this effect was not observed under the mastery-focus conditions. To re-iterate, it may be that mastery-focus conditions help to overcome the potentially negative effects of performing poorly. The evidence therefore seems to suggest that individuals may not need positive feedback to enjoy a task when a task is presented with a mastery-focus and that task outcomes may only play a small part in determining individuals' experiences.

Thus, while there seems to be plethora of evidence examining the differential effects of task orientation on task persistence and interest/enjoyment, apart from the Ryan et al and the Anderson and Rodin studies, task outcomes have rarely been examined in any systematic way in terms of a mastery-focus. When they have been - as they were in the Anderson and Rodin and Ryan et al studies – the results have been inconclusive. Even Feather's extensive research in the 60's that examined the effects of success and failure on measures such as task persistence and task choice were conducted with what Feather himself defines as "evaluative settings" (see Feather, 1961; 1963a, 1963b, 1963c, 1965, 1967, 1968, 1969; also see Feather 1982 for review on persistence). In a study where task orientation was manipulated, this study did not also manipulate task outcomes (Feather, 1959). Thus, (somewhat surprisingly), there appears to be little empirical basis for predicting whether success or failure while working on a task presented with a mastery-focus will lead to greater persistence. Indeed, if Dweck and Leggett are right in their suggestion that failure should not be an obstacle to continuing persistence, the theoretical basis to suppose that success and failure should lead to different levels of persistence is also unclear. Additionally, it is also theoretically and empirically unclear whether or not success and failure should lead to different task experiences. It is the issues of task persistence and task experiences under mastery-focus conditions that the experiments in chapter 3 aim to explore. In these experiments, participants were presented with a task in mastery-type way and they either did well (success condition) or poorly (failure condition). In a similar fashion to the Ryan et al (1991) and Anderson and Rodin (1989) studies, task persistence and interest and enjoyment were used as dependent measures.

1.7 Outline and rationale for experiments in chapter 3

The research question addressed in chapters 2 and 3 was how individuals would respond to differing levels of success and failure when a task was presented to them with a mastery-focus. The responses measured (i.e. the dependent measures) were persistence at the task and self-reported post-task enjoyment. Two issues were of particular concern. Firstly, it was unclear how long individuals who did poorly at a task would persist at a task relative to the group who performed relatively well. If Dweck and Leggett (1988) are correct in their assertion that failure in a task presented with a mastery-focus should present no obstacle to further persistence, then persistence at a task should be at least equal to that of a group who perform relatively well at the task. Secondly, it was unclear whether or not individuals who performed poorly would enjoy the task as much as those who did well. If mastery orientations cause individuals to become involved in the task, then performing poorly at a task with a mastery-focus might result in individuals experiencing the task as positively as those who do well. After all, if a mastery-orientation causes individuals to focus on the process rather than the outcome, then the outcome should not be predictive of their experiences.

A feature of the experiments in chapter 3 was the use of a strangely under-used method of measuring persistence. Instead of the free-choice paradigm most frequently adopted by researchers in the field of intrinsic motivation (e.g., Deci, 1975), participants in experiments 1-5 were given a task and allowed to persist with it for as long as they wished. In theory, this method of persistence should be no different to the free-choice persistence paradigm, in that participants are free to desist with the task anytime they want. Additionally, there are benefits to this type of persistence measure because instead of having to interpret when participants are or are not actually engaging with the task and use this as the measure of persistence (often requiring validation using inter-rater reliability co-efficients), persistence can simply be taken from the time the participant enters the room to the time they leave. This method has been employed in similar forms by several researchers. For example, Feather's (1962) review of persistence offers two different definitions of persistence. Feather was interested in how long individuals would stay with a task before moving onto a different task. This type of persistence he referred to as *"temporal"* persistence. The second paradigm employed by Feather (1959, 1961) was a sub-case of temporal persistence where individuals were given extremely difficult or insoluble tasks at which they continually failed. Feather believed this type of persistence scenario to be indicative of a particular aspect of individuals' behaviour, namely, their persistence in spite of information that suggested they were unlikely to be able to solve the task. Feather preferred to refer to this second type of persistence as *"resistance to extinction"*. The type of persistence in chapter 3 seems to more closely resemble Feather's first type of persistence, namely, *temporal* persistence.

One of the reasons for using this type of persistence measure was that it seemed to be analogous to the situation that most students are faced with in their everyday schooling. For example, students are often given statistics problems and asked either to solve the problems in terms of an equation they have just been given (mastery-focus) or are told their answers will be graded (performance-focus). It seems that it would therefore be useful to investigate how individuals respond to success and failure when they are instructed to engage with a task, but are given a free reign to desist whenever they want.

1.8 Ecological validity

A recurring problem for researchers in social psychology experiments is the degree to which phenomena observed in laboratory settings transfer to non-experimental settings. In this thesis, all the experiments are laboratory-based. Therefore, it seems useful to outline some of the objections that have been raised against the validity of phenomena observed in experimental social psychology in order that these issues can be addressed in chapters two and three.

Several authors suggest that the experimental situation itself is a highly relevant and underestimated confound. For example, Tedeschi and Reis (1981) suggest that rather than motives or habits, it is social contexts that serve as the causes for behaviour. Individuals may be motivated to behave in a particular way before they arrive at the experiment, but it is the context of the experiment that will determine their eventual behaviour. According to Page (1981), experimental situations are simply examples of symbolic interaction whereby participants try and make sense of their environment and behave in relation to their perception of this environment (see also Alexander and Rudd, 1981). Tedeschi and Reis (1981) further suggest that individuals are constantly trying to manage the impressions they give to others about themselves, and that experimental situations create a context where, compared to tasks done in non-experimental situations, tasks are likely to be experienced in crucially different ways. For example, an individual completing a crossword at home on their own is likely to experience the task differently compared to if they had to complete the same crossword in an experimental situation. In a similar fashion to how Nicholls (1984) suggests that individuals in achievement settings are constantly trying to prove competence (or avoid proving incompetence), Tedeschi and Reis suggest that all behaviour is either consciously or sub-consciously an attempt to manage impressions. On the face of it, these two positions seem remarkably similar but whereas Nicholls' analysis stops at a global disposition to prove competence, Tedeschi and Reis make a more specific point in that they question whether such dispositions are relevant when analysing behaviour in experimental settings. For example, Page (1981) suggests that participants act unnaturally in laboratory experiments because they either consciously or subconsciously are aware of the experimental hypothesis or are working towards a hypothesis of their own. So although individuals may indeed be trying to prove competence, this may in part be tied to a goal of "being a good subject" in the experiment. These two goals may, or may not be orthogonal but whatever the case, according to Page (1981) and Tedeschi and Reiss (1981), the experiences of the participant not only in the experimental situation, but *because of* the experimental situation, must not be underplayed (see also Alexander and Rudd, 1981; Orne, 1969).

The second criticism of experimental psychology relates to the demand characteristics inherent in an experiment, that is, the specific situational cues that indicate to the participant how they should behave. Page (1981, pp. 63) refers to this phenomena as "demand awareness" or "hypothesis awareness". According to Page, participants enter into an experiment with some sort of implied contract between themselves and the experimenter. The participant wants the experimenter to be pleased with their performance and will thus try to guess what the experimental hypotheses are, and, as far as possible, try and behave in way that supports these hypotheses. Experimenters, on the other hand, have a vested interest in the experimental outcome and therefore may, either consciously or subconsciously, unwittingly give cues to the participant as to how they should behave. Demand awareness may therefore result from a combination of the participant wanting to understand the purpose of the study and the experimenter wanting to tell the participant how to behave. Page (1981, pp. 64-74) details a series of studies that have investigated the effects of demand characteristics where post-experimental enquiry has revealed participants to be highly aware of the purposes of the experiments (e.g. Page, 1968), and how that awareness has allowed a different interpretation to be made with regards to the phenomena being observed. For example, Schafer and Murphy (1943) reported that the perception of an ambiguous figure-ground picture could be conditioned by rewarding participants for making specific types of responses. In subsequent non-rewarded sessions, the reward contingencies in the reward phase of the experiment predicted the way participants would view similar ambiguous figure-ground stimuli. Page was concerned that some stimuli were clearly not ambiguous, that is, it was unreasonable to claim that these stimuli could be seen in more than one way. However, after being rewarded for perceiving the (purportedly) ambiguous figure in one way, participants were prepared to claim that they saw other (purportedly) ambiguous figures in that way. Page (1968) replicated the Schafer and Murphy study but added a post-experimental enquiry session where participants were asked to say whether they had guessed the experimental hypotheses and whether they had the intention of trying to conform with these hypotheses. Page found the effect produced by Schafer and Murphy only occurred when participants were both aware of the hypothesis and reported a willingness to co-operate with the experimenter. Thus, although the phenomena observed by Schafer and Murphy may have been a real one, the post-experimental enquiry data in Page's study suggested that researchers should be wary about claiming effects that could be explained by the cues inherent in the experimental design.

Two further factors relevant to the experiments in this thesis are the issues of deception and participant sophistication. Page (1981) suggests that because experimenters are aware that participants are constantly trying to make sense of the experiment they are in, they construct elaborate cover stories in an attempt to deceive participants as to the true purpose of the experiment. However, such deception does not adequately take into account the fact that the participants are also aware that some deception may be taking place and have incorporated this knowledge into their analysis of what the experiment is about. For example, Page reports that during his pilot studies for the previously mentioned replication of the Schafer and Murphy experiment, the hypothesis of the experiment was identified most frequently by those participants who were taken from a pool of psychology undergraduates. Thus, even when deception is used, Page suggests that the participant population strongly determines whether or not the deception will be successful.

In this thesis, the true nature of the experiments was deliberately concealed from the participants. This technique is in keeping with methodologies employed in similar research areas. For example, in experiments where task orientation has been manipulated, the participants are either put under pressure to well by being told the task is a measure of creative intelligence (performance-focus), or are told nothing (e.g. Harackiewicz, 1979). The belief or otherwise of this information is crucial to the interpretation of the subsequent behaviour observed by participants. Similarly, in the free-choice paradigm methodology outlined by Deci (1975), participants are told that the experiment is over but that the experimenter has to leave the testing area to retrieve some questionnaires that the participant needs to complete. The participant is asked to wait in a room and whilst the experimenter is away, the participant's behaviour is secretly filmed. Again, the degree to which participants believe that they are in a free-choice situation will determine the length of time they persist with the task. The issues of deception and participant sophistication are particularly relevant to this thesis because in all the experiments detailed in this thesis, deception has been used and participants have been taken from a psychology undergraduate population.

Page (1981) suggests that social psychologists, as a matter of course, should employ post-experimental interviews as a part of standard experimental methodology. However, whilst this position is laudable, Page does seem to underplay the possibility that the post-experimental enquiry may in itself be fraught with demand characteristics. For example, asking participants about the true nature of an experiment is likely to create a suspicion in the participant that something was indeed suspect about the experiment (when in fact they were not suspicious to begin with). In this instance, it is not unreasonable to suspect that participants will start to look for plausible reasons for what the experiment was about and report a hypothesis that they did not believe and had not affected their behaviour. Whilst Page is prepared to claim that experimenters are subconsciously affected by their desire to achieve significant findings in their experiments, the same criticism could also be levelled at the methodologies used in a post-experimental enquiry session. It seems that if one was to probe and prompt a participant into responding after they have completed an experimental task, this may result in the same type of demand characteristics that the original experiment is being criticised for.

Additionally, by asking participants whether or not they had guessed the purpose of the experiment, there is no way of checking whether this understanding or realisation is post-hoc. Clearly, participants could realise the true nature of the experiment after they have completed it, but does this necessarily mean they were not affected by the experimental manipulations during the experiment? Also, even if participants were aware of the experimental manipulations, does this necessarily mean that the manipulation did not affect their behaviour? This seems like the argument that people routinely give about advertising in that they believe they are aware of the manipulation but are not affected by it.

At a broader level, Page seems to be suggesting that verbal reports may be reliable explanations of the causes of behaviour. This position needs to be placed in context with the counter claim, and considerable research, suggesting that verbal reports do not reliably correspond with actual behaviour. For example, in a study by Nisbett and Bellows (1977), participants were given resumes of potential job candidates together with their performance in a job interview. Participants were then asked to rate the candidates on a series of characteristics (e.g., sympathetic; intelligent). Before being given the resumes, participants were asked the degree to which certain factors (e.g.; whether or not the candidate spilled a cup of coffee during the interview; whether or not the person was attractive) would affect their judgements of the candidates. These factors were subtly incorporated into the resumes and job interview descriptions. Nisbett and Bellows found that participants' a priori beliefs about how certain factors would affect their ratings of candidates did not always match with their actual ratings. For example, participants who reported that the attractiveness of a candidate would not affect their intelligence ratings, when given the information that the candidate was attractive, consistently reported them to be of low intelligence. Bargh and Chartrand (1999) have recently reviewed the evidence on the usefulness of verbal reports as explanations of behaviour and found compelling evidence that such reports are poor predictors of actual behaviour (see also, Gollwitzer, 1999; Kirsch, 1999; Nisbett and Wilson, 1977).

Page and co-workers' suggestions can be viewed as the sceptical viewpoint. Indeed, whilst the experimental procedures adopted in this thesis have been derived from methodologies employed in the extant literature, this does not mean that these methodologies are not open to criticism. However, it is felt that Page's concerns may somewhat overplay the value of adopting a sceptical approach to the analysis of behaviour in task situations. In an attempt to present a balanced view, issues regarding the ecological validity for the experimental findings will be discussed both throughout and at in the final conclusions section of this thesis. In terms of the post-experimental enquiries demanded by Page (1981), for experiments 1 – 5 in this thesis, no structured post-experimental analyses were completed, though post-experimental comments during the debriefing sessions were noted (and in some cases acted upon in terms of amending future experiments). For experiments 6-10, participants were asked a series of questions in an unstructured fashion during debriefing sessions, and notes on the answers to these key questions were kept by the experimenter.

Thus whilst it is recognised that experimental situations can create artificial arenas for investigating social psychological phenomena, this does not necessarily mean that the phenomena do not exist. Even if a demand characteristic existed, this does not mean that it necessarily caused the phenomena observed in the experiment. Whilst Page and co-workers are right to caution researchers against the possibility of making Type II errors when analysing their data, they need to also be aware that in the process of criticising their own experiments, they may be committing Type I ones.

2 Introduction to Experiments 1-5

Chapter 1 presented the argument that when a task was presented with a mastery-focus, it was unclear whether or not participants who did well at the task would persist or enjoy the task any differently compared to those who did poorly. Chapter 2 presents an introduction to five experiments in chapter 3 that investigated this hypothesis.

2.1 Choice of experimental task

The purpose of Experiments 1-5 was to investigate how long participants would persist with task when they felt that they had either performed well or performed poorly. Fundamental to this investigation was to ensure, as far as was possible, that when comparing the two groups, participants did indeed feel that they had performed well or poorly.

Researchers have adopted a variety of methodologies to convince participants they have either done well or poorly. For example, a common method of inducing an experience of "success" has been to tell participants that they have done well at a task. Intrinsic motivation theorists believe that feelings of competence help to increase intrinsic motivation (Deci, 1975; Deci and Ryan, 1985) and experimenters in this field have attempted to induce feelings of competence by giving participants positive feedback after completion of a task (e.g., Deci, 1971; 1972; Boggiano and Ruble, 1979; Ryan, Mims and Koestner, 1983; Harackiewicz, Manderlink and Sansone, 1984; Ryan, 1982). Presenting the feedback in these studies has been administered in a variety of ways. For

example, feedback can take the form of a) evaluative feedback e.g., "very good, excellent"; b) normative feedback, e.g., ".. that's the best one I have seen so far" or, ".. your score falls in the 80^{th} percentile", c) improvement focussed feedback e.g., "that is the best work you have done so far", d) experimenter approval, e.g., "I really like what you have done", and e) person praise, e.g., "you really are a very fine artist".
(see Koestner, Zuckerman and Koestner, 1987). Other researchers have attempted to induce feelings of success by telling participants they have surpassed a normative standard (e.g. Brunstein and Gollwitzer, 1996; Harris and Covington, 1992; Brown and Gallagher, 1992; Baumeister and Tice, 1985; McFarlin, 1985; Miller, 1985). Normative criteria are performance standards that are known to the population being studied and can be general (e.g., I. Q.) or specific (e.g., average I. Q. for a particular population such as undergraduates).

In the studies above, participants have been left in no doubt as to how well they performed. However, the methodology I wanted to employ was one where participants were not explicitly told whether or not they were performing well, that is, I wanted them to have to rely on their own perceptions of their performance, not on verbal feedback or an advertised normative standard. This was because the task was supposed to have a mastery-focus and I wanted to minimise the number of cues that might imply to participants that they were engaging in a performance task.

There are several examples where this type of technique has been employed. For example, Miller and Hom (1990) asked participants to solve a series of anagrams that had been pilot tested to ascertain how many were solved in a certain time. Participants were either given a selection of hard anagrams or easy anagrams, and Miller and Hom assumed that solving high and low percentages of these anagrams would induce feelings of success or failure (see also Brown and Dutton, 1995 for similar procedure; also Sandelands, Brockner and Glynn, 1988; Feather, 1965, 1968, 1969). This type of feedback has been defined as self-administered because rather than the experimenter telling the participant how well they have done, the participant is not given any cues as to how well they have done, they have to decide for themselves how well they are performing.

However, one of the drawbacks of this experimental technique is that the experimenter can never be entirely sure that the participant has experienced the desired level of success/failure. In the Miller and Hom (1990; see also Miller and Klein, 1989) pilot studies, participants were given 15 (purportedly difficult) anagrams and it was found that they solved approximately 41% (6.08). Unfortunately, Miller and Hom do not report how difficult participants actually rated these tasks either in the pilot or the full studies. Other researchers who have adopted a similar approach to manipulating task outcomes by presenting either easy or difficult tasks, have augmented performance outcomes by explicitly advising participants how well they have done (e.g., Brown and Gallagher,1992; Brown and Dutton, 1995; McFarlin and Blascovich, 1984) suggesting that these researchers were aware that allowing participants to simply experience success/failure, may not have been sufficient for them to experience the amounts of success the experimenter desired.

Since it was intended that participants were to receive feedback in a procedure similar to the one adopted by Miller and Hom (1990), several pilot studies were conducted, to help address whether or not the manipulations in the experiments in this chapter would indeed induce feelings of success or failure.

Additionally, it was also important that participants in each of the success and failure groups achieved similar success rates. For example, if participants in the success group varied in the percentage of success they experienced, then differences in ratings may have been explainable by their different rates of success. By ensuring that all participants who were to receive success feedback experienced virtually identical rates of success, then any differences in subsequent ratings within the group could be explained by factors other than the amount of success they received.

However, the proposed persistence paradigm for the experiments in this chapter meant that participants were allowed to persist with the task for as long as they wanted, and it was not clear in advance how long this would be. This meant that by having the restriction of not administering feedback verbally, it was going to be difficult to control absolute success rates, because participants would be persisting with the task for different lengths of time and thus attempting (and succeeding and failing) at different rates. Devising a task that would be appropriate for this experiment was therefore not straightforward. For example, in research where persistence has been used as a dependent measure, tasks such as spill and spell word games (e.g., Harackiewicz, Abrahams and Wageman, 1987; Epstein and Harackiewicz, 1992), anagrams (e.g., Feather, 1963b, 1965, 196, 1969; Baumeister and Tice, 1985; Sandelands, Brockner and Glynn, 1988; Eisenberger, Kuhlman and Cottrell, 1992), the soma cube (e.g. Deci, 1972) and the NINA task (e.g., Ryan, 1982; Ryan, Mims and Koestner, 1983; Ryan, Koestner and Deci, 1991; Koestner, Bieneri and Zuckerman, 1992) have been the preferred types of problems participants have been given to engage with. Unfortunately, these tasks become inappropriate when feedback and the time of persistence is self-administered because they do not allow the level of success to be controlled with any great precision. If participants persist for different periods, even if they are presented with easy tasks to ensure that they solve many problems, there is likely to be proportionately different success rates for different participants. Thus, differences in subsequent experiences could be confounded by the actual level of success of each of those participants.

Harackiewicz, Manderlink and Sansone (1984) devised an interesting solution to how bogus feedback could be administered. Participants were asked to play a pinball game and given a target points total to attain. Unbeknownst to participants, the points totals displayed on the machine as the participants played at the pinball game were manipulated by Harackiewicz so that participants either passed or failed the targets they had been set. Harackiewicz reported that none of the participants reported being suspicious or aware that the feedback they had been given was bogus. Thus, Harackiewicz was able to precisely control the level of feedback participants received without having to support the feedback with verbal confirmation. It was this type of approach that the experiments in this chapter attempted to replicate.

Developing a pinball game proved not to be possible within available resources and investigations into the possibility of manipulating feedback by re-programming the software for a currently popular computer games such as "Tetris" also proved unsuccessful. Existing in-house computer programs such as a maze task where participants were given a brief arial view of a maze and then asked to navigate the maze from ground level also proved to be unsuccessful. In this task, participants received (bogus) feedback as they progressed through the maze. However, this feedback could only be set according to the number of responses a participant made. This meant that participants could actually return to their starting point in the maze but still receive positive feedback. Pilot testing revealed that although participants reported enjoying the task, they tended not to believe the feedback they were receiving. The eventual solution was to use a task that, for the purposes of this thesis, has been called the "stockmarket" task.

The game was a computer-based task that involved participants having to make predictions whether the stockmarket was going to go up or down. The purpose of the experiment was for participants to try to recognise patterns of fluctuations, similar to the way one would do if they were to analyse the movements of stocks and shares in the stockmarket. On each trial, participants were asked to predict whether the stockmarket would go up or down; they were then told whether their prediction was right or wrong. Based on this information, participants were asked to try and predict what the next outcome would be. The implication to the participant was that there was an underlying pattern to recognise and that once this pattern had been recognised, the "market" could be predicted. However, and in a similar fashion to the Harackiewicz et al (1984) study, outcomes were manipulated so that one group of participants achieved rates of success that I wanted them to, namely, 70% and 30%.

The success levels were determined via a series of pilot studies. In these studies, participants were interviewed after completing various versions of the stockmarket task to determine how much they enjoyed the task and how well they thought they had done at it. These participants were mostly Psychology postgraduates at Stirling University, purposely selected as it was felt that they would be sensitive to the methodological issues. They were all blind to the purpose of the experiment. They were also encouraged to make any other comments about the experiment. A (qualitative) analysis of their replies suggested that at success levels lower than 30%, participants began to report that the feedback they were receiving was not entirely unbiased, with several participants (correctly) suggesting that it was the software program that was determining their success levels. The success level for the "Failure" group was consequently set at the 30% level for Experiment 1

For the Success group, I worked on the assumption that 50% success would be experienced as average success. Anderson and Rodin (1989) reported that participants who were told they had performed in the 55th percentile were "disappointed, but not devastated". Thus, 50% success appeared not only to be a statistically average success rate, but an experiential one aswell. Because the Failure group was to achieve success rates of 20% below the 50% average, the success level for the "success" group was set at 20% above (i.e., 70%). Additionally, it was noted that in the pilot studies, participants reported having performed well at the task at all levels above 70%, suggesting that this rate of objective success actually did result in participants feeling they had done well .

2.2 Dependent measures

The dependent measures were the length of time participants persisted with the task, the number of predictions they made, and how they rated their experiences. These measures were similar to those used by Ryan et al (1991) and Anderson and Rodin (1989). In these studies, after the free-choice persistence phase, participants were asked a series of questions. The first two questions concerned interest and enjoyment. These questions have been identified as an operational definition of intrinsic motivation - that is, when individuals report high levels of interest and enjoyment in a task, they are also likely to be motivated for that task (see Deci, 1975). Researchers have combined these questions with a variety of others such as effort, freedom, fun, competence, but interest and enjoyment seem to be the measures that have consistently been used in the achievement-motivation and intrinsic-motivation literature to determine participants' post-task experiences (e.g., Elliot and Church, 1997; Harackiewicz and Elliot, 1996, 1994; Ryan, Koestner and Deci, 1991; Koestner, Zuckerman and Koestner, 1987; Ryan, 1982; Harackiewicz, 1979).

The third question asked participants to report how difficult they thought the task was. It was hoped that this question would help to validate the independent variables, namely, whether or not the two conditions were actually perceived as success and failure conditions. For example, if participants in the success group reported the task significantly less difficult than participants in the failure group, then this could be seen as validations that the different success manipulations had at least caused participants to experience the task in different ways.

The fourth and fifth questions examined not only how well participants thought they had done at the task, but also how well they thought they had done relative to others. Throughout this chapter, these questions will be represented by the acronyms "WellY" (How well did you think you did at the task?) and "WellO"(How well do you think others would do at this task?). The acronym "WYWO" is the acronym for a within-subject comparison between WellY and WellO, and represents an analysis of how participants thought they had done relative to others. For example, if WellY was greater than WellO, then this would mean that participants thought that their performance would be better than others.

A specific feature of the WellY and WellO questions was the *type* of self-analysis individuals had to perform. Questions for interest and enjoyment required individuals to reflect on themselves and to examine their own experiences. However, asking participants how they thought others would do at the task required a different reflective process, one that involved a comparison with an imagined other. The overall idea was to see whether or not participants, after performing either well or badly, would vary in their perceptions of the quality of their performance relative to others. As noted earlier, achievement-goal theorists such as Dweck and Nicholls claim that individuals are concerned with how they match up to others so I felt it was important to have a measure that tapped something about how participants were comparing themselves to others. For example, after experiencing success, individuals may find a task more interesting/enjoyable, but how does this experience of success affect how individuals feel about themselves? Would participants who performed well at the task experience an increase in self-confidence? Would performing poorly lead to a loss of self-confidence? After all, if mastery conditions are those where participants seek to master a task, then failure to do so may result in the recognition that the task is a difficult one. Thus, participants who do poorly at the task may perceive it as difficult to solve, but because they are not concerned with performing well, their goal of mastering the task, or understanding the difficulty of the task, might be fulfilled with no loss to self-confidence. This suggestion can be examined by analysing the answers to the questions 3 (how difficult/easy did you think that task was?) and 4 (How well did you think you did at that task?). In the "failure" conditions, if participants really are only trying to understand or master the task, then if they report the task as difficult, they should not report themselves as having done poorly at a task. To phrase this in terms of an attribution analysis, it would be unreasonable to feel that one has performed poorly at a task that one also feels is very difficult.

Additionally, in a similar fashion to the Ryan, Koestner and Deci (1991) and Anderson and Rodin (1989) studies, the length of time participants persisted with the task was correlated with their interest and enjoyment ratings. Recall that Ryan (1982) suggested that when participants persisted with tasks but did not enjoy them, this was indicative that they were engaging in "ego-involved" persistence as opposed to persisting because they were intrinsically motivated. By examining the relationship between time spent on the task and enjoyment and interest – or mood and activity as Anderson and Rodin (1989) refer to it as – I thought this would be a useful additional measure to help determine whether participants were persisting because they were enjoying themselves.

2.3 Use of 6-point Likert scale

Traditionally, researchers have anchored their questionnaire rating scales on a 7-point Likert scale (e.g. Grolnick and Ryan, 1985; Plant and Ryan, 1985; Koestner, Zuckerman and Koestner, 1987; Ryan, Koestner and Deci, 1991; Elliot and Harackiewicz, 1994, 1996; Elliot and Church, 1997). Such scales allow participants the option of choosing a middle rating score of 4. In the experiments in this thesis, a 6-point rating scale was used instead. There were several reasons for this decision.

Firstly, a 6-point scale offers no opportunity for participants to select a middle score. When a decision is marginal, participants are forced into deciding whether their rating of a particular measure is more or less than average - e.g., was that task more than averagely interesting (rating 4), or less (rating 3). It was hoped that by not allowing participants the option to select an essentially nebulous rating, this would force them to reflect more carefully on how they actually felt about the question they were being asked.

The second justification for using a 6-point rather than a 7-point scale requires the use of a hypothetical scenario. Say 20 participants were tested in a between-subjects design and asked to rate their experiences of x on a 7-point scale. If all participants were undecided about x, then they may all rate x at 4. Subsequent analysis of differences between the groups would result in $p = 1$. If a 6-point scale was used, then it would be reasonable to suspect that participants would select either a 3 or a 4 rating, but that these ratings would essentially be evenly distributed around a mean of 3.5 resulting in the same p value in a subsequent between-subjects analysis. So, in both groups who used the 6-point scale, five participants would rate x as 4, and the other five would rate x as 3. However, if four of those participants who rate x as 3 were to rate x as 4, then a significant difference between the groups would emerge ($p < .10$). Similarly, if three participants in group A change their 3 ratings to a 4 and two participants in group B change their 4 ratings to a 3, then significant differences in the ratings of x again emerge between the two groups ($p < .05$).

One concern regarding the above suggestions is that 6-point scales seem to be prone to Type II errors because small fluctuations in ratings, either for participants within a single group or participants in different groups, can result in statistically significant results. However, an alternative position is that when participants give a rating of 4 on a 7-point scale, they are essentially claiming ambivalence to what the question is asking. It may be that if they were forced to reflect on that question, they would make a more considered decision. A 6-point scale forces participants to consider whether they feel more or less than average about *x*. Thus, small fluctuations in ratings may actually be reflective of a true phenomena because the rating is based on a considered, as opposed to an ambivalent, judgement. Looked at in this way, non-significant between-subjects results where a 7-point scale has been used, especially when mean ratings cluster around 4, may actually be prone to Type II errors.

The third justification for using a 6-point scale follows from the logic outlined above. Because both groups will be introduced to the task in mastery fashion, the suspicion is that experiential differences will be subtle. If the logic outlined above is valid, a 6-point scale may be more sensitive to differences in ratings and therefore better able to pick up any differences between the two groups, should they emerge.

It should be noted that the use of a 6-point scale is also exploratory. The arguments against using a 7-point scale are not conclusive, but merely an attempt to identify a potential weakness in using such scales to identify subtle differences. The forced-choice paradigm has been adopted in many disciplines where participants are asked (forced) to make difficult judgments. For example, in the face-processing literature, participants are often asked to select which of two virtually identical faces is more attractive (e.g. Burt and Perret, 1997). In studies such as this one, no option is given to the participant to say that they have no preference for either face. The argument given for using a forced-choice paradigm in these instances is that the experimental manipulations are so subtle, participants are unlikely to rate attractiveness as a significant feature of either face. A forced-choice paradigm forces participant to make considered judgments about a feature the researcher is interested in. A 6-point rating scale is an attempt to achieve the same.

2.4 Task orientation manipulation

In order to engage a mastery-orientation, instructions to participants are based on those used in studies that have employed similar orientations (see Koestner, Zuckerman and Koestner, 1987 pp. 383-384). The aim was to minimise the possibility that performance was an issue, and to this end, the experiment was introduced to participants in advertising posters as an investigation of what features of games make them interesting. All participants were given exactly the same task instructions with the key orientation manipulation elements being the phrases "In this game, you have to predict what the market is going to do based on the outcome of previous predictions. You predict by pressing either the 'up' or 'down' button on the box-shaped console. You should then continue making predictions until you feel you have enough information to allow you to comment on what you feel are the characteristics of the game". It was hoped that this manipulation would minimise the pressure on participants to perform well at the task.

3 Experiments 1-5

This chapter presents five experiments examining how participants respond, in terms of their persistence and their self-reported enjoyment, to differing levels of success, when tasks are presented to them with either a mastery or a performance-focus. The first of these experiments examines how participants respond to a task presented to them with a mastery-focus.

3.1 Experiment 1: The effect of success/failure on task persistence and enjoyment when a task is presented with a mastery-focus.

The purpose of Experiment 1 was to determine how participants with a mastery orientation, who succeeded on a task would differ from those who failed. In particular, the goal was to determine whether they would persist longer and whether they would report greater interest and enjoyment.

3.1.1. Experiment 1 - Methods

Overview Participants were asked to play a stockmarket-prediction game and told that they were to comment on the characteristics of the game. The game involved predicting whether the market was going to go UP or DOWN on a succession of trials. Outcomes, - that is, whether participants' predictions were successful or not - were controlled by the experimenter. For the first 20 predictions, outcomes were manipulated so that 50% of the outcomes were "UP" and the other 50% "DOWN". This was to prevent participants developing a preference for either response button. Thereafter, outcomes were manipulated so that the Success group achieved a success rate of approximately 70%, whilst the Failure group achieved a success rate of approximately 30%. After playing the game, participants completed a questionnaire about their experiences of the task.

Participants 20 participants (17 female, 3 male) were recruited via the Stirling University Psychology department's participants panel.

Apparatus The stockmarket game was played on a Viglen ENVP15P PC with a software program for the stockmarket game written in Turbo Pascal. Attached to the PC was a normal QWERTY keyboard plus a specially prepared control panel that had only two buttons (right and left). One key was marked 'UP' and the other 'DOWN'.

The testing room was a small windowless room with nothing but the computer, a chair and the task instructions.

Procedure All participants were tested individually. When they arrived for their appointment, they were greeted by the experimenter and taken to the testing room. In the room was a computer. On the desk beside the computer was a sheet containing the experimental task instructions. The task instructions read as follows:

Stockmarket Prediction Game

The Game

In this game, you have to predict what the market is going to do based on the outcome of previous predictions. You predict by pressing either the 'up' or 'down' button on the box-shaped console.

In this version of the task, you MUST NOT make any notes[1]. You should rely solely on your memory.

When you are ready to start, press the ENTER key on the keyboard.

* You will be presented with a message asking you to make your prediction.

* Press either the UP or DOWN buttons on the box-shaped console to make your first prediction.

* After you make your prediction, the computer will reveal what the real outcome was.

You should then continue making predictions until you feel you have enough information to allow you to comment on what you feel are the characteristics of the game.

When you have 'finished', you just stop predicting. You do not need to press any buttons.

PLEASE START WHEN YOU ARE READY

In order to minimise experimenter effects, verbal contact was kept to a minimum. The experimenter simply asked participants to read the instructions (which were pinned up on the wall and attached to the desk), at which point the experimenter left the room[2].

---

[1]This instruction served two purposes. Firstly, it was hoped that participants would believe that the task was also a memory task and secondly, if participants did make notes, they might have noticed the manipulated pattern of outcomes.

[2]No comment as to the content of the experiment or its nature were given. This was felt to be especially important in view of the fact that the program required that the success schedule (30% or 70%) had to be set before the participant entered the room. As such, the experimenter was not blind to condition.

The task instructions asked participants to press the 'enter' key on a keyboard to initiate the program. The first message participants received was:

"Please indicate your prediction by pressing either the UP or DOWN buttons"

Thereafter, participants only pressed one of the two buttons (i.e. either 'UP' or 'DOWN') on the prediction console. Once they had made their prediction, they were presented with the message:

"Your prediction was [participant's prediction]"

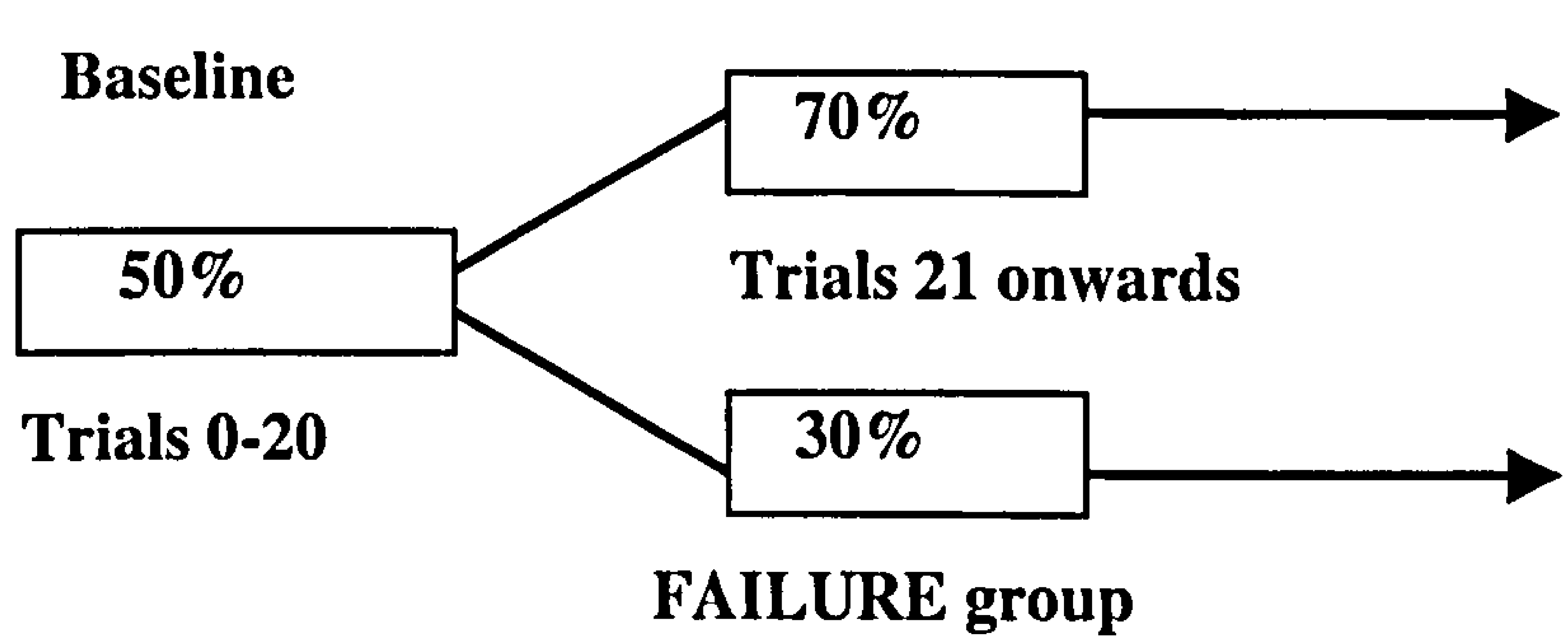
2.5 seconds later, the computer manipulated outcome (either "UP" or "DOWN") was presented for 2 seconds. Participants were then returned to the prediction screen.

"The outcome was [computer-generated outcome]"

Trials were set in blocks of 10. For the first 2 blocks, (i.e., 20 predictions), 50% of the outcomes were 'UP' and 50% were 'DOWN'. This was an attempt to prevent participants developing a preference for predicting just UP or DOWN. Because the outcomes were independent of participants choices during this phase, it was possible for participants to get all 20 predictions right or wrong simply by chance. An (arbitrary) cut-off point for inclusion in the analysis was a minimum rate of success above 60%, averaged across the entire experiment. For the Failure group, the cut-off point was a maximum score of 40%. The original sample contained 20 participants, but because of poor performance during the control period, two participants' performance from the Success group overall success rates were only 48% and 46% respectively and therefore, this data were excluded from subsequent analysis. The opposite pattern (i.e. two participants were highly successful during the baseline period) occurred for two participants in the failure group resulting in these participants achieving overall success rates of 43% and 41%. These data were also excluded.

Thereafter, participants in the Success group were told they were correct on 70% of the trials for each completed block (see figure 1 below for a schematic representation). For the Failure group, the success rate was set at 30%. The *sequence* of success was randomised within each block but participants always either achieved 7 out of 10 (success group) or 3 out of 10 (failure group). No indication was given to participants when to stop, though the computer program terminated after 30 minutes.

Fig 1 - Success/Failure schedule for Experiment 1



When participants finished the task, they left the room and were met by the experimenter who was sitting outside. Participants were taken back into the room and given a questionnaire. The questionnaire read as follows:

STOCKMARKET QUESTIONNAIRE

Your answers to the following questions will be ANONYMOUS

Please circle the number that corresponds most closely to what you thought of that task. [Circle only one number].

How BORING/INTERESTING was the task?

Boring	6	5	4	3	2	1	Interesting

How ENJOYABLE was the task?

Enjoyable	6	5	4	3	2	1	Not Enjoyable

How BADLY/WELL did you think you did?

Badly	6	5	4	3	2	1	Well

How BADLY/WELL did you think other people would do at this task?

Badly	6	5	4	3	2	1	Well

How EASY/DIFFICULT was the task?

Easy	6	5	4	3	2	1	Difficult

Additionally, participants were asked the following general questions about the characteristics of the game:

- What do you think were the main characteristics of the game?
- How could the task be improved?
- How would you describe this task to someone else?

- Any other comments?

These questions were intended to corroborate the cover story and were not included in subsequent analysis. If participants had reviewed the whole questionnaire before committing themselves to an answer, the fact that there were NO questions about the characteristics of the game might have made them suspicious as to the nature of the experiment.

Participants were then thanked for their participation in the experiment and asked if they had any questions. They were then debriefed. This was especially important to reassure participants in the 30% success group that they had not done badly at the task. In addition, participants were additionally given a full written description as to the real nature and purpose of the experiment and asked not to discuss the experiment with other students.

Dependent Measures

Task Persistence - The computer recorded the time from when participants pressed the 'enter' key to the time of their last prediction.

Task Experience - The answers to the questionnaire items.

3.1.2 Experiment 1 - Results

20 participants were tested. 4 participants' data (two from each condition) were excluded from the final analysis.

Participants' answers on the questionnaire were scored so that high values represented positive ratings: high enjoyment, high interest, thought that task was easy, thought they had done well at the task (WellY), thought others had done well at the task (WellO). This meant that the interest, WellY and WellO answers were reverse scored. Throughout this thesis, all the questionnaire data has been scored in this way[3].

Table 2: Means and standard deviations (in italics) for all dependent measures for stockmarket Experiment 1 (n=16)

	70% success		**30% failure**	
	Mean	***s.d.***	**Mean**	***s.d.***
Time spent on task	688 **	*301*	1090 **	*274*
Number of responses	92.7	*59.8*	139.6	*47.0*
Interest	3.0	*0.93*	2.4	*0.74*
Enjoyment	2.5	*0.76*	2.2	*0.89*
Task Difficulty	3.2 **	*1.28*	1.6 **	*0.74*
WellY	3.4 **	*0.74*	1.5 **	*0.76*
WellO	3.4	*0.52*	3.0	*1.41*

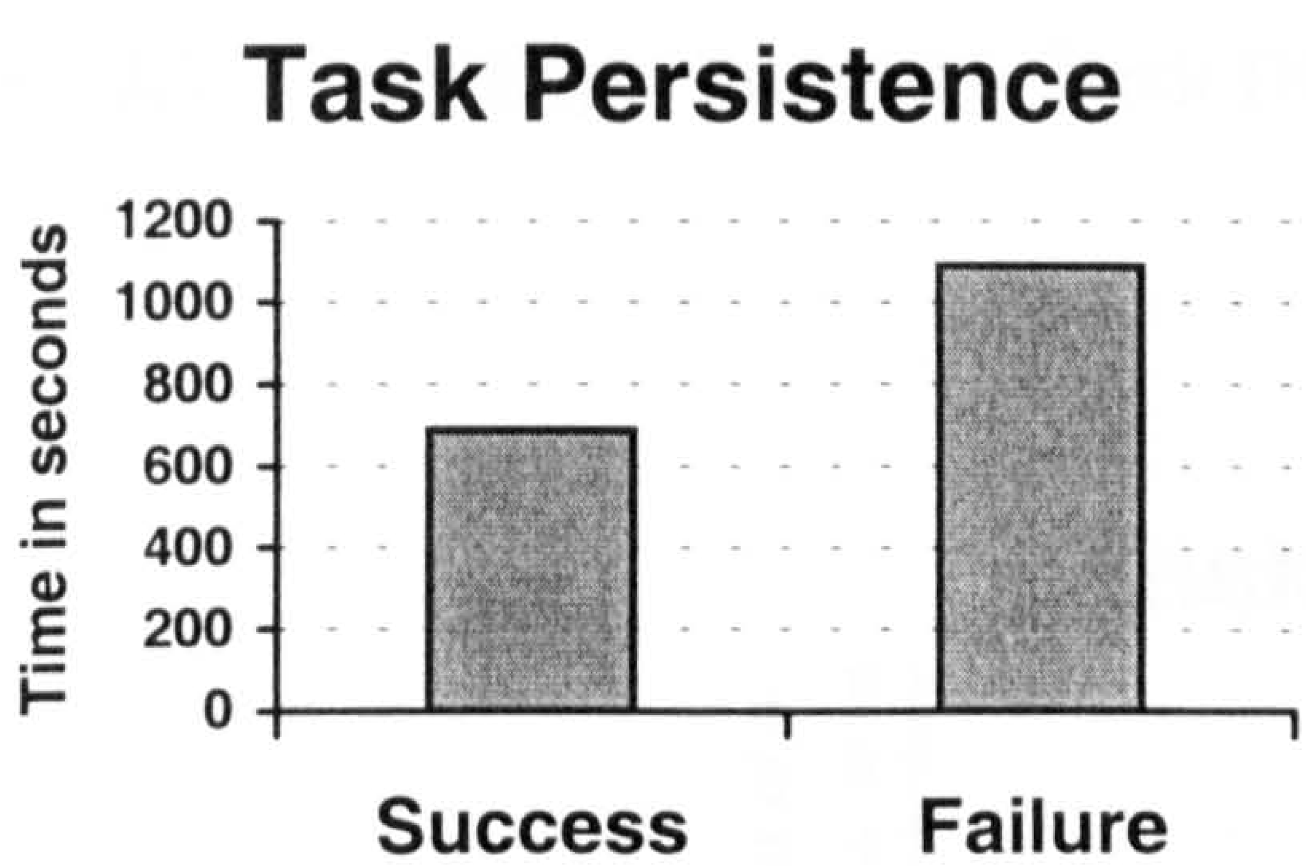
** $p < .01$

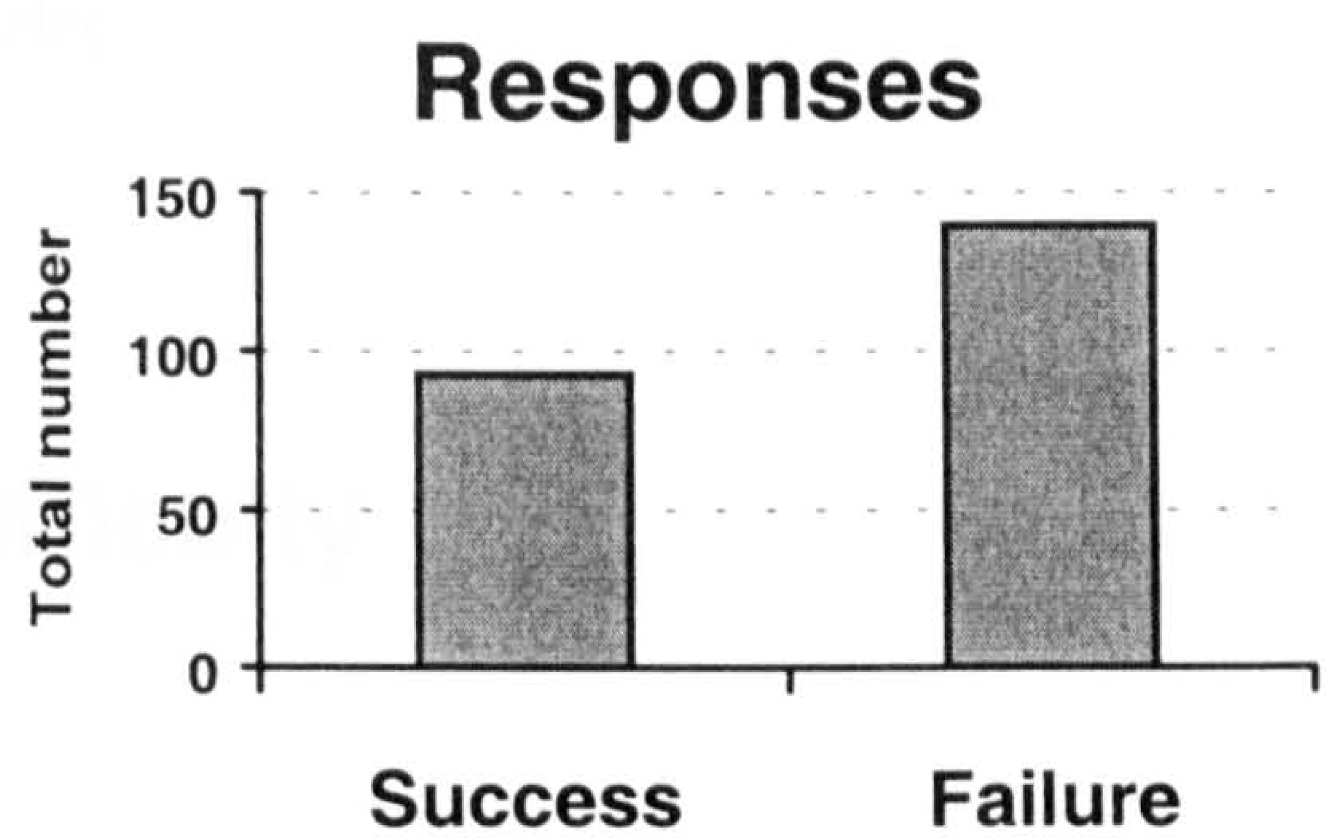
[3] For example, for Experiments 4-10, the task difficulty question on the actual questionnaire was anchored so that high ratings represented high difficulty. However, for the purposes of analysis, this was reverse scored so that a high rating represented low difficulty.

All data (in this experiment and this chapter) was analysed using between group t-tests except for the WellY vs. WellO analysis, which was analysed with a within-subjects t-test.

Table 2 above shows the means and standard deviations for the dependent measures of task persistence and the average ratings for the questionnaire items. Significant differences between the Success and Failure groups were observed for task persistence, task difficulty and WellY. The results for each dependent measure are dealt with individually in the following sections.

3.1.2.1 Experiment 1 - Task Persistence and Number of responses





Analysis revealed that participants in the Success group persisted for significantly less time at the task compared to participants in the Failure group, $t\ (14) = -2.79, p < .05$. In terms of responses, that is, the number of predictions that participants made, those in the success group also made less predictions, but this difference was not significant (p = .10)

3.1.2.2 Experiment 1 - Enjoyment and Interest

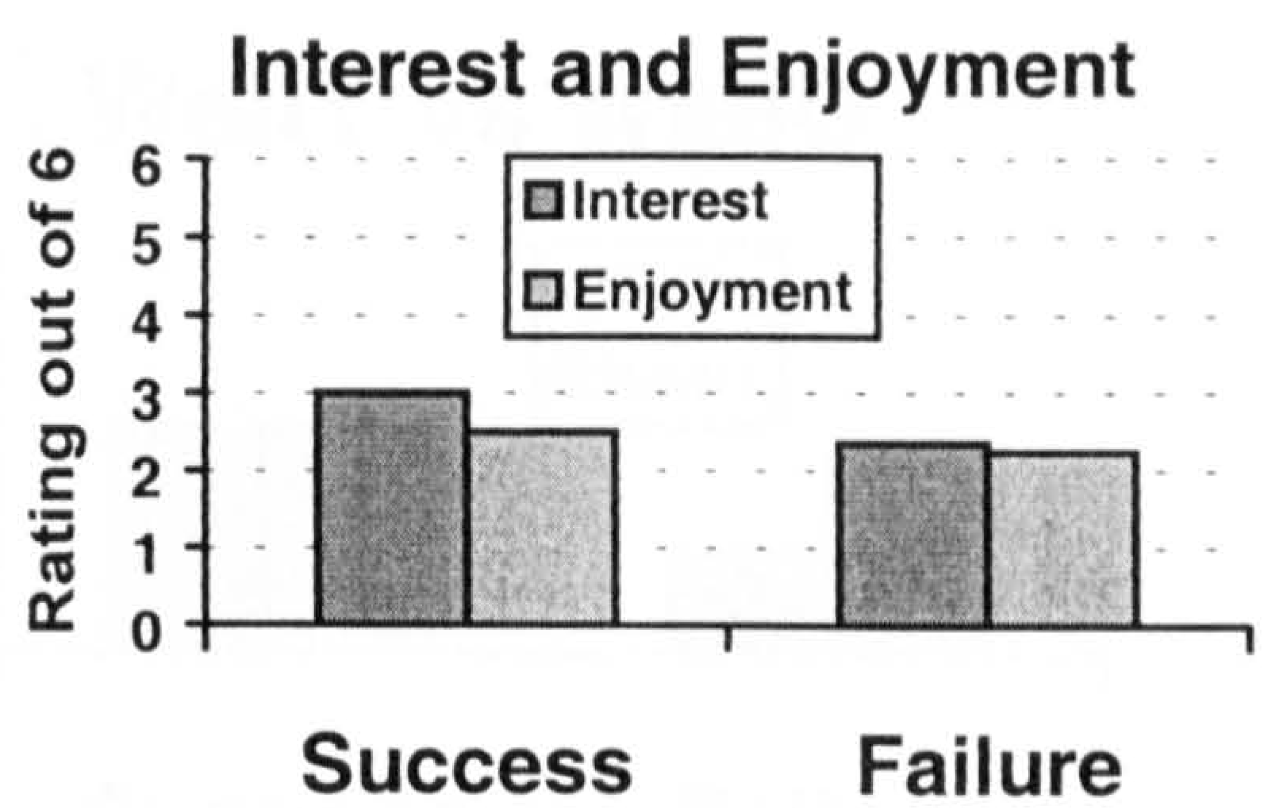
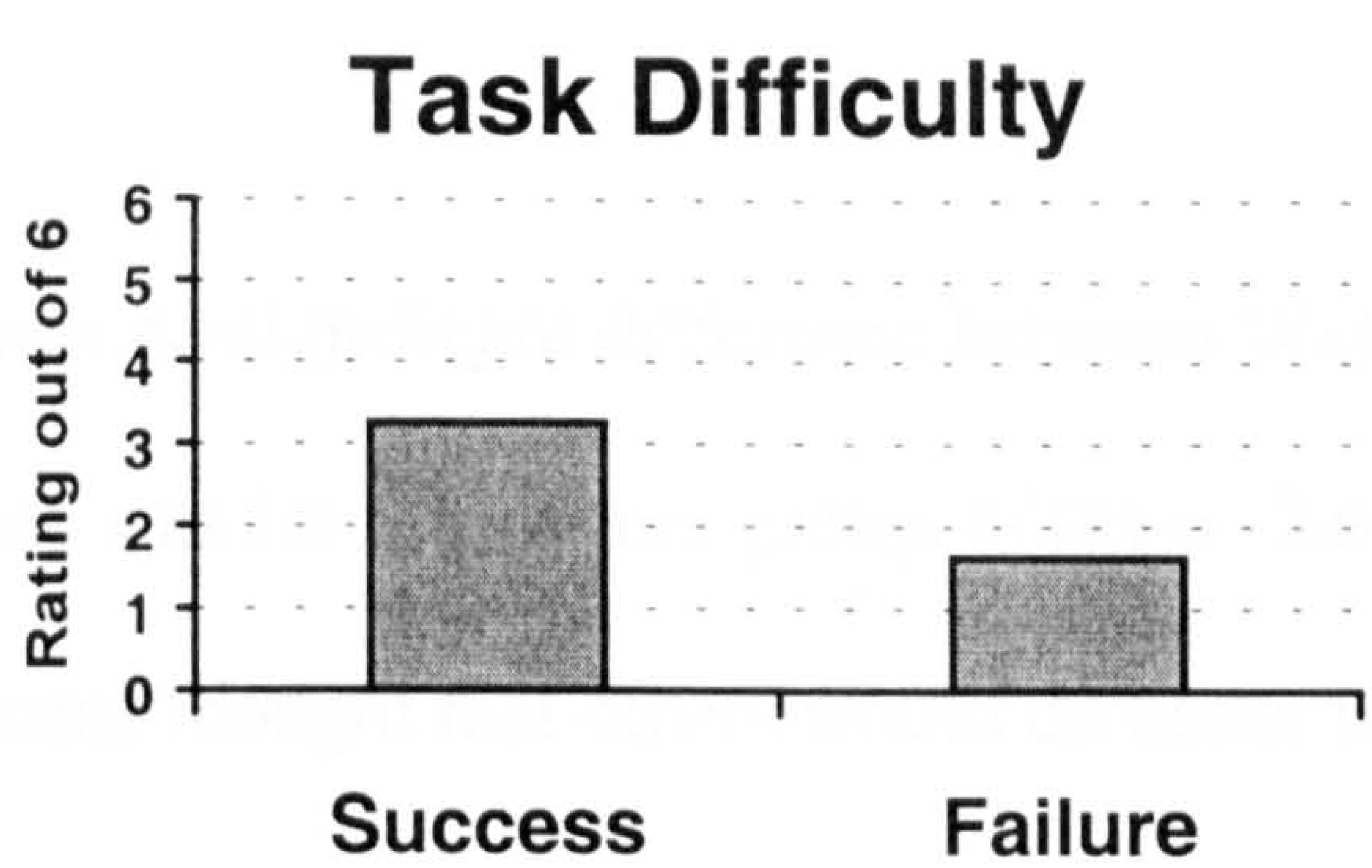


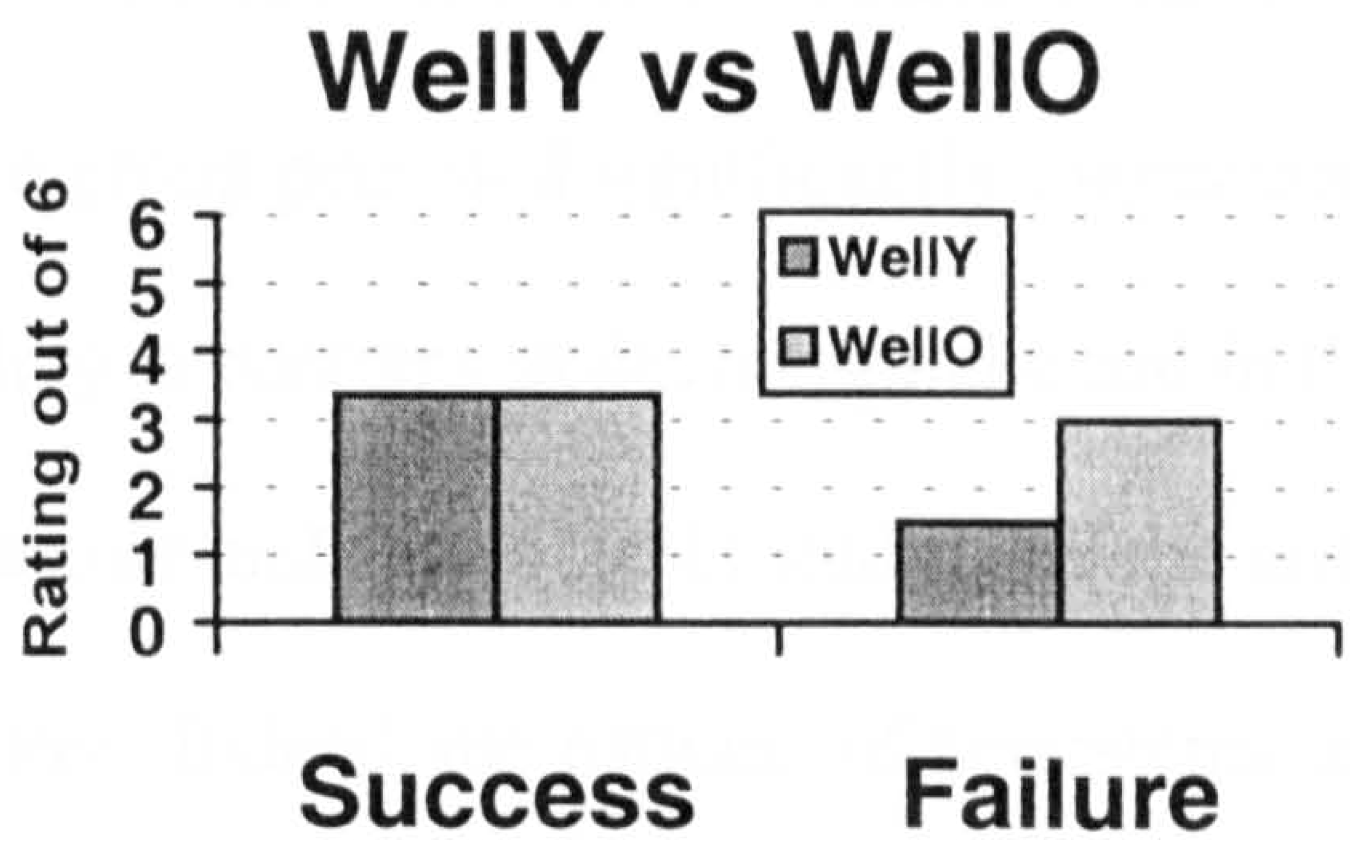
Table 2 reveals that participants in the Success group enjoyed the task more (M = 3.0 vs M = 2.37) and also found it more interesting (M = 2.5 vs M = 2.25), compared to the Failure group. However, neither difference was significant.

3.1.2.3 Experiment 1 – Task Difficulty



Participants in the Success group rated the task as significantly easier compared to the Failure group, $t\ (14) = 3.1$, $p < .01$.

3.1.2.4 Experiment 1 – WellY vs. WellO



WellY When participants were asked to rate how well they had done, those in the Success group rated their performance significantly higher than those in the Failure group, $t\ (14) = 5.00$; $p < .001$.

WellO There was no significant difference between the two groups for how well they thought others (WellO) would do at the task.

WYWO There was no significant difference between WellY and WellO for the Success group, but there was for the Failure group, $t\ (14) = -2.65$, $p < .05$. For the Failure group, participants thought that others would do better than they did.

3.1.2.5 Experiment 1 – Mood and activity

There were no significant correlations between time spent on the task and reported interest and enjoyment.

3.1.3 Experiment 1: Discussion

Experiment 1 showed that when the stockmarket task was presented with a mastery-focus, the Failure group persisted significantly longer compared to the Success group ($p < .05$). This finding is contrary to the marginal trend in the opposite direction reported in the Ryan, Koestner and Deci (1991) study, and the null findings in the Anderson and Rodin (1989) study. Indeed, the patterns of persistence more appropriately resemble the results for the performance-focus groups reported by Ryan et al (1991), where participants in the no feedback group persisted longer than those who received positive feedback, although the difference was not significant.

Several factors in Experiment 1 could help to explain why participants persisted for significantly longer in the Failure group. Firstly, unlike the Ryan et al (1991) and the Anderson and Rodin (1989) studies, performance in the Success group was compared to that of a group who experienced failure. In the Anderson and Rodin study, participants who were told they had performed in the 55th percentile reported feeling disappointed but not devastated. In Experiment 1, the mean WellY rating for the Failure group was 1.50 suggesting that this group considered their performance to have been substandard. Additionally, the Failure group rated both how difficult they thought the task was, and how well they thought they had done at the task significantly lower than the Success group. It would therefore seem that participants in the Failure group experienced a level of failure that was more then just mildly disappointing. Thus it might be that when participants experience failure, as opposed to no feedback or mild negative feedback, it is this factor that may cause them to persist for longer at a task, even when it is presented with a mastery-focus.

A second aspect of the procedure that could explain why participants persisted so long after performing poorly was the use of a post-task interview. Although the task was presented with a minimal emphasis on performance, participants were aware that they would be asked questions about the task after they had completed it. It is therefore plausible that they also might have been concerned about being asked about their performance and therefore had engaged with the task with a performance rather than a mastery-focus. For example, in the more traditional free-choice paradigm where participants are either secretly filmed or observed via one-way mirror, it is reasonable to suspect that participants would not expect to be challenged about their behaviour in this session when the experimenter returned to give them their questionnaires. In the paradigm adopted in Experiment 1, participants were specifically advised that they would be asked about their experiences of the task and they might have felt embarrassed about having to report that they had done poorly. If, as a result, they had experienced the task with a performance-focus, then the patterns of persistence would resemble those observed in the performance-focus conditions in the Ryan et al (1991) study.

A third possible explanation for the greater persistence after failure can be derived from the suggestion that when a task is presented with a mastery-focus, individuals may persist longer with it in an effort to try and master or understand it (e.g., Dweck and Leggett, 1988; Nicholls, 1984). For example, it may have been that participants were not concerned with performing well but were simply striving to prove to themselves that they could solve the task. Whilst this argument is certainly plausible, it does not explain why participants in the Failure group experienced the task significantly more negatively relative to those in the success condition. For example, there were significant differences between the WellY and WellO ratings for participants in the Failure group, but not for the Success group. Additionally, although the differences in interest and enjoyment ratings were not significant, participants in the Failure group tended to report less positive experiences of the task. If participants in the Failure group were simply striving to master the task rather than being concerned with performance, it is not clear why they reacted so negatively when they did not succeed.

Ryan (1982) proposed a form of performance-focus which could account both for the greater persistence of participants in the Failure group and the more negative feelings. He suggested that when individuals become concerned about doing well, they may continue to persist at a task even though they are not enjoying it (see also, Ryan, Koestner and Deci, 1991; Baumeister and Tice, 1985). The data from Experiment 1 suggests that despite the intention to present the task with a mastery-focus, the type of persistence that was observed seemed to meet the Ryan's criteria for ego-involvement. Thus, it may have been that the task was actually experienced with a performance-focus.

3.2 Experiment 2: The effect of success/failure on persistence and enjoyment when participants improve or get worse at a task.

Participants in Experiment 1 who performed poorly persisted more than those who did well. One reason suggested for this was that the method employed to measure persistence may have inadvertently caused participants to become concerned about their performance, resulting in the type of ego-involved persistence identified by Ryan (1982), that is, greater persistence after failure but lower ratings of interest and enjoyment. Thus, it may be that when persistence is measured using the method employed in Experiment 1, even when attempts are made to present that task with a minimal emphasis on evaluation, participants will still experience a desire to perform well at it.

However, in Experiment 1, the differences between the groups for their ratings of interest and enjoyment were not significant. So it is still unclear whether those in Failure condition experienced the task any more interesting and enjoyable than those in the Success group. They may have recognised that they had done poorly and may have lost confidence in their ability, but it is still unclear why these factors did not transfer to their interest and enjoyment ratings.

One explanation for the lack of significant differences in the interest and enjoyment ratings may be that individuals in the Success group simply did not experience sufficient success to feel that they had truly been successful. Given that the highest mean interest, enjoyment or WellY rating was 3.37 out of 6 (for WellY in the success group), it may be that the 70% success rate was simply too low. If so, increasing the success levels to 80% or 90% in a subsequent experiment should logically result in more positive ratings. Additionally, reducing the failure levels (to 10%) may also have the effect of polarising the Success and Failure group experiences, and thus give greater scope for separating out the interest and enjoyment ratings. However, by increasing success levels from a baseline 50% straight to 80% or 90% may lack plausibility and lead to individuals externalising their performance (i.e., it was the computer that controlled the success [external], not my own personal performance [internal]). To avoid this, success rates in Experiment 2 were changed gradually rather than abruptly.

Experiment 2 therefore employed a slightly amended methodology in an effort to see whether or not there really were differences in interest and enjoyment ratings between the two groups. In Experiment 2, there were two manipulations. Firstly, the optimum success participants could achieve in the Success group was increased from 70% to 90%. Secondly, rather than have participants achieving a success rate of 90% straight after the baseline phase, an improvement schedule was introduced. After the initial baseline period, participants improved by 10% in each successive block of 10 trials (e.g. 60%, then 70%, then 80% etc.) until they reached a 90% success rate.

The improvement schedule was also an attempt to make the task more ecologically valid. It was felt that when individuals engage in a new task, they expect to learn at least something. For example, it is unlikely that given a problem-solving task one would experience no change in one's ability to work out the solution, especially if the task has a prediction sequence, as the one in this experiment does. Also, the increase in performance to 90% might have been unbelievable if participants went straight from the baseline 50% to 90% in just 10 trials. By including the improvement schedule, it was hoped not only that participants would feel that they were doing well, but that this improvement would minimise suspicion that performance outcomes were being determined by the computer program, rather than their own personal efforts.

Participants in the Failure group received the same success pattern as the Success group but in the opposite direction. This was expected to increase the feelings of failure by promoting a feeling of not only failing, but getting gradually worse.

Following the findings from Experiment 1, it was expected that:

- Participants in the Failure group would again persist for longer at the task than participants in the Success group.
- Participants in the Failure group would now find the task significantly less interesting and enjoyable than participants in the Success group.

3.2.1 Experiment 2 - Methods

Overview Experiment 2 was conducted in the same manner as Experiment 1. Participants were asked to play a stockmarket-prediction game and told that they would be asked to comment on the characteristics of the game. The levels of success were manipulated again, but this time so that participants in the Success group experienced gradual improvement in performance, whilst participants in the Failure group experienced a decline.

Fig 2 - Success/Failure schedule for Experiment 2

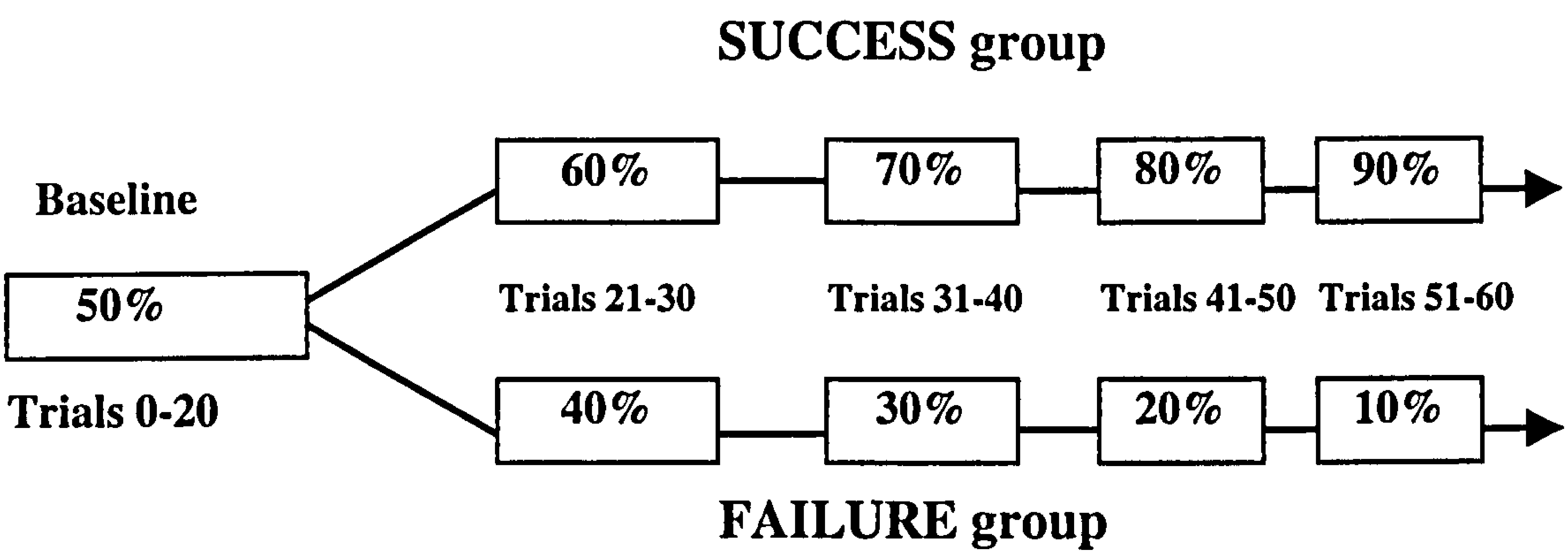


Figure 2 above shows a schematic for the success/failure schedule for Experiment 2. As with Experiment 1, for the first 20 predictions, outcomes were manipulated to prevent participants developing a key choice preference (i.e. to stop them from just predicting either ‘UP’ or ‘DOWN’ all the time). For the next 10 predictions, participants in the Success group achieved a 60% prediction success rate, whilst the Failure group achieved a 40% success rate. For the next 10 predictions, the success group achieved a 70% success rate whilst the Failure group achieved a 30% success rate. This trend continued until the Success group achieved a 90% success rate whilst the Failure group achieved a 10% success rate. After trial 60, the Success group achieved a consistent 90% success rate whilst the Failure group achieved a consistent 10% success rate.

Participants 17 subjects (9 female, 8 male) were recruited from a visiting Open University Summer School (A102 - Arts Foundation course). One set of data was discarded after the participant revealed that they were receiving counselling for a gambling related addiction. It was felt that such a condition might have seriously affected performance.

Unlike the previous experiment, these participants either freely volunteered to take part or agreed to take part after being asked by the experimenter. There was no financial inducement or obligation on the part of the participants to take part in the experiment.

Apparatus All testing took place in a specially prepared room in the participants' halls of residence, between 8.30pm and 11.30pm every weekday night. The apparatus consisted of the same equipment used in Experiment 1 and the same software program, this time written to incorporate the improvement schedule.

Procedure The procedure was identical to experiment 1 except for the "improvement schedule" outlined in the overview for this experiment. The same questionnaire used in Experiment 1 was used in Experiment 2.

3.2.2 Experiment 2 - Results

All analyses were analysed with between-group t-tests except for the WellY vs. WellO analysis which was analysed with a within-subjects t-test.

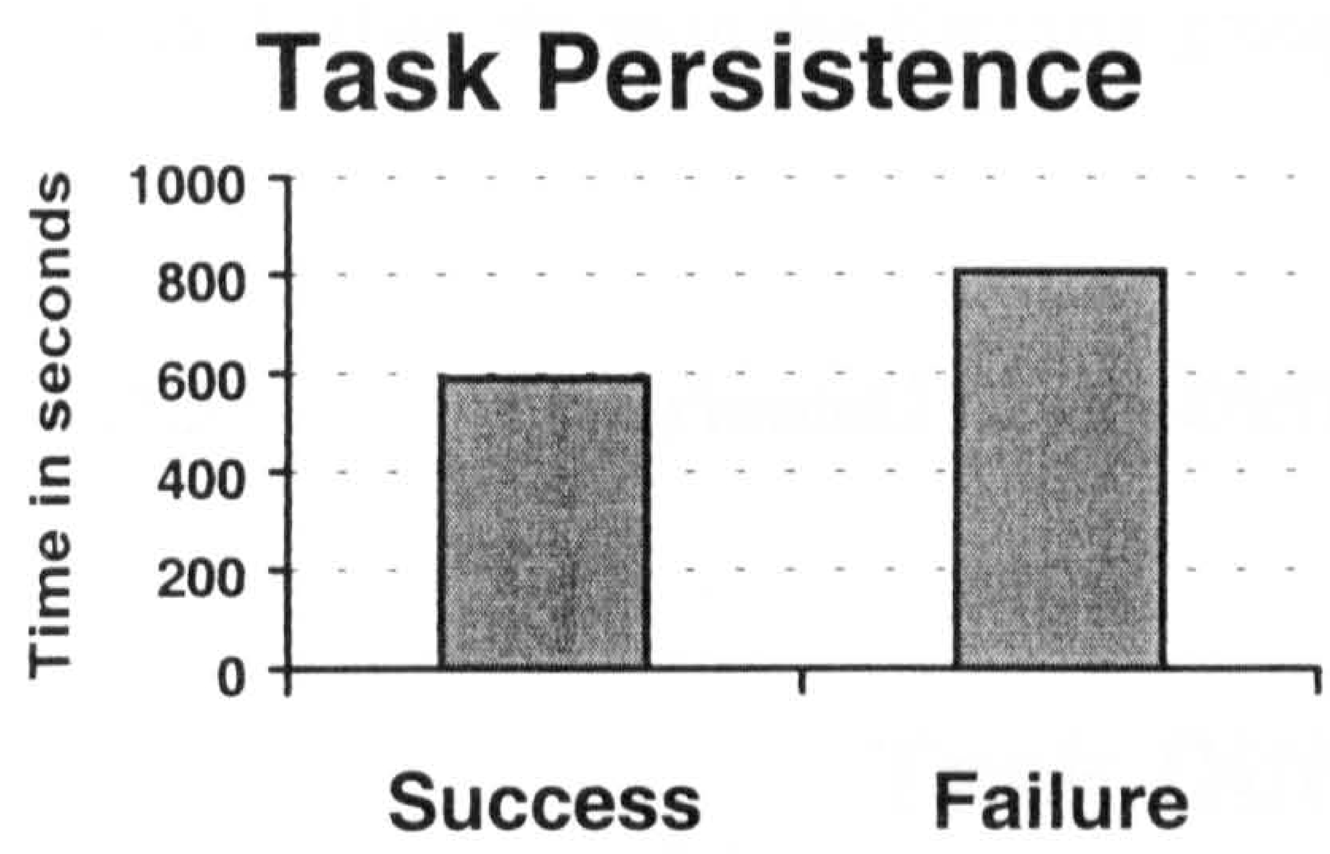
Table 3: Means and standard deviations (in italics) for all dependent measures for stockmarket Experiment 2 (n=16).

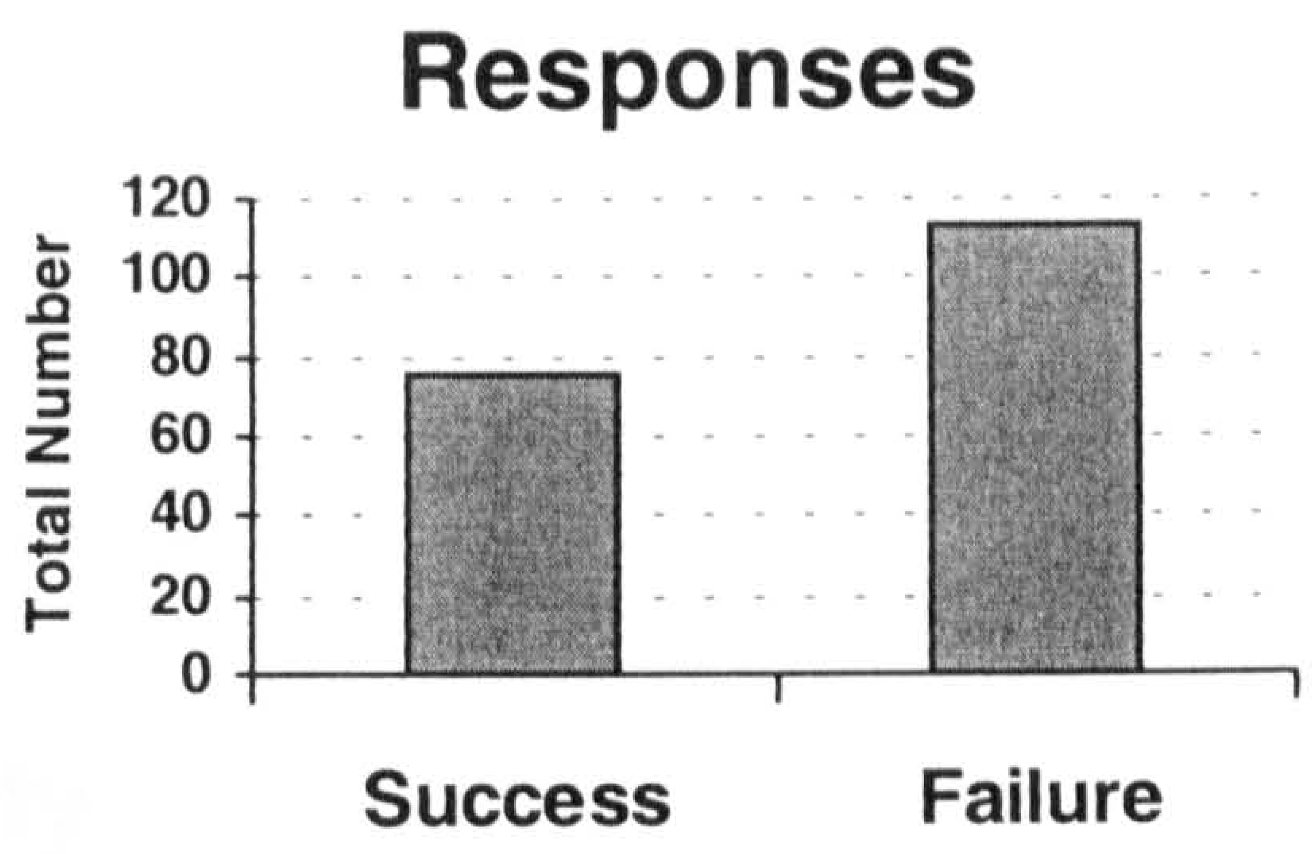
	90% success		**10% failure**	
	mean	***s.d.***	**Mean**	***s.d.***
Time spent on task	549.2 *	*52.8*	804 *	*274*
Number of responses	75.4 **	*13.4*	113 **	*47.0*
Interest	3.1	*1.27*	3.0	*0.74*
Enjoyment	3.3	*1.49*	3.9	*0.89*
Task Difficulty	4.0	*1.69*	3.1	*0.74*
WellY	3.6	*1.06*	2.9	*0.76*
WellO	4.5	*0.76*	3.5	*1.41*

* $p < .05$ ** $p < .001$

Table 3 above shows the means and standard deviations for the dependent measures of task persistence and the average ratings for the questionnaire items. Significant differences between the Success and Failure groups were observed for task persistence, task difficulty and WellY. The results for each dependent measure are dealt with individually in the following sections.

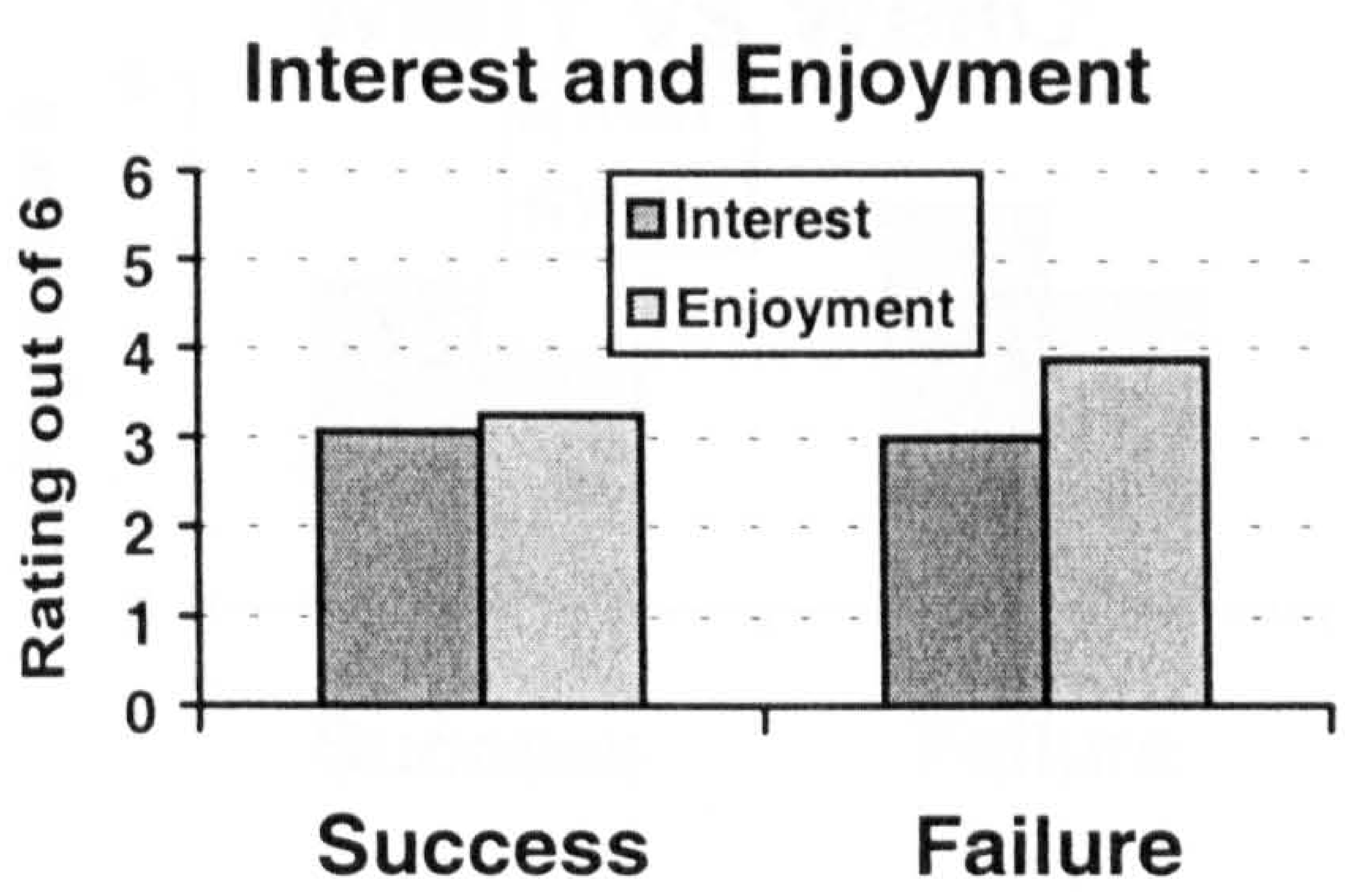
3.2.2.1 Experiment 2: Task Persistence





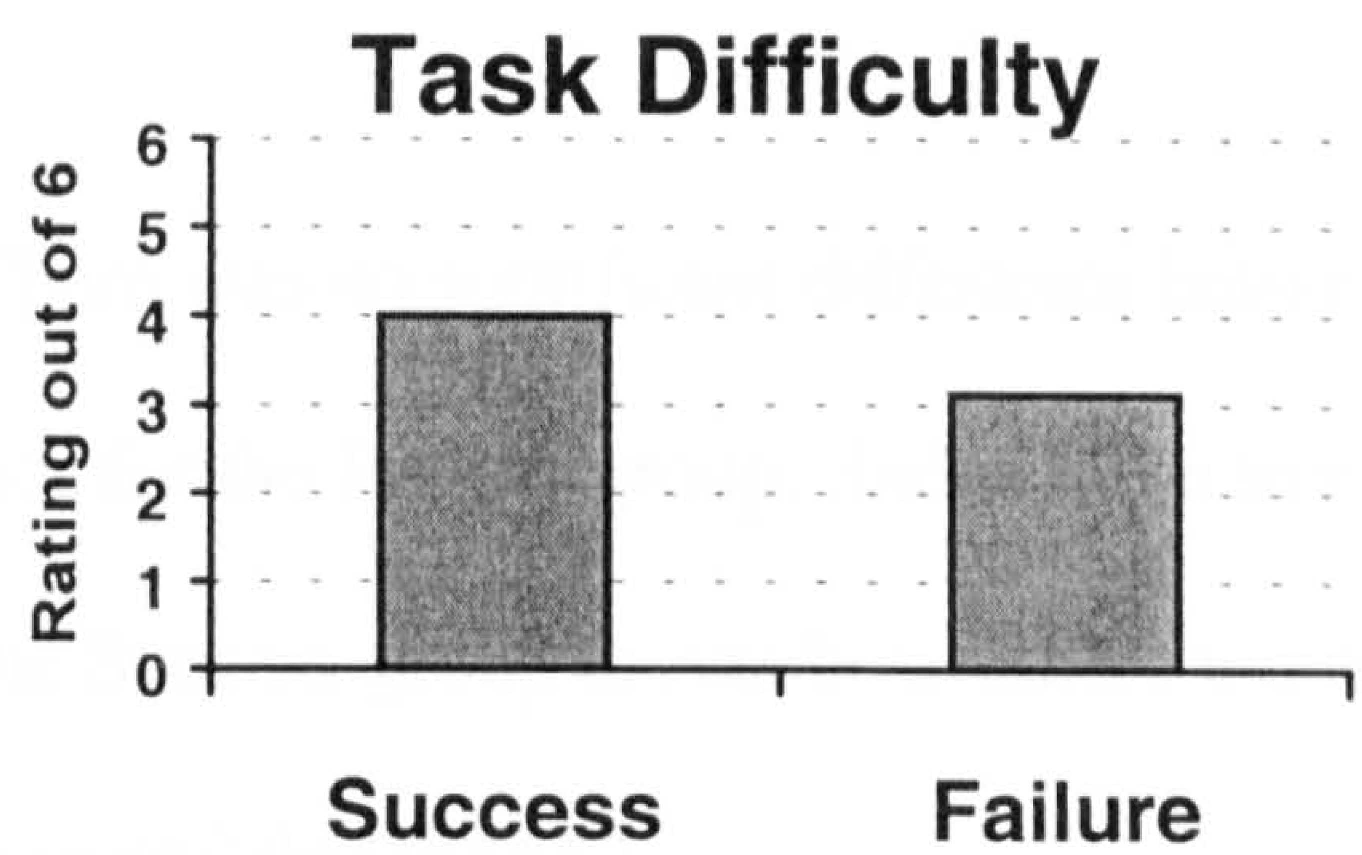
Analysis showed that there was a significant difference between persistence in the Success group and the Failure group, $t\ (14) = -2.84$, $p < .05$. As with Experiment 1, participants in the Failure group persisted longer at the task. The significant difference in the amount of time participants spent at the task was also reflected in the number of responses they made, $t\ (14) = -4.69$, $p < .001$. Participants in the Failure group made significantly more responses.

3.2.2.2 Experiment 2: Interest and Enjoyment



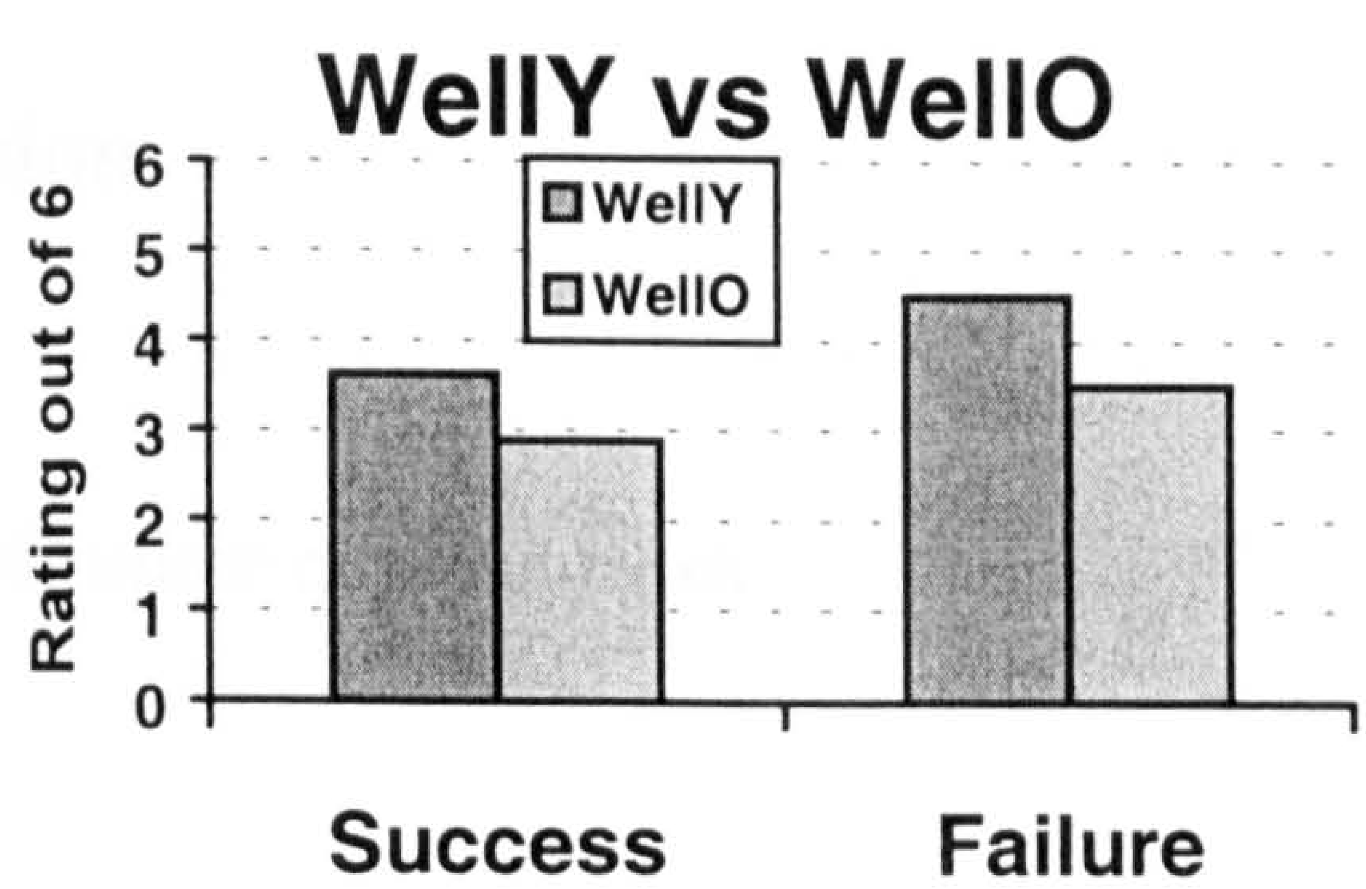
There were no significant differences between the groups with relation to how interesting or enjoyable participants found the tasks. With regards to enjoyment, contrary to hypothesis, participants in the Failure group tended to rate the task more enjoyable than those in the Success group.

3.2.2.3 Experiment2: Task Difficulty



As with Experiment 1, participants in the Success group reported the task as easier than those in the Failure group, but the difference was not significant.

3.2.2.4 Experiment 2: WellY vs WellO



<u>WellY</u> In a similar fashion to Experiment 1, participants in the Success group rated their performance higher than those in the Failure group, but this time the difference was not significant.

<u>WellO</u> There were no differences between the Success and Failure groups for how well they thought others would do at the task.

<u>WYWO</u> There was no significant difference between WellY and WellO for the Success group nor for the Failure group. Indeed, the trend in Experiment 2 was for participants in the Success group to rate how others would do at the task (WellO) much higher than they would do (WellY).

3.2.2.5 Experiment 2 – Mood and Activity

In a similar fashion to Experiment 1, there were no significant correlations between time spent on the task and reported interest and enjoyment.

3.2.3 Discussion

Task Persistence and Number of Responses

The hypothesis that participants in the Failure group would persist at the task longer than participants in the Success group was supported. This difference in persistence was also reflected in the number of responses made. These findings replicated the findings from the first experiment.

However, whilst the within-experiment persistence data for Experiment 2 confirmed the findings observed in Experiment 1, a more perplexing issue arises when the persistence times for the Failure groups are compared between the two experiments. Because of the greater levels of failure in Experiment 2, it would be reasonable to expect that participants would persist for even longer in the failure condition in Experiment 2 compared to Experiment 1. This hypothesis was not supported, and, moreover, the patterns of persistence were in the opposite direction, namely, less, not more, persistence in Experiment 2 (Failure group persistence, Experiment 1, M = 1090 secs; Failure group persistence, Experiment 2, M = 804 secs).

One reason that this prediction was not supported may lie in the types of attributions participants made about their performance. Weiner (1985, 1992) suggests that one determinant of persistence is how believable individuals perceive their performance to be. Recall that in chapter 1 (section 1.4), Weiner's analysis suggested that individuals can attribute their behaviour to either internal causes (e.g., it was due to my *ability*) or external causes (e.g., the task is very difficult and there is little that I can personally do about that difficulty). According to Weiner, when individuals fail, their responses to that failure will depend on their attributions. It may be useful to examine Weiner's analysis more closely to explain the differences in persistence times for the failure groups in Experiments 1 and 2.

In Experiment 2, it was particularly noticeable that during the post-experimental briefing, participants in the Success group tended to attribute their success to their own ability, whereas participants in the Failure group tended to attribute their performance to the computer. A typical comment was *"..it seemed to me that no matter what I was predicting, the computer was coming up with the opposite answer"*. If in the Failure group, individuals' attributions were indeed external (e.g. my performance was due to the computer), one consequence is that failure would not necessarily imply that the individual was doing badly at the task. If so, this would explain why participants in the Failure group in Experiment 2 persisted less compared to the same group in Experiment 1.

However, the above suggestion seems to beg the question why, if participants in the Failure group in Experiment 2 realised that the outcomes were being manipulated, they did not persist less than participants in the Success group? If the participants in the Failure group had realised that the computer was determining the outcomes, one might think that they would have desisted immediately. One explanation is that it would take participants in the Failure group some time to come to this conclusion. There was no suggestion in the instructions that outcomes were computer generated; this conclusion had to be arrived at via some additional cognitive process. For participants in the Success group, success was apparent, they did not need to look to any justification for their performance. The task was presented to them as one that was controlled by them, they were successful, so therefore that success must be because of their own personal abilities. As the task was presented to participants in the Failure group in the same way, performance was also due to their personal abilities. So, in failing, looking to justify that failure required a secondary justification paradigm, one that was counter to their original perception that the task outcomes were regulated by them. At first, they would have reacted to the failure by increasing effort or trying other strategies but when they continued to fail would have needed to justify their poor performance by attributing it to external factors rather than ability. What I am suggesting here is that developing such a justification schema took longer. In terms of search-level strategies, participants in the Success group only needed to search the level where the case that they were presented was true, namely that the outcomes they were achieving were a result of their own personal performance. The Failure group not only needed to overcome the suspicion that they were performing badly because of their own efforts, but needed to generate an alternate hypothesis which allowed their personal feelings of competence to be satisfied, hence they persisted longer and made more responses.

So, it makes sense that participants in the Failure condition should persist longer when compared to those in the Success condition. However, when comparing the level of persistence in the Failure group in Experiment 1 with same group in Experiment 2, because participants in Experiment 2 may have had a stronger tendency to attribute their performance externally ("it was the computer, not me") because their level of Failure was unbelievably low, then it is reasonable to suspect that, compared to Experiment 1, they would leave the task earlier.

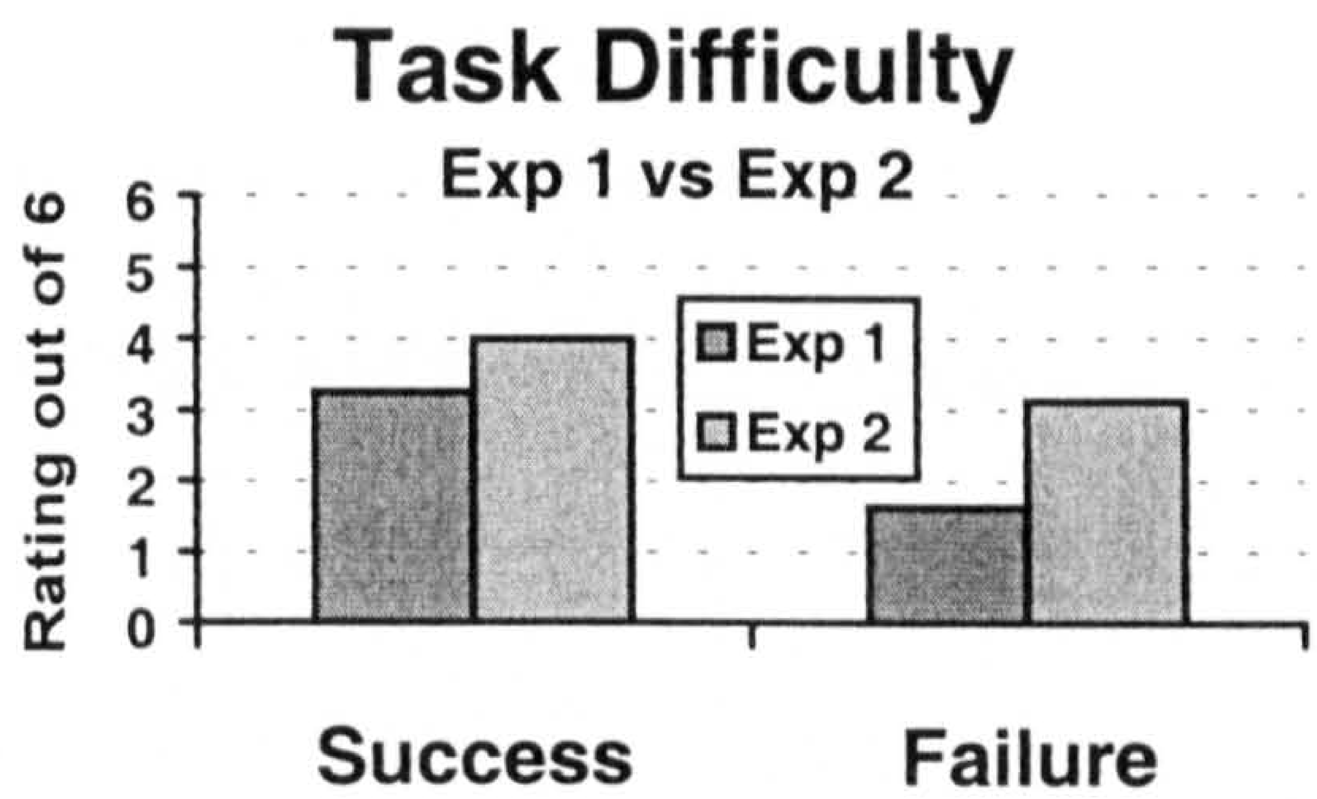
If participants in the Failure group in Experiment 2 attributed their poor performance to the computer, this could also to explain some other aspects of the results. For example, there were no significant differences between how interesting or how enjoyable participants in the two groups found the task. Indeed, participants in the Failure group rated the task as *more* enjoyable than those in the Success group. This anomaly is surprising but explicable. If, as previously mentioned, participants in the Failure group were attributing their performance to the computer, then, in one sense, they had "worked out" what was happening. In terms of success, they had "cracked the code" –

that is, no matter what they did, the computer would give the opposite outcome and/or, that success was not controlled by them. If, as Bandura (1990) suggests, self-efficacy is an important determinant of behaviour in task situations, it is plausible that such an experience was, at least to some degree, competence enhancing. It was as if participants were saying to themselves "well, although I failed at that task, I was smart enough to figure out *why* I failed". If, as Deci (1975; see also Deci and Ryan, 1985) suggest, interest and enjoyment are correlates of competence, then it would not be surprising that the Failure group reported relatively high levels of interest and enjoyment. This could be because participants in the Failure group at least had the satisfaction of working out what the experiment was about.

Task Difficulty

One further question that requires addressing is why, given that participants in the Success group achieved a 90% success rate, they did they not rate the task as significantly easier than those in the failure group? One possible explanation may have been the high ratings of the task difficulty from the failure group.

To elaborate, if the average task difficulty ratings in Experiments 1 and 2 are compared (see graph below), participants in the Success group for Experiment 2 found the task easier than the same group in Experiment 1. This seems reasonable given that participants in the success group in Experiment 2 achieved a higher success rate. However, the same trend is <u>also</u> true for the Failure group. Moreover, the difficulty rating for Experiment 2 for the Failure group was similar to the ratings for the Success group in Experiment 1. So rather than the greater amount of failure in Experiment 2 causing individuals to rate the task as more difficult, the greater failure seemed to have the opposite effect, namely that participants rated the task as easier.



However, when the results are interpreted in terms of the attribution analysis, the increased task difficulty ratings do not seem so anomalous. If participants did not believe the outcomes they were achieving, they would have been less likely to view the task as difficult. This would account for why, despite performing so poorly, they still rated the task as relatively easy.

Conclusions for Experiment 2

Experiment 2 was successful in that it produced similar persistence findings to Experiment 1 and increased ratings of interest, enjoyment and task difficulty. However, it appears that the experimental manipulations for the Failure condition had the unintended consequence of increasing the quality of experience for these participants. It appears that reducing the failure criteria to 10% changed the type of attributions participants made about their performance, resulting in them feeling <u>more</u> rather than <u>less</u> positive about their achievements. Thus while the improvement to 90% manipulation seemed to be a useful innovation, the reduction to 10% may not have been. One possible course of action could be to retain a 90% level of success but compare it to a decline to 30%.

3.3 Experiment 3: Persistence and interest/enjoyment responses to success (improvement in performance) and failure under *performance-focus* conditions

In section 3.1.3, it was suggested that the patterns of persistence and types of experiences that were reported seemed to be more appropriate to what Ryan (1982) had labelled "ego-involved" persistence. Ryan (1982) had suggested that when participants failed at a task that was presented with an emphasis on evaluation (i.e., a performance-focus), participants would use the subsequent free-choice period to try and recover the self-esteem they had lost by their earlier poor performance. In Ryan's (1982) study, when asked to comment, participants reported not having enjoyed engaging with the task. The phenomenon identified by Ryan (1982) was a mismatch between task persistence and task experiences, that is, persistence without enjoyment. This phenomenon was also observed in Experiments 1 and 2, but this time under what was supposed to be a mastery-focus. In the discussion section of Experiment 1, it was suggested that because participants had to complete a post-task questionnaire, this might have caused them to become concerned about their performance.

One way of validating the claim that participants were performance-focused would be to replicate the paradigm used in Experiments 1 and 2, but this time present the task with a performance-focus. If participants were indeed performance-focused in Experiments 1 and 2, then the results with a specific performance-focus should be exactly the same as those found earlier with an apparent mastery-focus. Experiment 3 therefore investigated how success/failure at a task which was presented with a performance-focus would affect participants' persistence, interest, and enjoyment.

Rather than participants being asked to comment on the characteristics of the game (as they were in Experiments 1 and 2), to induce concerns about their performance, participants were this time told that performance at the game was reflective of their problem-solving ability.

Additionally, because the failure manipulation in Experiment 2 resulted in participants reporting that they felt the task outcomes were determined by the computer, this manipulation was amended. Instead of participants attaining a final success level of 10%, the final success level was 30% (i.e., the same level used in Experiment 1).

In Experiment 3, the same hypotheses as Experiment 2 were generated

- Participants in the Success group would persist less at the task than participants in the Failure group.
- Participants in the Success group would find the task more interesting and enjoyable than participants in the Failure group.

3.3.1 Experiment 3 - Methods

Overview In most respects, Experiment 3 was conducted in the same manner as Experiment 2. However, unlike Experiments 1 and 2, participants were told that the task was one that measured problem-solving abilities. Outcomes, - i.e. whether participants predictions were successful or not - were manipulated in the same fashion as in Experiment 2 except that the lowest failure rate in the failure group was 30% rather than the 10% employed in Experiment 2.

Participants 24 subjects (12 female, 12 male) were recruited from the Open University Summer School (A102 - Arts Foundation course). In a similar fashion to Experiment 2, these participants either freely volunteered to take part or agreed to take part after being asked by the experimenter.

Apparatus All testing took place in the same specially prepared room used in Experiment 2 and utilised the same computer hardware and software program.

Procedure The procedure was identical to Experiment 2 except for a change in the task instructions and the amendment to the "improvement schedule" for the Failure group only.

Fig 3 - Success/Failure schedule for Experiment 3

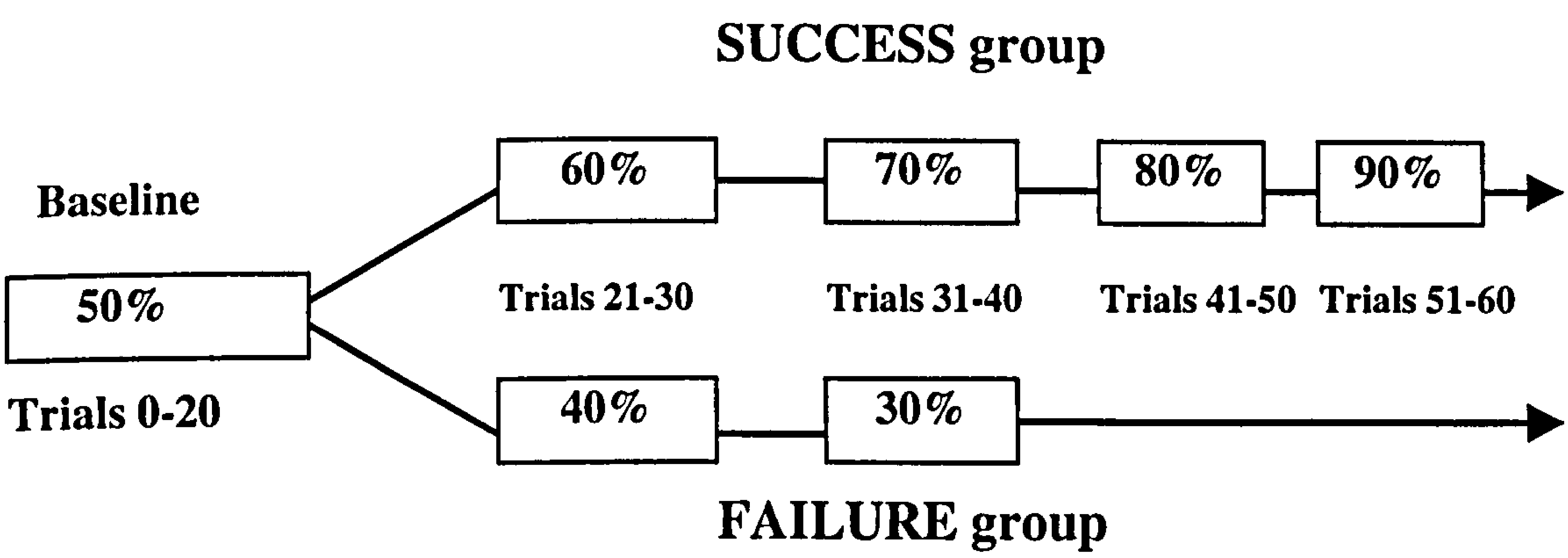


Figure 3 above shows a schematic for the success/failure schedule for experiment 3. As with Experiment 1, for the first 20 predictions, outcomes were manipulated to prevent participants developing a preference for the 'UP' or 'DOWN' button. For the next 10 predictions, participants in the Success group achieved a 60% prediction success rate, whilst the Failure group achieved a 40% success rate. For the next 10 predictions, the Success group achieved a 70% success rate whilst the Failure group achieved a 30% success rate. This trend continued until the success group achieved a consistent 90% success rate. However for the Failure group, their success rate remained at 30% once this level was reached, rather than going on to 10% as in Experiment 2.

The task instructions were altered to make performance more salient and therefore make the task more performance-focused. Participants were told that;

> *"..we are interested in canvassing the views from people who are studying for a degree. The game is supposed to reflect problem solving abilities"*

Otherwise, instructions and the rest of the procedures and dependent measures were identical to Experiments 1 and 2.

3.3.2 Experiment 3 - Results

24 participants (16 female, 8 male) from the Open University were tested. Data was discarded if participants in the Success group performed below a 60% success rate. For the Failure group, the criterion was set at a success rate below 40%. No data needed to be discarded for this study.

All data was analysed using between group t-tests except for the WellY vs WellO analysis which was analysed with a within-subjects t-test.

Table 4

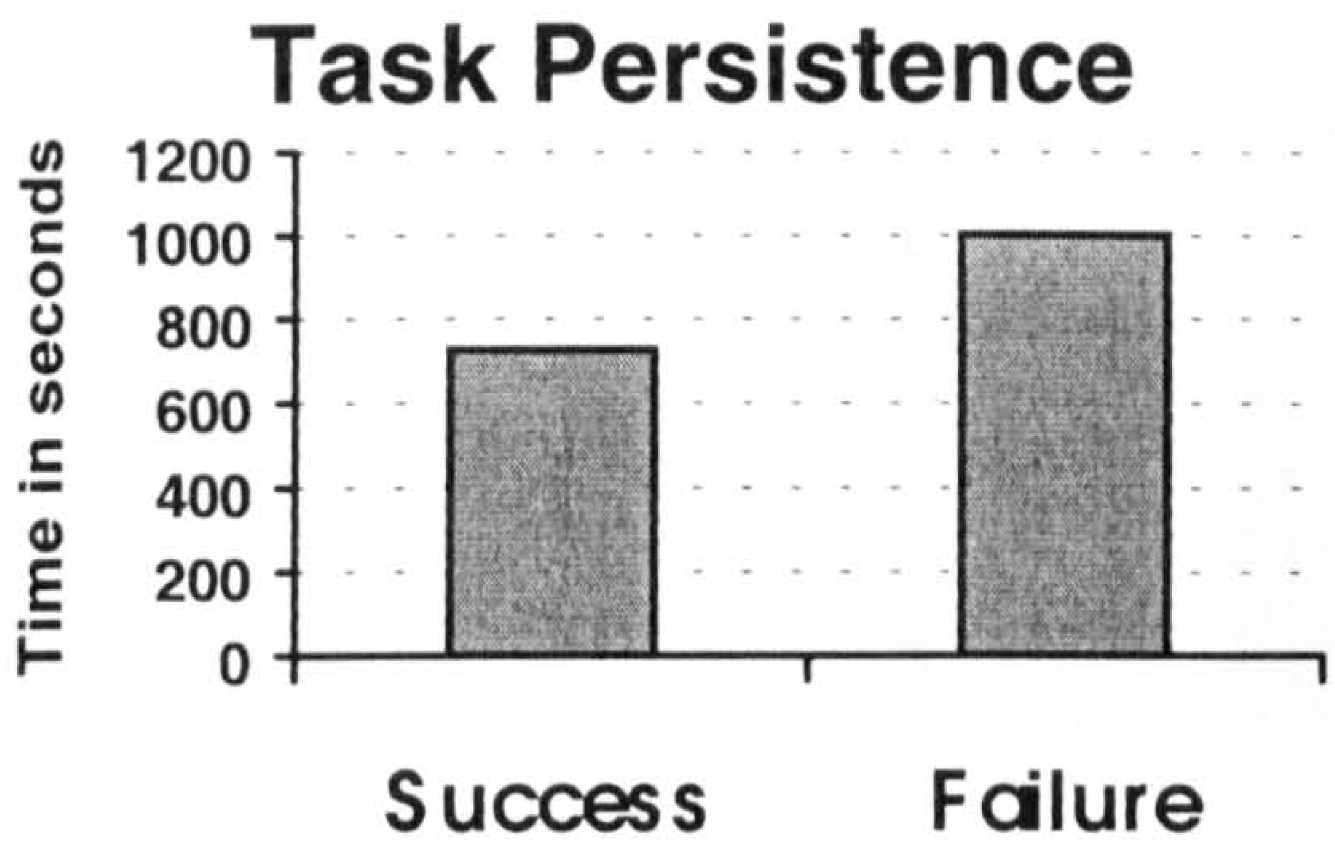
Means and standard deviations (in italics) for all dependent measures for stockmarket Experiment 3 (n=24).

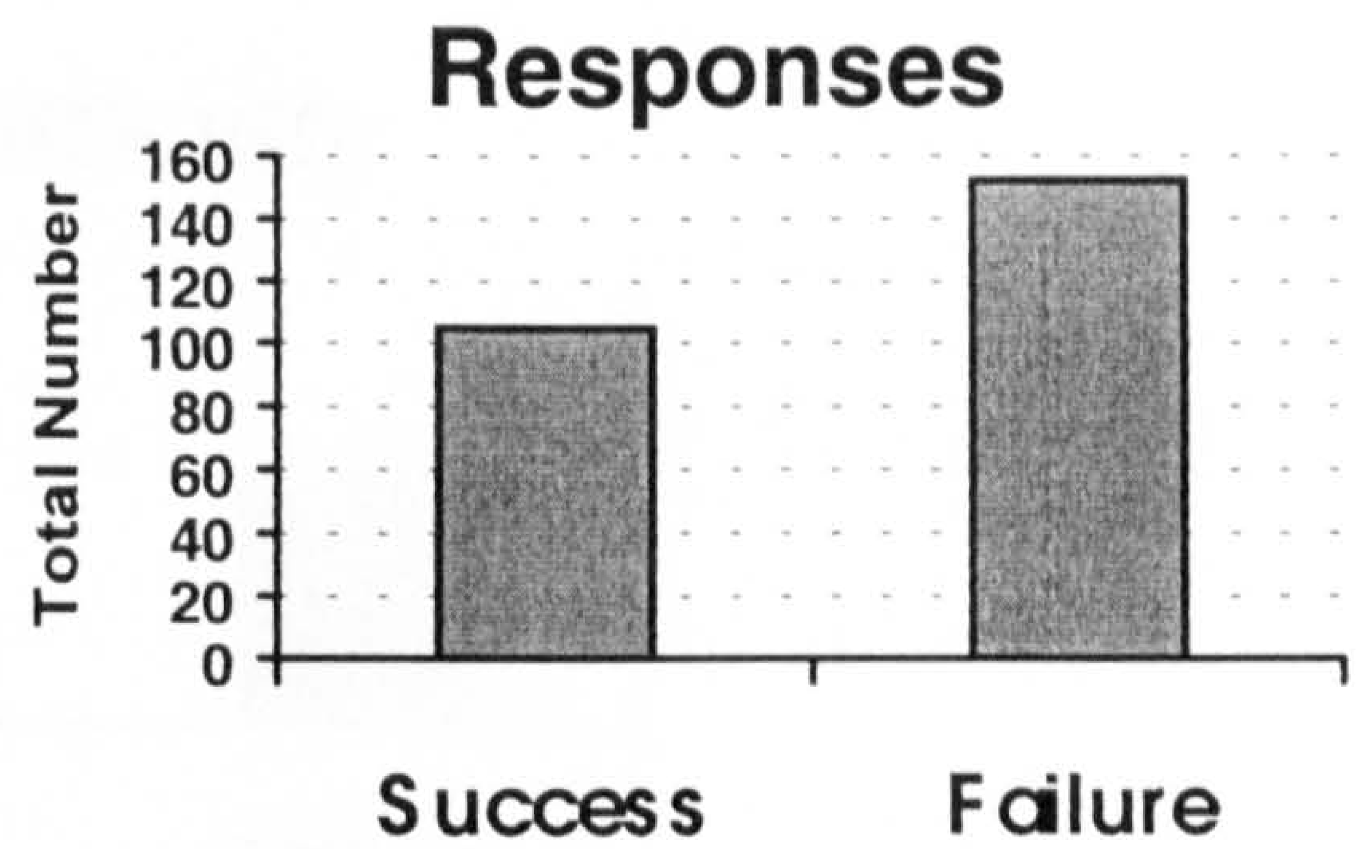
	90% success		**30% failure**	
	Mean	***s.d.***	**mean**	***s.d.***
Time spent on task	727	*235*	1002	*394*
Number of responses	104.9	*42.5*	152.3	*67.7*
Interest	3.7**	*1.44*	2.1**	*1.00*
Enjoyment	4.0***	*1.35*	2.0***	*0.95*
Task Difficulty	4.4*	*2.92*	2.9*	*1.66*
WellY	3.6**	*0.1*	2.2**	*0.89*
WellO	4.4*	*1.31*	3.0*	*1.01*

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 4 above shows the means and standard deviations for the dependent measures of task persistence and the average ratings for the questionnaire items. Significant differences were observed whereby the Success groups rated the questionnaire items more positively. In terms of persistence, the Failure group persisted for longer at the task but the difference was not significant t (22) = -2.08, p = 0.53. In terms of responses, the Failure group made more predictions, but again the difference was not significant, t (22) = -2.06, p = 0.54. The results for each dependent measure are dealt with individually in the following sections.

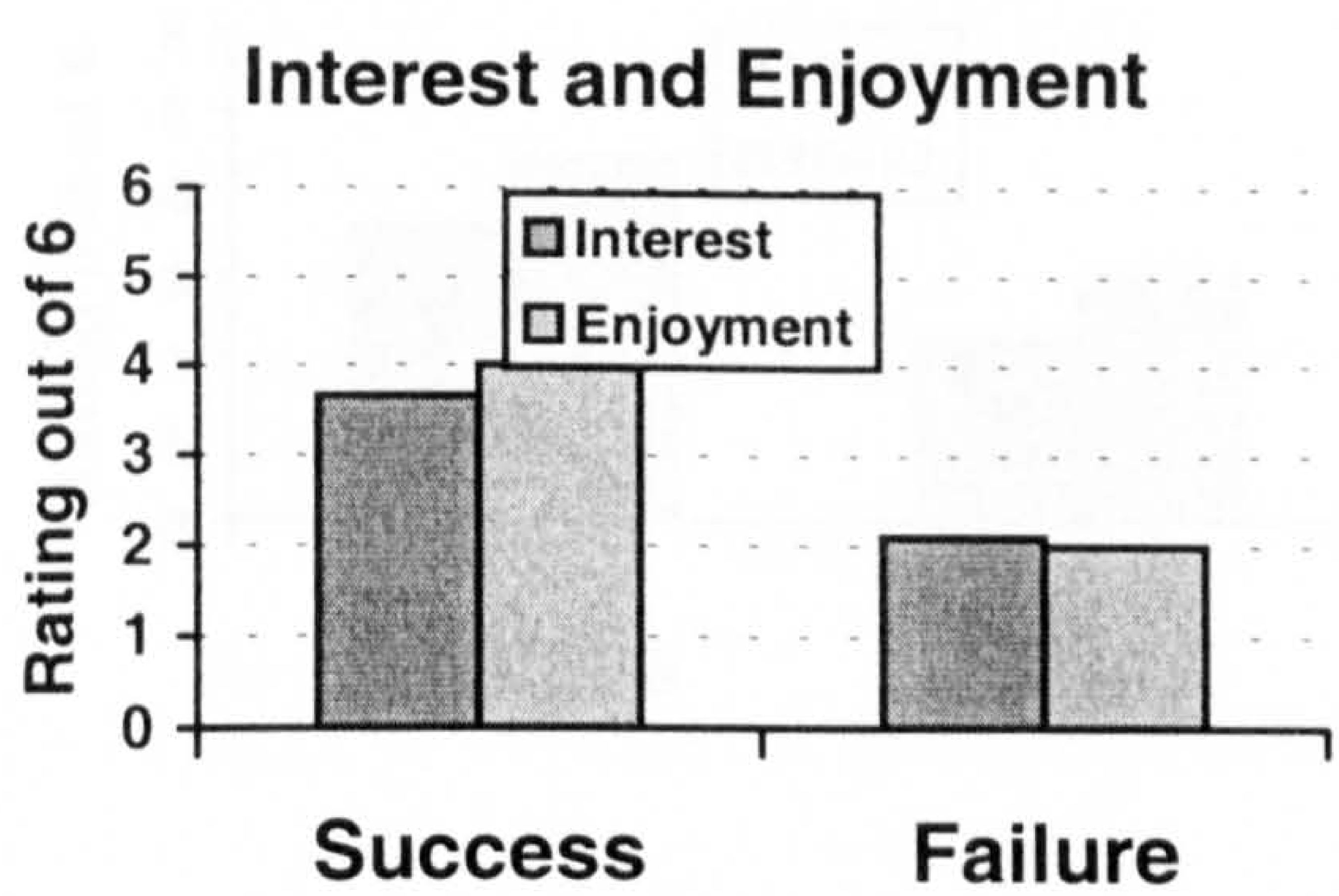
3.3.2.1 Experiment 3 - Task Persistence + Number of Responses





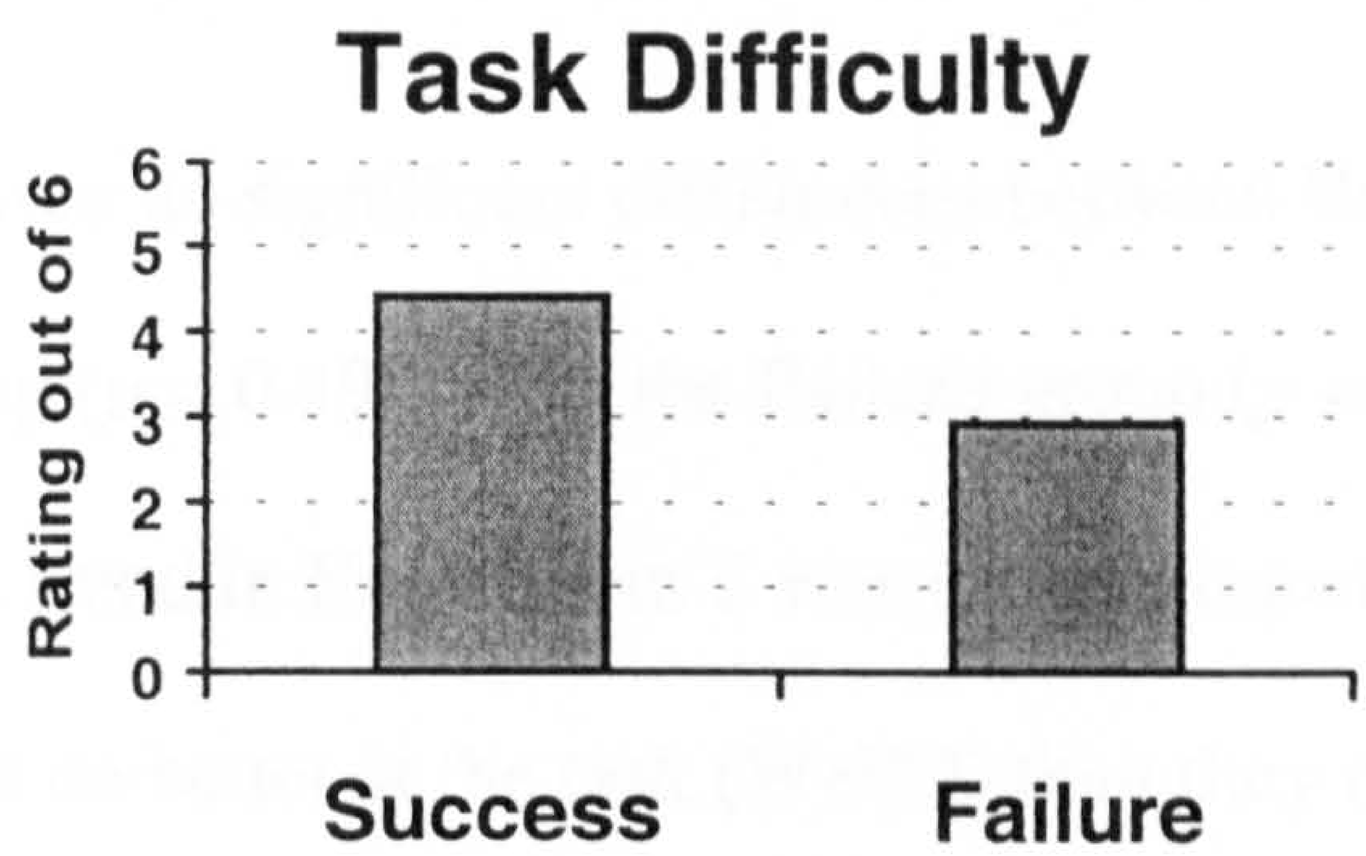
There was a (virtually) significant difference between the two groups for task persistence and responses ($p = 0.053$ for Persistence; $p = 0.054$ for Responses). In line with Experiments 1 and 2, participants in the Failure group persisted for longer with the task compared to those in the Success group.

3.3.2.2 Experiment 3 - Interest and Enjoyment



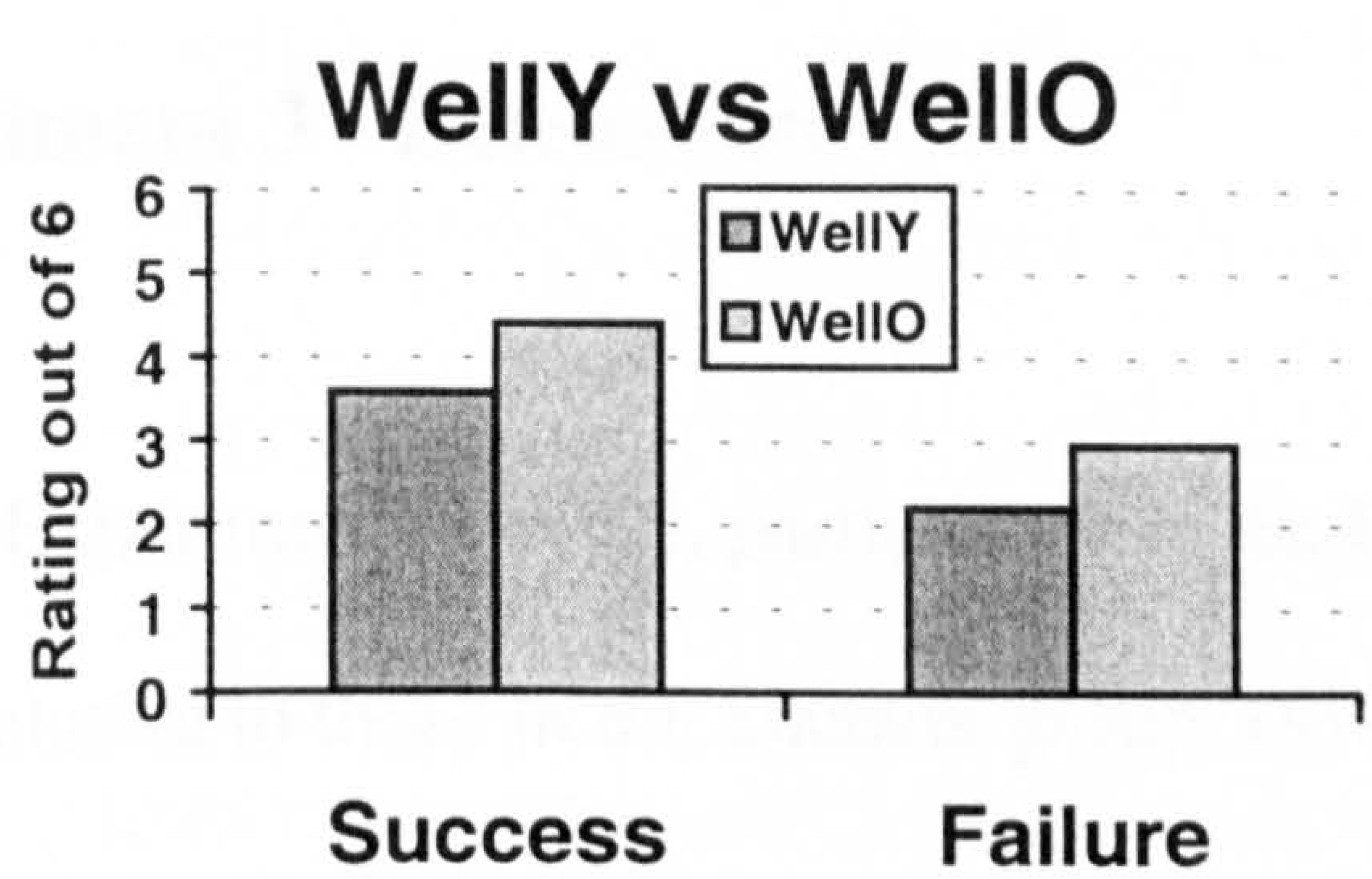
In this experiment, participants in the Success group experienced the task as both more interesting, $t\ (22) = 3.14,\ p < .01$ and enjoyable, $t\ (22) = 4.20,\ p < .001$ compared to those in the Failure group.

3.3.2.3 Experiment 3 - Task Difficulty



Participants in the Success group reported the task significantly easier compared to the Failure group, t (22) = *2.45*, $p < .05$.

3.3.2.3 Experiment 3 - WellY vs. WellO



<u>WellY</u> In a similar fashion to Experiment 1, participants in the Success group rated their performance (WellY) higher than those in the Failure group, t (22) = *3.56*, $p < 0.001$.

<u>WellO</u> There was a significant difference between the Success and Failure groups for how well they thought others (WellO) would do at the task, t (22) = 2.45, $p <$

.05. Participants in the Success group thought that others would do significantly better at the task compared to those in the Failure group.

WYWO There were no significant differences between WellY and WellO for either the Success group ($p = 0.49$) or for the Failure group ($p = 0.07$). However, as with Experiment 1, the trend in Experiment 3 was for participants in the Failure group to believe others would do better at the task (WellO) than they did (WellY).

3.3.2.5 Experiment 3 – Mood and Activity

In a similar fashion to both Experiments 1 and 2, there were no significant correlations between time spent on the task and reported interest and enjoyment.

3.3.3 Experiment 3 - Discussion

In a similar fashion to Experiments 1 and 2, participants in the Failure group persisted for longer at the task relative to those in the Success group. However, although they persisted for longer, participants in the Failure group experienced the task as less interesting and enjoyable. This persistence/experience disassociation is the same phenomenon observed by Ryan (1982; see also Ryan, Koestner and Deci, 1991) when they presented tasks to participants with a performance-focus. So whilst Experiments 1 and 2 were presented with a mastery-focus, and the same task was presented with a performance-focus in Experiment 3, all three experiments resulted in the same patterns of persistence and subsequent ratings of interest and enjoyment. The suggestion therefore is that regardless of the way that the stockmarket task was presented, participants seemed to be approaching it with a performance-focus. This hypothesis will be discussed at greater length in chapter 4 when the findings from all the experiments in chapter 3 are brought to together. At this point, the evidence seems to support the suggestion made in the discussion section of Experiment 1 (see section 3.1.3) that the methodology used in that experiment may have inadvertently caused the task to be experienced as performance-focused.

A further point of interest concerns the relatively low levels of interest and enjoyment reported by the participants in Experiments 1, 2 and 3. At a methodological level, because neither the average mean interest nor enjoyment rating, for any experiment, has exceeded 4, it is questionable whether participants were strongly interested (or what might be labelled "intrinsically motivated") in this activity. The task was presented on the posters as a 'Prediction game'. The game itself was very simple and was generally not well received by those participating in it. During the post-experimental briefing sessions, participants reported that they expected more information during the game, or at least something more interactive. Unlike the Soma cube (see Deci, 1975, pp. 32-133) and NINA tasks (e.g. Ryan, 1982; Ryan, Mims and Koestner, 1983; Ryan, Koestner and Deci, 1991) traditionally used in intrinsic motivation experiments, the prediction task did not appear to be perceived as interesting. As such, it is questionable whether participants were <u>ever</u> interested by the task. The relevance of this issue may be important in assessing whether or not the stockmarket task is an appropriate one for producing a mastery-focus. If, as suggested, the differences in interest and enjoyment ratings between the groups emerged in Experiment 3 (and not in experiments 1 and 2) solely because performance-focus heightened concerns for doing well at the task, then it seems unlikely that when the stockmarket task is presented with a mastery-focus, differences in experiences between the two groups will be observed. This point regarding the interest (or nature of the experimental task) is taken up as an experimental design issue for Experiment 4.

Summary and Conclusions

The manipulation of the task instructions to include a performance goal led to the previous trends for all the dependent measures in Experiments 1 and 2 reaching significance (or virtually significant - e.g. persistence in Experiment 3, $p = .054$; responses $p = .053$). As such, Experiment 3 offered evidence that (relative) failure results in greater task persistence compared to success under performance-focus as well as mastery-focus conditions but that participants in the Failure condition report lower levels of interest and enjoyment. Moreover, given that the pattern of persistence and interest and enjoyment ratings were similar to those observed in Experiments 1 and 2, there is a strong possibility that the method of measuring persistence in Experiments 1 and 2 might have caused participants to become performance-focused. Thus, it may be that in Experiments 1 and 2, although efforts were made to present the task with a mastery-focus, the fact that participants were aware they needed to answer questions in a post-task interview might have caused them to become overly concerned about performing well.

3.4 Experiment 4 – Initial and free-choice persistence responses to success and failure

In Experiments 1-3, participants persisted longer with the stockmarket task when they performed poorly compared to when they performed well. This pattern of persistence occurred regardless of whether the task was presented with a mastery or performance-focus. One possible explanation for this was that the method of measuring persistence resulted in participants becoming concerned about performing well, leading them to approach the task with a performance-focus, even when efforts were made to present the task with a mastery-focus.

One way of testing whether it was the method of measuring persistence that caused participants to persist as they did would be to allow them two opportunities to engage with the task, an unlimited persistence phase and a free-choice persistence phase. If it was the form of persistence that was causing participants to behave as they did, then they should persist in the same way in both the initial and free-choice persistence phases.

Experiment 4 therefore set out to extend the findings from Experiment 1-3 to investigate how participants would persist with the task in the different persistence phases. The paradigm employed in Experiment 4 was one where participants, in a similar fashion to experiments 1-3, were given a task under mastery conditions and told to persist with the task for as long they wanted. When they indicated they had finished the task, a free-choice persistence session, similar to the procedure outlined by Deci (1975, pp. 132-138) was employed.

In the conclusion section for Experiment 2, it was suggested that the stockmarket task might have reached a ceiling in terms of participants' ratings of interest and enjoyment. It was noted that consideration should be given to changing the task in a subsequent experiment because it was felt that whilst the rating scale was a 6-point one, the effective ceiling, in terms of interest and enjoyment, was only about 4. This meant that even when participants did very well at the tasks, they still only rated it at a maximum rating of 3.2 (Experiment 2). Indeed, the highest interest and enjoyment rating was observed when participants actually did poorly at the task (M = 3.9, Experiment 2, Failure condition) or when the task was specifically performance-focused (M = 4.0, Success condition). This meant that when comparing the mean interest and enjoyment ratings between the groups, unless participants disliked the task a lot, it was difficult to observe any statistically significant differences between the groups. Thus, it was felt that using a more interesting task might help to create a higher ceiling for the interest and enjoyment ratings

Unlike intrinsic motivation experiments (e.g. Manderlink and Harackiewicz, 1984; Epstein and Harackiewicz, 1992; Blanck, Reis and Jackson, 1984), the pilot studies for the stockmarket task were aimed at testing whether or not the software program worked and the levels at which participants experienced success and failure, rather than testing whether the task was an interesting one (it did not seem to be an issue at the time). In fact (and in retrospect), notes taken during the pilot studies, and during the actual experiments, revealed that participants who completed the stock market task frequently commented on how limited the task was. For example, participants reported that they had expected the task to be more interactive (e.g., they envisaged having the ability to buy and sell shares in an effort to influence the market) and expected more information (e.g. prices of shares; categories of shares). The stockmarket task was designed with an emphasis on controlling the levels of success that participants experienced. Making the task more interactive would have made it more difficult to ensure that participants experienced similar levels of success and failure. The alternatives that were considered prior to selecting the stockmarket prediction task were discounted because they failed to adequately control for levels of success (see section 2.1). However, by limiting the scope of the experimental task in this way, the interest of the task might have been also reduced.

3.4.1 Tasks used in Free-choice persistence and studies where persistence has been used as a dependent measure

Free-choice persistence

Free-choice persistence researchers have traditionally used spatial tasks in their studies of intrinsic motivation. For example, Deci, (1972) used the Soma cube task. This puzzle consists of seven pieces of plastic that can be manipulated to make different shapes. More recently, researchers have tended to use the NINA task where participants are required to find as many instances as possible of the word NINA in a cartoon drawing (e.g. Ryan, 1982; Ryan, Mims and Koestner, 1983; Plant and Ryan, 1985; Koestner, Zuckerman and Olsson, 1990; Ryan, Koestner and Deci, 1991; Koestner, Bieneri and Zuckerman, 1992).

Another category of task frequently employed by free-choice persistence researchers has been word games. For example, Epstein and Harackiewicz (1992) used a game called 'Boggle'. In this game, letter cubes are jumbled up and arranged randomly in a 6 x 6 matrix. The participants were given a pre-set time to identify as many three-or- four-letter words by linking letters horizontally, vertically or diagonally. Harackiewicz, Abrahams and Wageman (1987) report using a virtually identical task, except that the letters in the matrix were pre-set by the experimenters, presumably to control for difficulty. Blanck, Reis and Jackson (1984) also used a task which they presented as a 'spill and spell' task. Here participants were given 30 letter cubes and were required to make as many words out of them as they could, again against the clock. Blanck et al (1984) report that they successfully pilot-tested their task for intrinsic interest.

Another form of word game has been an anagram task (Baumeister and Tice; 1985). However, the usage of anagram tasks in studies which employ a free-choice persistence paradigm appears to be limited.

Task used for other forms of persistence

Feather has used a variety of tasks in his studies where persistence has been the dependent measure. For example, in one study participants were given a card-sorting task where they had to sort a pack of cards into categories within a certain amount of time (Feather, 1959). In another, participants were given a task that involved having to trace over a particular geometric shape without taking the pencil off the paper or retracing over a line (Feather, 1961). Feather's later studies used anagrams as the experimental task (Feather, 1963b; 1965; 1968 and 1969). However, these studies did not directly test persistence. Instead, after manipulating success and failure outcomes, participants were asked to indicate their expectations and confidence of future success (e.g. Feather, 1963b; 1965; 1968), or were asked to make attributions about their performance (Feather, 1969). These predictions were thought to be indicators of future persistence.

More recently, Sandelands, Brockner and Glynn (1988) have used six-letter anagrams, some of which they manipulated so that they were unsolvable. Persistence was measured as the time spent on these unsolvable anagrams. In a study by Eisenberger, Kuhlman and Cotterell (1992), participants completed two experimental tasks. In the first task, they either performed well or poorly. In the second task, they were given a series of unsolvable anagrams and their persistence with these anagrams was used as the dependent measure for how they responded to their either good or poor performance on the first task.

For the studies that have used anagrams, it is interesting to note that no rationale for why anagrams were used is given in any of these studies. Also, none of these studies pre-tested the intrinsic interest of the task.

3.4.2 Selection of task for Experiment 4

Despite the history of the successful employment of the Soma task and the NINA task in free-choice studies, it was felt that the anagram tasks might also be an interesting task for participants to engage in. For example, anagrams consistently appear in various British newspapers and have also been used on the T.V. game show "Countdown", a show particularly popular with students. Indeed, because of this latter point, and because some participants might be poor at solving anagrams, a second task was also included. This task was the numbers task that is used in this show. In this task, participants are given six numbers plus a target number. They are then required to multiply, subtract, divide, or add any or all of the six numbers by one another to achieve a numbers as close as possible to the target number.

Two sets of anagrams were prepared, an easy and a hard set. Pilot testing was conducted by giving 10 volunteers a series of number and letter puzzles and recording the time they took to answer each problem. The anagrams which were solved the quickest (less than 30 seconds) were used for the Easy group, and the anagrams which took the longest to solve (over 4 minutes) were used for the Hard group. The anagrams seemed to fall into two distinct categories. The easy anagrams were generally between 4-6 letters long, whilst the hard anagrams were generally between 7-12 letters. For the numbers tasks, the numbers in the task which were included in the Easy group took between 1 and 3 minutes to solve whereas for the Hard group, solution times were over 6 minutes and were generally not solved at all. Appendix 1 lists all the problems in the two conditions (including the solutions).

Additionally, all pilot-tested volunteers were interviewed to gauge how interesting/enjoyable they thought the anagram tasks were. No formal analysis was completed on participants' responses, but several of the volunteers had also been volunteers for other potential tasks in this thesis, and it was encouraging to note the general agreement among these participants that the anagram and number task were far more interesting than the stockmarket task.

3.4.3 Fun Question

In addition to the interest and enjoyment questions used in Experiments 1-3, a third question was included to examine whether participants were enjoying their experiences. This question asked participants "How much Fun would you say that task was?". This question has been validated as a correlate of interest and enjoyment and has been used by researchers to assess participants' levels of intrinsic motivation (e.g., Grolnick and Ryan, 1985; Plant and Ryan, 1985; Koestner, Zuckerman and Koestner, 1987, Harackiewicz and Elliot, 1983; Elliot and Harackiewicz, 1994; 1996; Elliot and Church, 1997). The purpose of including this question was twofold. Firstly, it increased the suite of dependent measures that could be used to help determine whether participants were experiencing the task positively. Secondly, the mastery-focus induction in Experiment 4 was presented in terms of asking participants to report how much fun they thought the task was. It may have therefore been surprising to them if, after completing the task, they were not asked anything about how much fun they had.

3.4.4 Attribution and "stop" questions

In the first 3 experiments, participants often seemed to be attributing success and failure to themselves (internal attribution) or the computer (external attribution). For example, in Experiment 2, it was speculated that participants in the Failure group rated the stockmarket task as interesting because they had worked out what the computer was doing and this lead them to attribute their failure to external factors.

In Experiment 4, participants were asked a series of additional questions in an attempt to provide a more objective record of the reasons for their behaviour during the task. To check for attribution, participants were asked whether they thought their performance at the anagrams was due to their own ability (internal attribution) or to the inherent difficulty of the task (external attribution). In addition, participants were asked a series of questions about their reasons for desisting at the task during the initial persistence phase of the experiment.

3.4.5 Methods

Overview In a similar fashion to Experiments 1-3, this was a two-condition experiment (this time labelled Hard vs. Easy). The experiment ran in two stages; the first measured initial persistence at the task and the second, free-choice persistence. Participants were randomly assigned to one of the two conditions and were told that their task was simply to comment on how much fun they thought a problem-solving task was (i.e. the mastery orientation manipulation). The task involved solving a series of anagrams and number problems, similar to those tackled by contestants in the T.V. game show "Countdown"

In the first phase of the experiment, participants were taken into a room and asked to work out a series of problems that were displayed to them on a computer screen. Participants were asked to just keep going until they had sufficient information to comment on how much fun they though the task had been. When participants indicated that they were ready to answer some questions about the task, they were asked to wait in the testing room while the experimenter analysed their results. While they were waiting, the experimenter casually mentioned to the participant that they were free to try a few more problems. The experimenter then left the room for 6 minutes. The free-choice measure of persistence was taken during this 6-minute period.

When the experimenter returned, participants were given a short questionnaire that asked them about their experiences of the task.

Participants 31 participants (20 female, 11 male) were recruited via the Stirling University Psychology department's subject panel. Participants volunteered as part of a first year departmental course requirement.

Apparatus The task stimuli were prepared using Superlab software and were presented to participants on a Mackintosh PC (performa 6200).

The testing room was a small windowless room containing just a table, a chair, the computer, and a set of task instructions.

Procedure Participants were recruited via a poster requesting volunteers for a words and numbers game. This game was described as similar to a popular British television game show called "Countdown". The posters presented the experiment with the title "Countdown for Fun".

When participants arrived for their appointment, they were taken to the testing room. They were randomly assigned to either a Hard or Easy problem-solving condition. They were given a sheet on which they were to record their answers (see Appendix 3)[4]. The experimenter emphasised that the aim was not to get answers right or wrong, but simply to comment on how much fun they thought these types of tasks were. Participants were asked to read a set of instructions which were pinned to both the desk and the wall. These instructions read as follows:

<u>Countdown for Fun Instructions</u>

Thank you for agreeing to take part in this experiment.

All I want you to do is to try and solve the tasks until you have some idea of how much Fun you think they are e.g. how enjoyable, interesting they are.

Please

- Use the answer sheets to show your working out.

- Do NOT move onto the next problem until you have either solved the previous one or given up

It is not important that you get the answers right, I just want your opinions of whether or not the tasks are fun and why?

You do not need to try to solve ALL of the problems on the program as there are just too many and you will be here for a few hours if you do.*

When you think you are in a position to give your views, please come and see me next door.

If you have any questions, please see me now, otherwise, you can start when you are ready.

[4]There were several reasons why participants were asked to record their answers on an answer sheet rather than just typing them directly into the computer. Firstly, the strategy allowed the experimenter to make a plausible excuse for keeping the participant waiting while the experimenter went away to "analyse" the responses. Also, the sheet allowed for further analysis to be conducted on the patterns of success and failure for each participant.

* The first three participants tested tried to solve all the problems presented to them. As a result of this, the experimental instructions were slightly amended to emphasise that participants were *not* to try to solve all of the problems as there was ".. just too many and you will be here for a few hours if you do".

The experimenter then left the room. In a similar fashion to the stockmarket experiments, the experimenter restricted contact with the participant in an effort to minimise experimenter effects.

The experimental task The task comprised of a series of anagrams and numbers puzzles. The anagrams varied from 4-12 letters. All problems were solvable. When participants turned to the computer, the first screen read as follows:

> Thank you for agreeing to take part in this experiment.
>
> All we want you to do is to try and solve the puzzles and tell us how much "FUN" you think they are i.e., how much do you enjoy doing them.
> It is not important that you get the answers right or wrong – we are primarily interested in how much you enjoy doing the tasks.
>
> Press the spacebar for more information

The next screen read:

> **The tasks**
>
> You will be presented with either an ANAGRAM or a NUMBERS task.
>
> The numbers task is similar to the ones on the T.V. program "COUNTDOWN".
>
> Press the spacebar to see an example.

Before each task was presented, a "preparatory" screen would appear announcing what type of task was next e.g. "O.K., now for a numbers game". The purpose of the preparatory screen was to allow participants a rest between tasks; this rest period was not counted in persistence time (see dependent measures section later).

The number tasks involved participants having to use a series of single numbers in order to calculate a larger target number. For example,

Target = 554 Number series = 1, 2, 2, 8, 10, 25.

In the number series, participants were allowed to add, subtract, divide, or multiply any number from another number (or the sum of any combination of the other numbers), but were only allowed to use any number once. Participants did not have to use all the numbers, just some combination of them to achieve the target sum.

A possible solution to the above example is as follows (numbers in the series in ***bold italics***)

10 + ***1*** = 11; 11 x ***2*** = 22; 22 x ***25*** = 550; ***8/2*** = 4; 550 + 4 = <u>554</u>

When participants pressed the "enter" key for the first time, they were presented with one practice anagram e.g. ALTBE (ans = TABLE) and one practice number task, together with a message telling them that these examples were to give them a feel for the tasks.

Once participants had completed the practice problems, they were presented with the instructions for the test phase. These instructions were as follows:

"OK, now for the test items.

You should try to work through the tasks as quickly as you can.

If you cannot solve a problem, just move onto the next one by pressing the space bar.

If you are ready to start, press the space bar to continue"

When the participant pressed the space bar to start the test items, this started the clock for the initial persistence measure. The participant then continued with the puzzles until they felt that they had completed sufficient anagrams. Once they had finished solving the puzzles, the participant left the room and met with the experimenter who was waiting outside.

The experimenter then advised the participant that he would need to ".. *go away for a few minutes to analyse the results and get a questionnaire for them to fill out*". The participant was instructed to stay in the room (because there were other participants doing experiments in the same testing area), but that they were free to try some more puzzles if they wished. This phase of the experiment was the free-choice persistence measure.

The experimenter returned to the room after exactly 6 minutes with a questionnaire. If the participant had attempted any more anagrams or numbers during the free-choice period, the first key the participant had pressed re-started the timer.

The participant was then given a questionnaire asking them to rate on a 6-point Likert scale (see also Appendix 3):

- " How much FUN would you say that task was?" (6=Fun; 1= Not Fun)
- " How much did you enjoy solving the problems?" (6 = A lot; 1= Not at all)
- " How interesting was the task as a whole?" (6 = Very interesting; 1= Uninteresting)
- " On the whole, how difficult was the task?" (6 = Easy; 1 = Difficult)
- " How well do you think you did at the task?" (6 = Well; 1= Badly)
- " How well do you think others would do at that task?" (6 = Well; 1= Badly)
- "Think about how you did. On the whole, was this down to your own ability or to the inherent difficulty/easiness of the problems you had to solve?" (6= Yourself; 1= The problems)

To obtain information on why participants stopped persisting at the task, participants were asked to rate the following statements as TRUE (6)/UNTRUE(1) on the same 6-point Likert scale.

- Stop1 You had enough information to be able to comment on how much fun the task was
- Stop2 The tasks were too EASY
- Stop3 The tasks were too HARD
- Stop4 You got bored
- Stop5 You felt you had spent enough time on the tasks

Participants were then asked if they had any questions concerning the experiment; they were fully debriefed as to the nature and purpose of the study and thanked for their participation. Again, this debrief session was particularly important for participants who attempted the hard problems in order to assure them that they had done well at the task.

Dependent measures

There were three dependent measures:

- Initial task persistence - The time from when the participant pressed the key to initiate the test items to the time when the experimenter entered the room to disengage the timer, minus the time the preparatory screen was displayed.
- Free-choice persistence - The time from when the participant pressed a key during the free-choice period to the time when the last problem was selected, minus the time the preparatory screen was displayed.
- Answers to the questionnaire items

For the initial persistence measure, the time between the problem being selected and the time to the next preparatory screen was taken to be the period when participants were persisting with the task. The total time for all the problems attempted were added up. The times when the "preparatory" screen was displayed were excluded because this was time when participants were resting between problems rather than working on them.

For the free-choice persistence measure, if a problem was on the screen when the experimenter came into the room after the participant had indicated that they had finished the task, the experimenter would press the space bar three times quickly in succession. This would enable the experimenter to recognise on the printout when session one finished and when session two started.

Calculating the total free-choice persistence time was not straight-forward. This was because simply recording the amount of time from the time any key was depressed in session two to the time that the last key was pressed in session two would not necessarily reflect the amount of time the participant might have been spending at a problem from the time of selecting the final task, and the experimenter returning to the room. For example, in the initial persistence session, participants chose when to stop persisting with the task so when the experimenter entered the room, it was reasonable to conclude that the participant had finished with the task that was on the screen. In the free-choice session, however, because it was the experimenter who determined the moment when they returned to the room, it was unclear whether the participant was engaging with the task at the exact moment when the experimenter arrived back in the room. To cater for this, the experimenter made a mental note of whether the participant was working on the task when the experimenter re-entered the room for the second time by noticing whether the participant was working on a problem or simply sitting back waiting for the experimenter to return. If the participant was working on the task, the time from the commencement of the last problem to the experimenter's arrival was included as a persistence measure. If the participant was not, the time to the completion last problem selected was taken as the free-choice persistence time[5].

---

[5]Traditionally, the amount of time spent engaging in the target activity has been obtained using either covert filming or by the experimenter (or second observer) observing the participant, usually through a one-way mirror (e.g. Koestner, Zuckerman and Koestner, 1987). However, the testing time window coincided with a departmental course that required use of the rooms where covert observation was possible.

3.4.6 Experiment 4 - Results

Of the 31 participants who completed the experiment, three sets of data (all female) had to be discarded because participants desisted because they had a prior engagement (e.g. another lecture to attend, a bus to catch).

Table 5

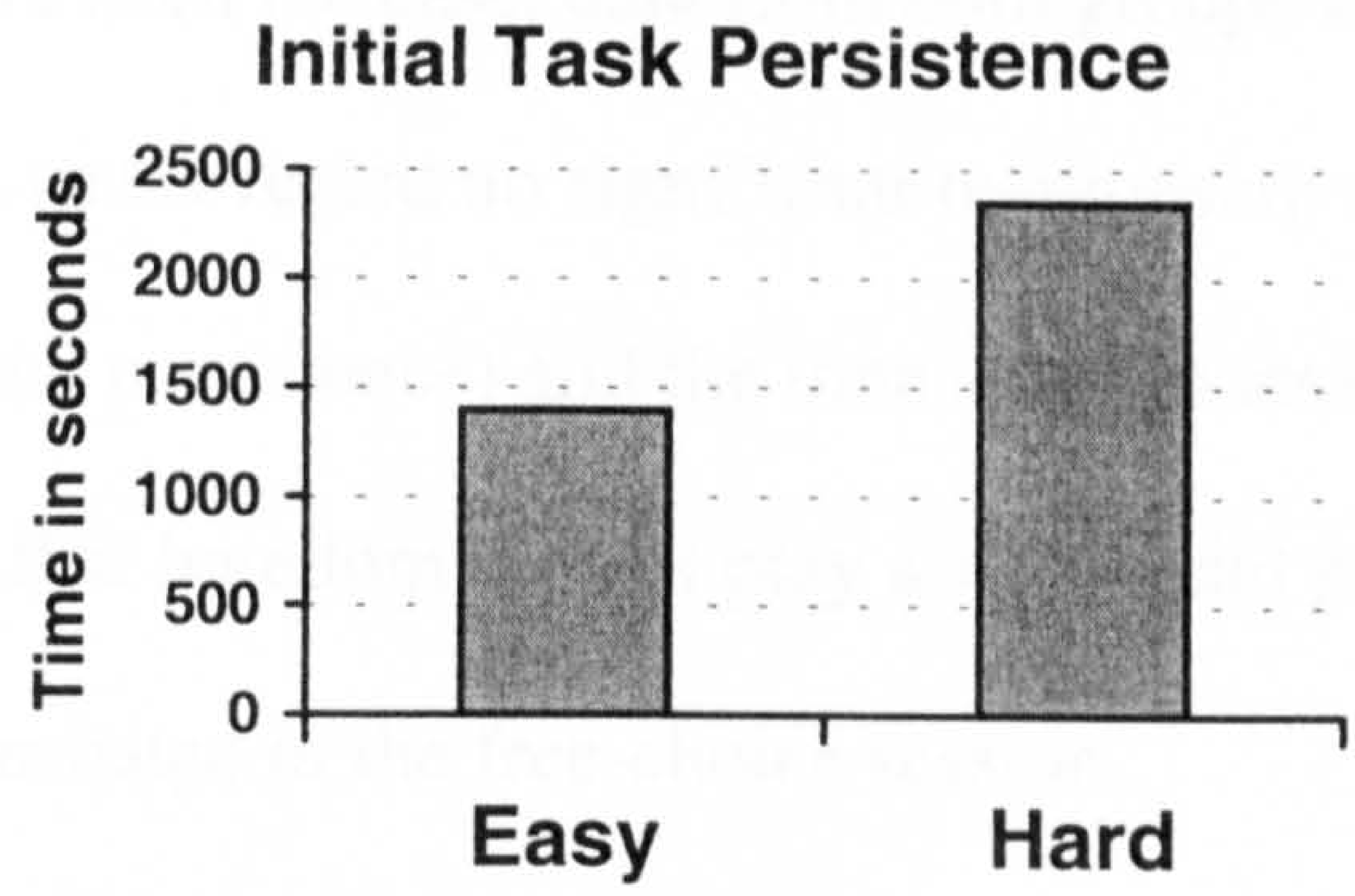
Means and standard deviations (in italics) for all dependent measures for anagrams and numbers experiment - Experiment 4 (n=28).

	Easy anagrams		Hard anagrams	
	mean	***s.d.***	**mean**	***s.d.***
Initial persistence	**1398****	***634***	**2347****	***752***
Problems solved	27.7	*13.0*	21.3	*9.4*
Free-choice persistence	217	*146*	120	*164*
Problems solved	**5.0***	***3.5***	**1.1***	***1.8***
Fun	3.9	*0.9*	3.1	*1.2*
Enjoy	4.3	*1.1*	3.5	*1.4*
Interest	3.8	*1.2*	3.2	*1.2*
Task Difficulty	3.4	*0.8*	2.4	*1.4*
WellY**	**4.1***	***1.2***	**2.1***	***1.1***
WellO***	4.4	*0.9*	4.1	*1.0*
Attribution	3.4	*1.1*	4.1	*1.3*
Stop 1	4.9	*1.4*	4.2	*1.5*
Stop 2	2.3	*1.3*	1.6	*0.8*
Stop 3	2.4	*1.7*	3.5	*1.3*
Stop 4	3.8	*1.2*	3.4	*1.8*
Stop 5	4.6	*1.1*	4.1	*1.3*

Measures in **bold** are significant   * = p < 0.05   ** = p < 0.01

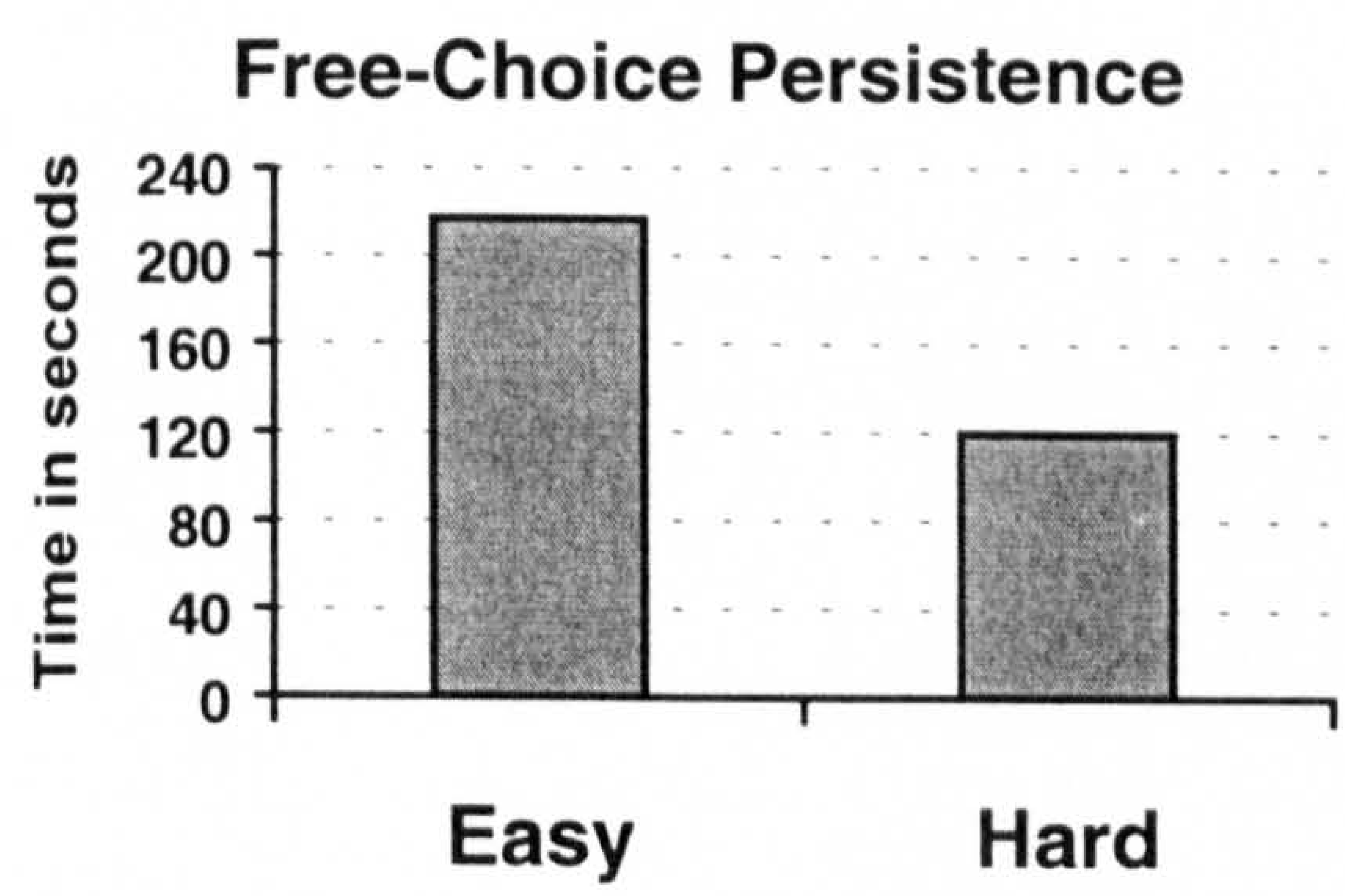
Table 5 above shows that there were significant differences between the Easy and Hard groups for the dependent measures of initial persistence, the number of problems attempted during the free-choice session, task difficulty and WellY.

3.4.6.1 Experiment 4 – Initial Task Persistence



The differences between the Easy and Hard groups for initial persistence were highly significant. Participants completing the Hard tasks persisted more than participants completing the Easy tasks, $t\ (26) = 3.61$, $p < 0.001$. This result follows the same trend in the previous stockmarket experiments.

3.4.6.2 Experiment 4 - Free-choice Task Persistence



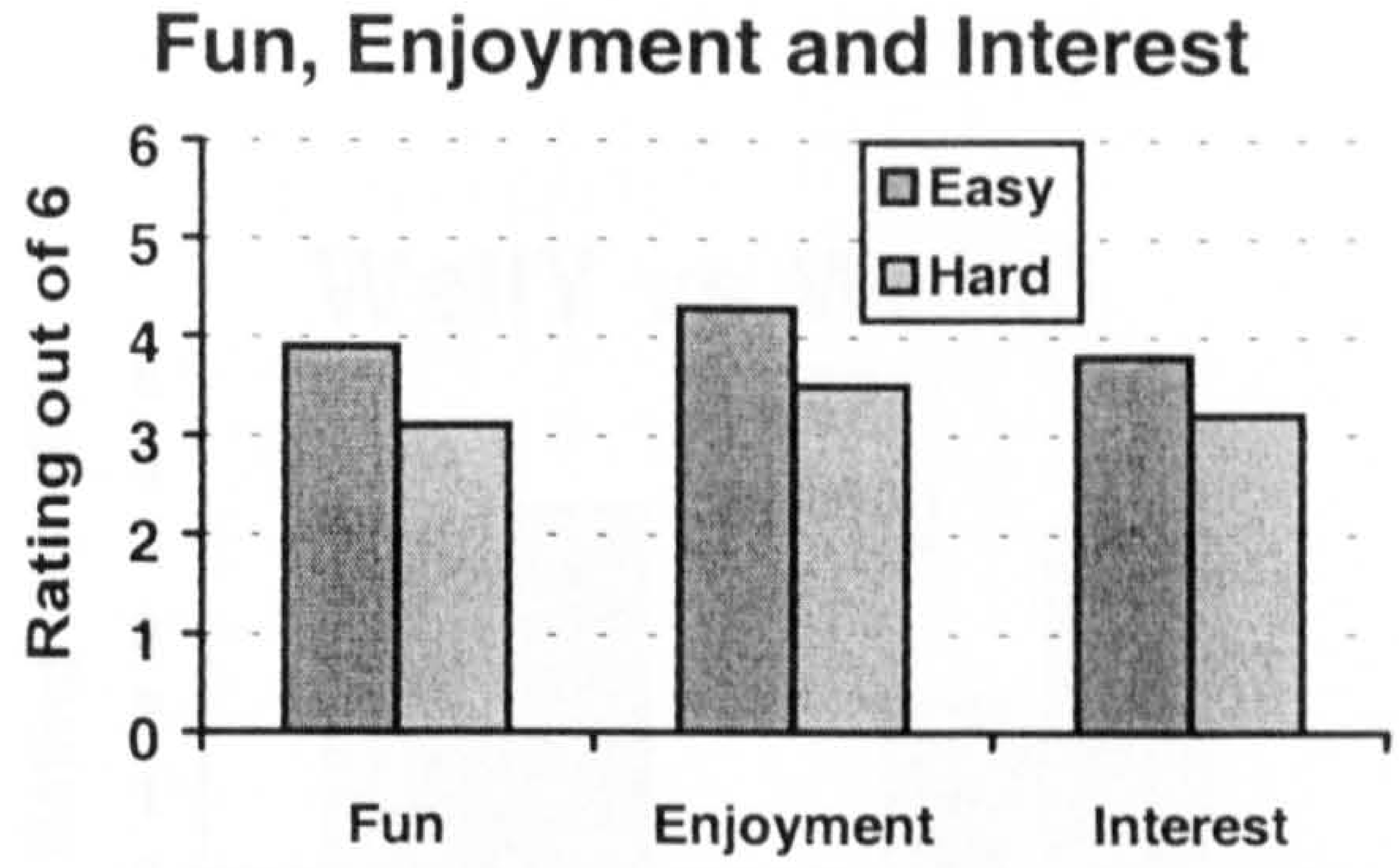
The opposite results were observed in the free-choice persistence session: Participants completing the Easy tasks persisted more compared to participants completing the Hard tasks. However, these differences did not reach significance.

One reason why participants in the Hard group persisted for less time in the free-choice session could have been boredom. Perhaps the longer participants persisted with the task in the first session, the less likely they were to engage with the task in the subsequent free-choice period because they had become bored with the task. To test whether this might have been the case, data from both groups were collapsed. A Pearson's r correlation test revealed no significant relationship ($r = .17$) between time spent in session 1 (initial persistence) and the time spent in session 2 (free-choice period). This suggests that boredom did not play a significant part in determining whether participants persisted in the free-choice session.

3.4.6.3 Experiment 4 - Number of Problems tackled in both sessions

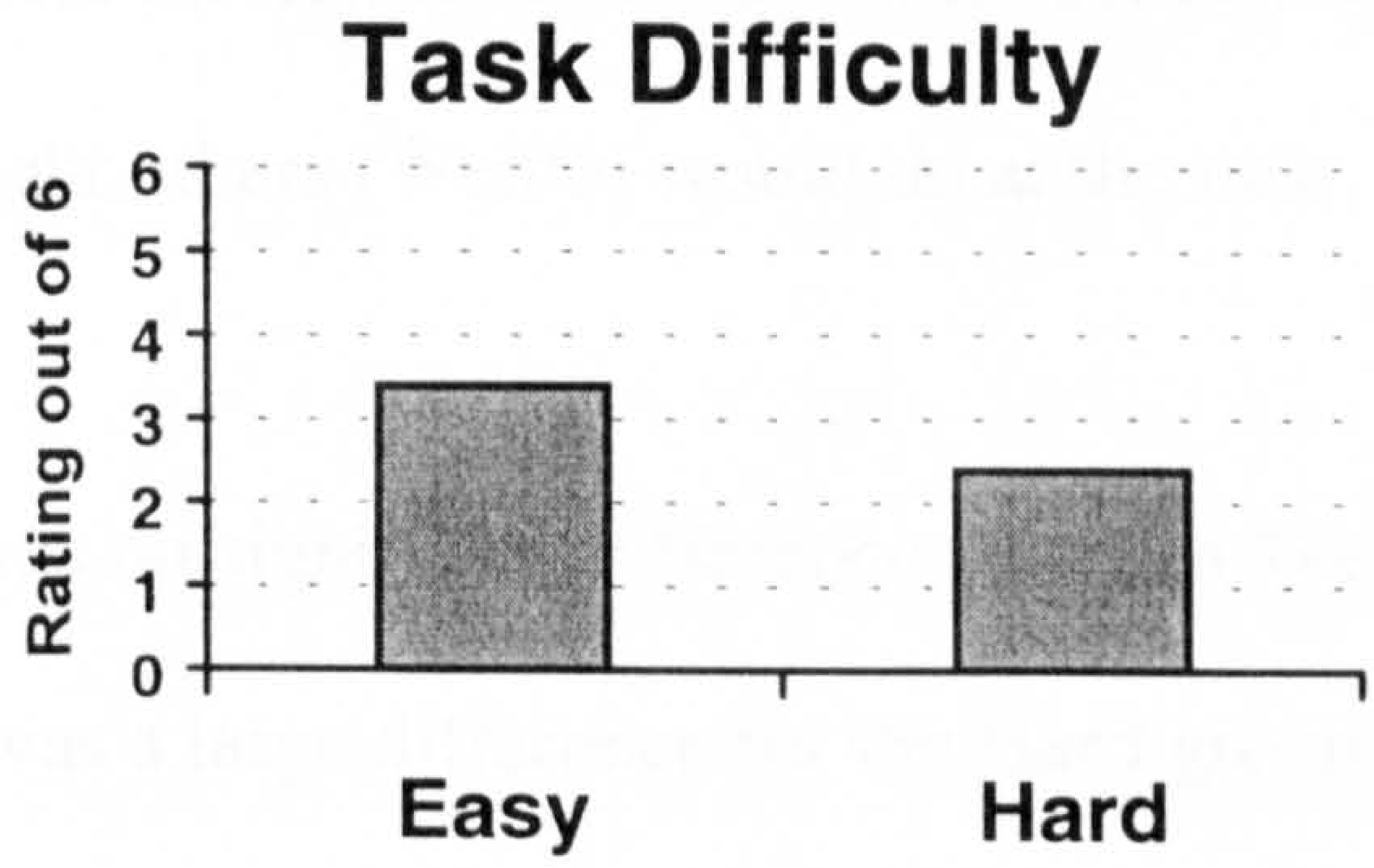
The number of problems attempted in the initial persistence session was not significantly different ($p = 0.15$) between the groups. However, during the free-choice session, participants in the Hard group (M = 1.1) attempted fewer problems than those in the Easy group (M = 5), $t(26) = -3.65$, $p < .001$.

3.4.6.4 Experiment 4 – Fun, Enjoyment and Interest



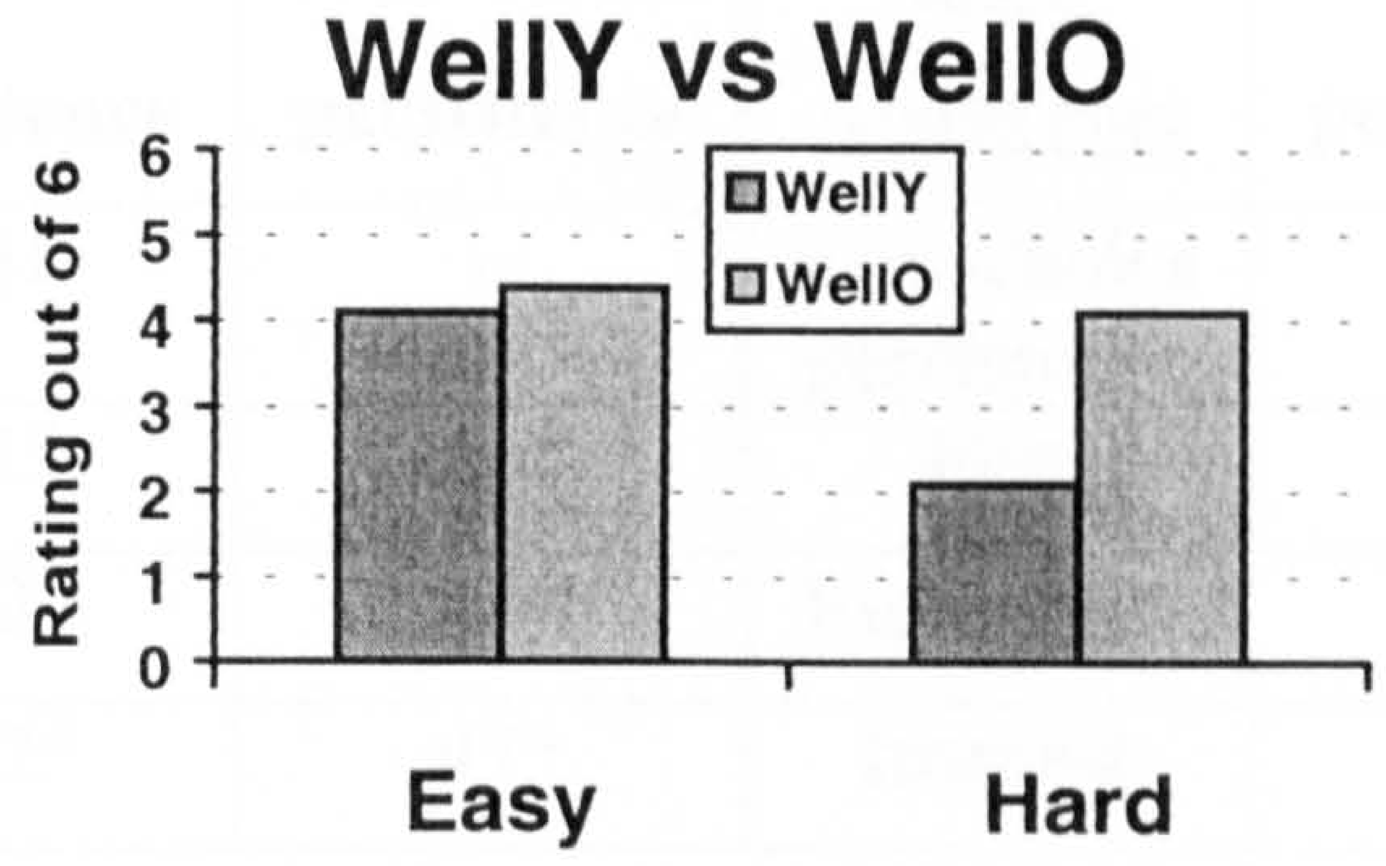
Although participants in the Easy group rated the task more positively for the dependent measures of fun, enjoyment and interest, none of these differences were significant.

3.4.6.5 Experiment 4 – Task Difficulty



Participants in the Easy group reported the task as significantly easier than those in the Hard group, t (26) = -2.34, $p < .05$.

3.4.6.6 Experiment 4 - WellY vs. WellO



<u>WellY</u> Participants in the Easy group rated their performance (WellY) higher than those in the Hard group, $t\ (26) = -4.53, p < .001$.

<u>WellO</u> There were no significant differences between the Easy and Hard groups for how well they thought others (WellO) would do at the task

<u>WYWO</u> There was no significant difference between WellY and WellO for the Easy group, but there was a large difference for the Hard group, $t\ (12) = -5.03, p < 0.001$. As in Experiment 1 and Experiment 3, participants in the Hard group thought others would do better than they had done at the task.

3.4.6.7 Experiment 4 – Mood and Activity

Unlike Experiments 1-3, there were significant relationships between time spent on the task and reported experiences. However, this was largely only true for the free-choice period; for the initial persistence period, the data mirrored the findings observed in Experiments 1-3.

Table 6: Correlation table for the dependent measures of Initial persistence, free-choice persistence, fun, enjoyment and interest for Experiment 4.

Hard Anagrams	**Initial persistence**	**Free-choice persistence**	**Easy Anagrams**	**Initial persistence**	**Free-choice persistence**
Free-choice persistence	- .31	--	**Free-choice persistence**	.25	--
Fun	- .25	. 56*	**Fun**	. 18	. 48*
Enjoyment	.03	.48*	**Enjoyment**	. 10	.43
Interest	- .52*	.47*	**Interest**	- .15	.54*

$* = p < 0.05$

Table 6 above shows that for the Hard group, there was a significant (negative) correlation between the amount of time they spent on the task in the initial persistence phase and the their ratings of how much they enjoyed the task. With regards the free-choice phase, there was a significant (positive) correlation between this measure and the fun, enjoyment and interest ratings. For the Easy group, there were no significant correlations between the amount of time they spent on the task in the initial persistence phase and the their task experiences, but there was for the free-choice phase. In this phase, there was a significant (positive) correlation between this measure and the fun and interest ratings.

3.4.6.8 Experiment 4 - Attribution question: Was the difficulty of the task due to your own ability or the inherent difficulty of the task?

Although the Hard group attributed their performance more to their own abilities than to the inherent difficulty of the task (M = 4.1 vs. Easy M = 3.4), this difference was not significant.

3.4.6.9 Experiment 4 - Stop Questions

There were no significant differences between the groups for any of the questions relating to why participants stopped persisting at the task.

3.4.7 Discussion

There were four main findings from Experiment 4

- On the initial problems, participants who tackled the Easy problems persisted less than those tackling the Hard problems.
- In the free-choice persistence session, there was a non-significant trend that participants in the Easy group persisted with the task more than those in the Hard task.
- There was a non-significant trend for participants to report the Easy task as more interesting and enjoyable.
- Participants in the Hard group rated their performance as likely to be worse than that of others doing the same task.

In a similar fashion to the stockmarket experiments, participants who attempted the Easy problems persisted less during the initial phase than participants who attempted the Hard problems. However, during the free-choice persistence phase of the experiment, although there were no significant differences between the groups, there was a trend for participants attempting the Hard problems to persist less (M = 120 secs) than participants completing the Easy (M = 217 secs) ones.

This pattern of persistence is particularly curious: Why should participants persist for so long during the initial persistence phase after performing poorly and then stop doing so in the free-choice persistence phase. Similarly, after doing well, why should participants persist for a relatively short time but then in the subsequent free-choice period start to persist with the task again. In considering these findings, there seems to be several possible explanations. However, it is probably premature to speculate on possible explanations at this point because significant differences between the groups were not observed during the free-choice period.

The lack of significant results is curious, especially given the large differences in means between the two groups. One explanation might be that due to the method of recording the free-choice persistence, may have inadvertently inflated free-choice persistence times. Recall that when the experimenter entered the room at the end of the 6-minute free-choice session, a mental note was made of whether the participant was engaging with a problem. If the participant was attempting a problem, then when the experimenter reviewed the print-out, the time from the selection of the last problem to the end of the free-choice period was included as "persistence" time. If the participant was <u>not</u> engaging with the task, this time was <u>not</u> included. This methodology could have caused problems in cases where a participant worked on one single task throughout the free-choice session. If the experimenter believed the participant to be engaging with a problem when they entered the room, the conclusion would be that they had engaged with the problem for the full six-minutes. However, the experimenter could have been wrong and the participant might have been only working on the task for only <u>some</u> of the time. This possibility was considered in advance and as a further check, during the debriefing session, participants were casually asked if they had engaged in the task while the experimenter had left the room. In all cases, the experimenter's judgement matched what participants reported. However, even this additional check would not fully cover the situation whereby the participant only participated with the final task for some of the time rather than the full 300 seconds. In short, the methodology used to calculate the free-choice persistence measure appears to be fundamentally flawed.

However, despite the suggestion that the timings for free-choice persistence were not 100% accurate, it was found that more participants in the Easy group engaged with the task during the free-choice period. A micro-analysis of the raw data shows that 11/12 participants in the Easy group attempted further problems during the free-choice period whilst only 6/12 did so in the Hard group. Also, of the 6 who persisted in the Hard group, 4 participants persisted for the full 6 minutes. This suggests that after performing well, participants were more prepared to attempt further problems. So whilst the calculated persistence times during the free-choice period might not be fully reliable, there seems to be good evidence that participants who solved the Easy problems were exhibiting some sort of desire or motivation to continue with the task. In retrospect, the method used to assess the free-choice persistence times might have been flawed resulting in a measure of free-choice persistence open to conjecture regarding the time that participants spent at the task.

Despite changing the task that participants were given, there were still no significant differences between the groups for the main dependent variables of fun, interest and enjoyment. As in Experiment 2, it might be that the perceived difficulty of the task helps to explain the lack of significant differences. To help explain the point further, it is perhaps useful to firstly examine the data for the task difficulty question.

Although there was a significant difference between the groups in how easy participants found the task, the mean rating for the Easy group (mean = 3.4) was still lower than expected. This rating was lower than the ratings given by participants in Experiments 2 and 3 of the stockmarket experiments (exp 2 mean = 4.0; exp 3 mean = 4.42). This is even more surprising given how well participants performed in this condition. The average success rate in the Easy group was 90.1% compared to 32.9% in the Hard group. The lowest individual success rate in the Easy group was 34%, with six participants scoring above 90% and five scoring between 80% and 90%[6]. This supplementary data suggests that participants who solved the Easy problems actually performed extremely well at the task they were set. It is therefore surprising they should have thought the task to be so difficult after performing so well at it.

One possible explanation for why participants did not rate the task as particularly Easy might lie in the inclusion of the numbers task. The Easy anagrams appeared to be experienced as very easy. Only three failures were noticed in all of the anagrams attempted, and the average time to solve an anagram in the Easy condition was 15.3 seconds. The numbers tasks, however, were only *relatively* easy, that is, relative to Hard tasks. Although the Easy number tasks were mostly solved, there were more failures compared to the Easy anagrams. In addition, the numbers tasks took longer to solve than the anagrams. If participants were succeeding quickly on an anagram task and then taking longer to solve the next problem, albeit correctly, this change in

---

[6]One participant's individual data could not be analysed as the participant inadvertently took their working sheets away with them. This participant rated that their performance as poor (rating = 2) and it is likely that they performed poorly at the task.

performance might have affected their perception of how Easy they thought the task was overall. There is some evidence to support this last statement. For example, when participants were rating the tasks, two participants rated the numbers and anagram tasks separately. In debriefing, three other participants commented that they preferred one task to the other.

Thus, the ratings in the Easy group might have been an average based on two different perceptions, one for the numbers task and one for the anagrams task. It is therefore plausible that the inclusion of the numbers task, intended to give individuals who were not good at anagrams a chance to succeed on another task, had the effect of making the task, on the whole, more difficult.

A further analysis of the pilot data reveals that persistence times for the numbers tasks were consistently higher than for the anagrams in both the Easy and Hard groups. Also, it was noted that in Experiment 4, participants in the Easy groups did not give up on the anagram task (and solved all of them), but they did give up on some of the numbers tasks. So although it appeared that the Easy group were solving far more problems and generally being more successful, this does not tell the whole tale about what intra-task experience they were having. It seems that despite solving many of the problems, they were still experiencing some level of failure.

This lack of complete (100%) success for participants in the Easy group might help to account for the lack of self-report rating differences between the Easy and Hard groups for the dependent measures of fun, interest and enjoyment. It might have been that the numbers task caused the task, as a whole, to be experienced as more difficult than if participants had just been given the anagram problems alone. Thus, at a methodological level, the numbers task seemed to undermine the success manipulations in the Easy problems condition.

In Experiment 4, participants were also asked whether they thought their performance was due to their personal ability (internal attribution) or the inherent difficulty of the problems (external attribution). The purpose of this question was to ascertain whether participants believed that the task they were given to do was soluble. For example, in Experiment 2, several participants in the Failure group reported that they thought that the computer determined their performance, not themselves. If these participants had been asked to attribute their behaviour to internal or external causes, it would have been likely that they would have rated their performance as externally caused. However, the results revealed that there were no significant differences between the groups. In one sense, this finding is encouraging because it suggests that participants (especially those who attempted the Hard numbers and anagrams), accepted that all the problems were soluble. Indeed, during the de-briefing sessions, unlike Experiment 2, participants in the Hard group never stated that they were suspicious that the problems might have been insoluble. However, if they really did believe that the problems were solvable, then the only causation for their behaviour should have been internal, that is, they should have attributed their performance to their own abilities and rated their performance as internally caused. However, a one sample t-test comparing attribution ratings against a hypothetical mean of 3.5 showed no significant differences for either group.

One possible explanation is that some of the participants who attempted the Hard problems did make an external attribution for their performance, but that this attribution was that the task was a difficult one. One possible way to check this would to see whether or not there was a (within-group) relationship between how difficult participants thought the task was and their attribution ratings (i.e., the harder the task, the more likely performance was externally attributed). However, there were no significant correlations for either group.

A further possible explanation for why there were no significant differences for this measure is that the question might have been ambiguous or confusing, and thus, the ratings became an artefact of mid-point reporting. It is difficult to test this hypothesis using the data from Experiment 4, but on reflection, this question was the one that several participants queried. However, it was felt that it was premature to discard this question based on the null findings from one experiment.

A series of questions were also included in an attempt to determine possible reasons why participants stopped persisting with the task during the initial persistence phase. Surprisingly, there were no significant differences between the groups for any of these questions. However, it is particularly interesting to note that the statement participants rated as most true was the one in which they were asked whether they had stopped because they had enough information to comment on the tasks (mean rating collapsed across both groups = 4.57). The second highest rated question was whether they stopped because they thought they had spent enough time on the tasks (mean rating collapsed across both groups = 4.39). Given that these were the explicit task instructions and were rated higher than "Because the task was too EASY" (1.96), "Because the task was too HARD" (2.96) and "Because you got BORED" (3.61), it seems that participants reported obeying the explicit task instructions rather than reporting being affected by their performance outcomes. It is also worth noting that participants in the Easy group did differ significantly in their rating of task ease and task difficulty as factors in their decisions to stop. The same is true for the Hard question, where participants in the Hard group did not rate their reason for stopping as being due to the questions being too hard.

So, whilst there is evidence that participants experienced the tasks differently (e.g. see section on task difficulty and how well did you think you did at the task), it appears that they were unaware - or not willing to admit - that this might have been influencing their persistence. Given that participants' initial persistence was influenced by task difficulty, it seems that the apparent lack of awareness of this influence suggests that their ratings were much more about what they believed they should have experienced, rather than what they actually experienced. This point is taken up further in the conclusions to this chapter when the issue of demand characteristics is examined in more detail.

Conclusions

Experiment 4 showed that when participants performed poorly at a task presented to them with a mastery-focus, they persevered with this task initially, but in a subsequent free-choice session, they seemed to lose their motivation to engage with the task. The opposite patterns of persistence were exhibited by participants who performed well at the task. However, it appears that the use of the numbers task in Experiment 4 affected participants' perceptions of task difficulty, resulting in the Easy task being perceived as more difficult than was intended. It is suspected that this problem also resulted in the failure to find significant differences between the two groups in terms of the dependent measures of fun, interest and enjoyment.

3.5 Experiment 5: Unlimited persistence, free-choice persistence and ratings of interest/enjoyment after differing levels of success.

Experiment 4 provided tentative support that participants may persist in different ways across different persistence measures, depending on whether they do well or poorly at a task. Unfortunately, significant differences between the two groups were not observed for the key dependent measures of free-choice persistence and self-reported ratings of fun, interest and enjoyment. It was suggested that the failure to find differences between the two groups for these measures was because of two methodological problems. Firstly, it was felt that whilst the anagrams in the Easy condition were experienced by participants as easy, the numbers tasks were not. It was therefore suspected that the numbers task undermined the potential successful experiences of participants in the Easy condition. Secondly, the method used to calculate the free-choice persistence times may have resulted in these times being artificially inflated for participants in the Hard condition. This was because free-choice persistence times were inferred rather than precisely measured by observing actual behaviour during the session.

A solution to the problem caused by the apparent difficulty of the numbers task was to remove these tasks and just give participants the anagram task. Alternatively, the numbers tasks could be made even easier, but given that in Experiment 4, even the simplest of the numbers tasks took longer to solve than the hardest of the Easy tasks, the numbers task appeared to be a confound rather than a useful variable to encourage feelings of success. Additionally, it would seem methodologically more efficient to only give participants one task to complete.

Experiment 5 was therefore a replication of Experiment 4 but this time, amendments were made to redress the methodological issues identified in Experiment 4. Firstly, participants were only presented with one task, namely, the anagrams. Secondly, free-choice persistence was measured using the more traditional method espoused by Deci (1975), whereby participants were covertly filmed and persistence measured from filmed evidence of their behaviour. In identical fashion to Experiment 4, tasks were presented to participants with a minimum emphasis on performance or evaluation. This was the mastery-orientation manipulation.

Following the patterns observed in Experiment 4, the experimental hypotheses were as follows:

- In the initial session, participants who attempted the Easy anagrams would persist *less* compared to those who completed the Hard anagrams.
- In the free-choice session, participants who attempted the Easy anagrams would persist *longer* compared to those who completed the Hard anagrams.
- Participants would find the task more fun, interesting and enjoyable after completing the Easy anagrams.

3.5.1 Methods

Overview In a similar fashion to Experiment 4, Experiment 5 was a run in two stages, an initial persistence phase followed by a free-choice persistence phase. Participants were randomly assigned to one of the two conditions (Easy or Hard anagrams) and were asked to comment on how interesting they thought the task was (i.e. mastery orientation).

In the first phase of the experiment, participants were taken into a room and asked to work on a series of anagrams. They were asked to keep going until they had sufficient information to comment on how interesting they thought the task was. When participants indicated that they were ready to answer some questions about the task, they were asked to wait while the experimenter got a questionnaire for them to complete. While they were waiting, the experimenter casually mentioned that they were free to try a few more anagrams. The experimenter then left the room for 6 minutes. The free-choice persistence measure was taken during this 6-minute period using covert-filming. Covert filming was achieved via an unobtrusive security camera mounted on the wall in sight of, above the eye-level of the participant.

When the experimenter returned to the testing room, the participant was given a short questionnaire that aimed to assess their interest, enjoyment and general experience of the task.

Participants 24 participants (17 female, 7 male) were recruited via the Stirling University Psychology department's participants panel. Participants volunteered as part of a first year departmental course requirement.

Apparatus The task stimuli were 330 specially prepared anagrams. All the anagrams were individually printed on a 110 mm x 40 mm laminated card. 210 of the anagrams were Easy, these anagrams ranged between 4-6 letters. 120 of the anagrams were Hard and these ranged between 7-12 letters. Details of these anagrams and their solutions are presented in Appendix 4.

The testing room was a small room with several tables and stacked chairs. One table and one chair were set out for the participant to complete the experimental task.

Procedure Participants were recruited from the Stirling University Psychology department's volunteers' panel via a poster which read "*I am conducting a series of experiments which investigate what makes solving puzzles such as anagrams interesting. You will be required to simply try a few puzzles and report how much you enjoyed doing them*".

Participants were pre-allocated to their group by the experimenter tossing a coin to determine which task, either Easy or Hard, each participant would complete.

When participants arrived for their appointment, they were taken to the testing room. On the desk was a box containing the anagrams and a pile of worksheets for participants to enter their answers (see Appendix 5 for example of worksheet). Each anagram had a code written on it e.g. E34. This code related to the anagram number in Appendix 4 and was used to help identify the correct solution to the anagram. The purpose was to convince participants that all the anagrams were soluble (during the pilot testing period, some volunteers had questioned whether the anagrams, especially the Hard examples, were actually soluble). Participants were asked to record the code of each anagram they attempted on the worksheet. The reason given was that this code allowed the experimenter to provide participants with the solution to any anagrams if they later wanted to check their answers.

Next to the work desk was a coffee table on which there were three magazines, the University student's paper, a psychology magazine and an issue of Q magazine (a popular music paper). In one corner of the ceiling was the security camera that was used to film the participants. Participants were seated so that the camera could film their actions but participants were placed in a position where the camera was not directly in line of sight. No comment was made by the experimenter regarding this camera; no participant subsequently reported being aware of the camera.

Also on the table was a list of task instructions. The participant was asked to read the task instructions and unless they had any questions, to begin with the task. As in earlier experiments, verbal contact with the participant was kept to a minimum in an attempt to minimise experimenter effects. The task instructions read as follows:

Task Instructions

In this task, I would like you to try and solve some anagrams. It is up to you how many you do, I am only interested in your experiences of the task.

Your Task Choose an anagram from the box, record the code on the answer sheet, and try to solve it.

Put your solved anagrams in the "Solved Anagrams" box. If you can't solve an anagram, just place it to the side of the box and try another.

When you think you think you have completed enough anagrams to be able to answer a few questions on your experiences of the task, please come and see me next door.

As the experimenter left the room, he started a wrist stopwatch. When the participant came out of the room to indicate that they had completed the task, the experimenter stopped the watch and made a mental note of the time. The participant was then taken back into the room and the experimenter gathered up their worksheet(s). The experimenter then said that he would like the participant to complete a short questionnaire but that the experimenter would need to get this from another area of the department as the usual printer was not working. The participant was asked to wait while the experimenter ran off a copy of the questionnaire. Participants were asked to stay in the room until the experimenter returned, but told that they were free to read the magazines that were on the table, to try a few more anagrams or just to wait. When the experimenter left the room, he went to the room next door that contained the video-recorder and activated the recorder. At this time, the experimenter wrote down the on-line persistence time on the back of the participants worksheet.

When 6 minutes of filming had elapsed, the experimenter turned off the recorder, returned to the room, and presented the participant with the questionnaire. The experimenter left the room whilst the participant completed the questionnaire.

When the participant had indicated that they had completed the questionnaire, they were debriefed as to the nature of the study and permission was sought to view the videotape. The participant was then thanked for their participation and asked not to divulge the nature of the experiment to other students.

Dependent measures

There were three dependent measures:

- Initial task persistence - This was the time from when the experimenter left the testing room for the first time, to the time that the participant left the testing room and indicated to the experimenter that they had completed the task.
- Free-choice persistence - Free-choice persistence was measured by a second experimenter blind to experimental, condition and hypotheses. Although the free-choice persistence paradigm has been the favoured method for measuring intrinsic motivation, the actual specific measure of behaviours for inclusion in the free-choice persistence period is imprecise. For example, Ryan, Koestner and Deci (1991) report that *" .. the amount of time the subjects spent working on the target activity was surreptitiously recorded ... "* As this was the first time that the free-choice period had been used in this thesis, and as the free-choice persistence was to be measured by a volunteer whose academic discipline was not Psychology, target behaviours for inclusion in the analysis were specified as follows.

  "Using the timer display on the video-recorder, review the tape and identify the amount of time the participant spends working at the anagrams. Feel free to rewind and/or pause any section of the tape to help make your judgements. "Working on the anagrams" includes such behavioural features as the time that the participant spends i. handling an anagram ii. "considering" or appears to be thinking about an anagram, even if they leave the workdesk to do this, iii. writing down any notes/workings out. If you have any queries, please ask".

- Answers to questionnaire items – these were the same questions used in experiment 4 (see Appendix 3).

3.5.2 Experiment 5 - Results

24 participants (17 females, 7 males) were recruited via the Stirling University psychology department's volunteer panel.

Table 7

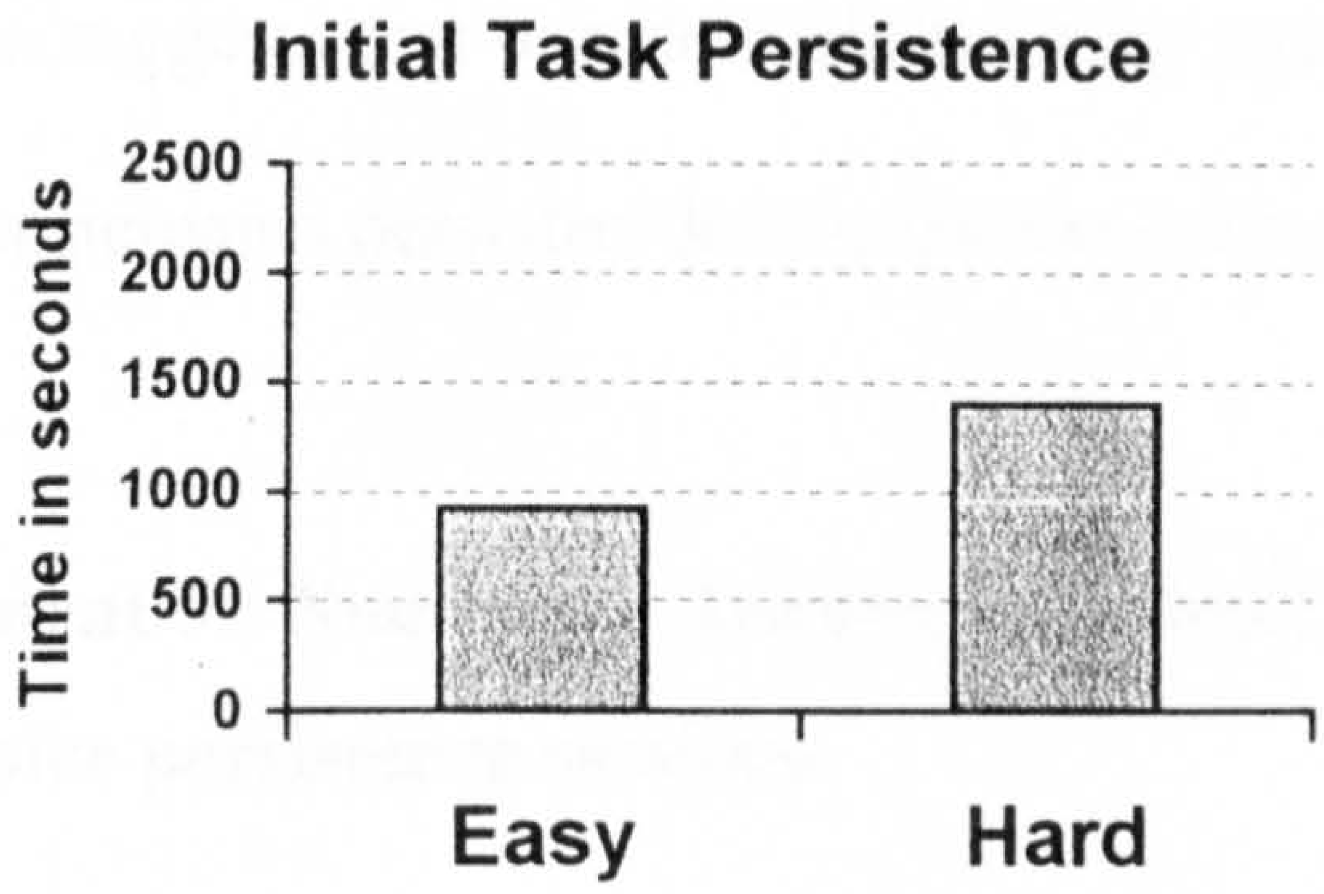
Means and standard deviations (in italics) for all dependent measures for anagrams and numbers experiment - Experiment 5 (n=24).

	Easy anagrams		Hard anagrams	
	mean	*s.d.*	**mean**	*s.d.*
Initial persistence	**925***	***365***	**1399***	***538***
Problems solved	**29.8****	***14.1***	**14.0****	***5.39***
Free-choice persistence	**246.1***	***88.8***	**153***	***109***
Problems solved	**8.6****	***4.1***	**0.7****	***0.96***
Fun	**4.2***	***0.84***	**3.2***	***1.14***
Enjoy	**4.0****	***1.04***	**2.7****	***0.75***
Interest	**4.2***	***0.72***	**3.3***	***0.78***
Task Difficulty	**4.4****	***0.67***	**2.0****	***0.92***
WellY	**4.2****	***0.87***	**2.1****	***1.0***
WellO	4.2	*1.03*	3.7	*0.75*
Attribution	3.2	*1.03*	3.8	*1.19*
Stop 1	5.1	*1.24*	4.2	*1.47*
Stop 2	2.2	*1.34*	1.4	*0.51*
Stop 3	2.3	*1.6*	3.3	*1.3*
Stop 4	3.7	*1.3*	3.2	*1.75*
Stop 5	4.7	*1.2*	4.2	*1.34*

Measures in **bold** are significant * = $p < 0.05$ ** = $p < 0.01$

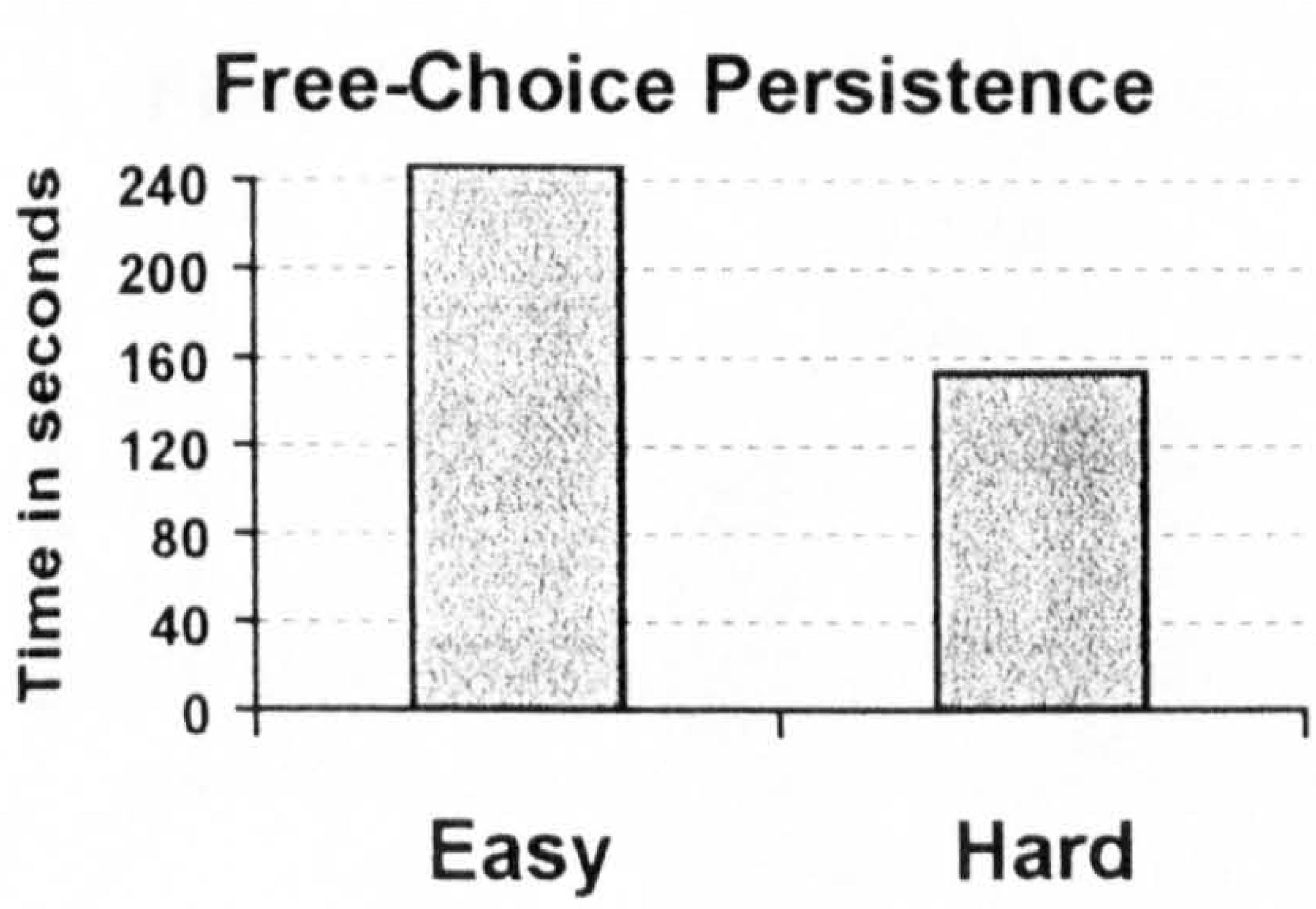
Table 7 above shows that in Experiment 5, there were significant differences between the Easy and Hard groups for all the key dependent measures except WellO. However, in a similar fashion to Experiment 4, there were no differences between the groups for the attribution and stop questions. The results are described in more detail in the following sections.

3.5.2.1 Experiment 5 - Initial Persistence



The differences between the Easy and Hard groups for on-line persistence were significant. Participants completing the Easy anagrams persisted less than participants completing the Hard anagrams, $t\ (22) = 2.53$, $p < 0.05$. This result is the same as the previous 4 experiments in this chapter.

3.5.2.2 Experiment 5 - Free-choice Persistence



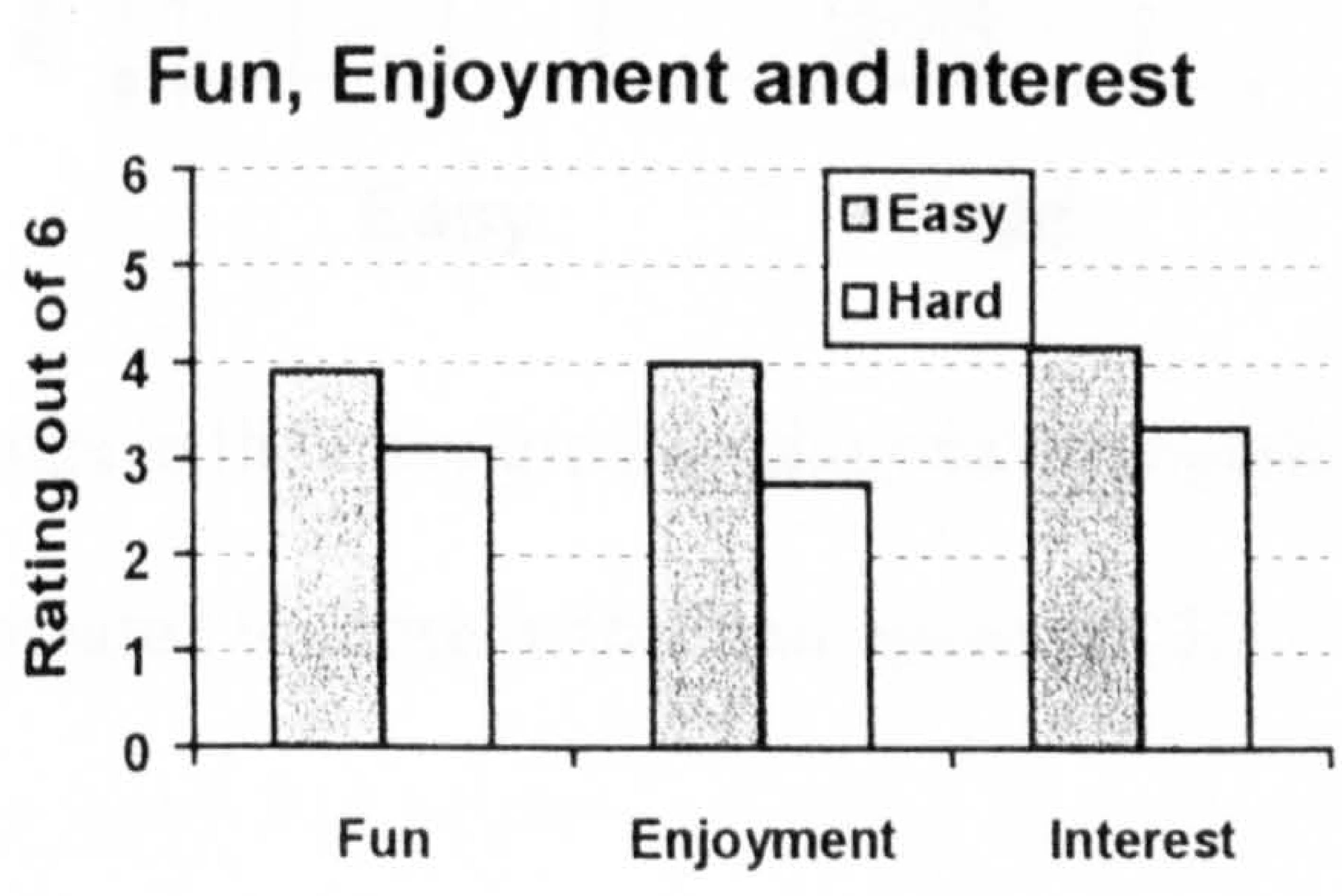
Participants in Easy group persisted for significantly more time with the anagrams during the 6-minute free choice period, $t\ (22) = -2.30$, $p < .05$.

As mentioned in Experiment 4, one reason why participants in the Hard group persisted less in the free-choice session could have been boredom resulting from spending more time on the anagrams in the initial phase. To test whether this might have been the case, data from both groups were combined. A Pearson's *r* correlation test revealed no significant relationship between time spent in the initial period and the time spent in the free-choice period. This suggests that boredom did not play any significant part in determining whether participants persisted during the free-choice session.

3.5.2.3 Experiment 5 - Number of Anagrams tackled during the initial and free-choice persistence sessions

Participants solving the Easy anagrams attempted significantly more anagrams during the initial , $t\,(22) = -3.64$, $p < 0.01$, and free-choice persistence sessions, $t\,(22) = -6.44$, $p < .001$.

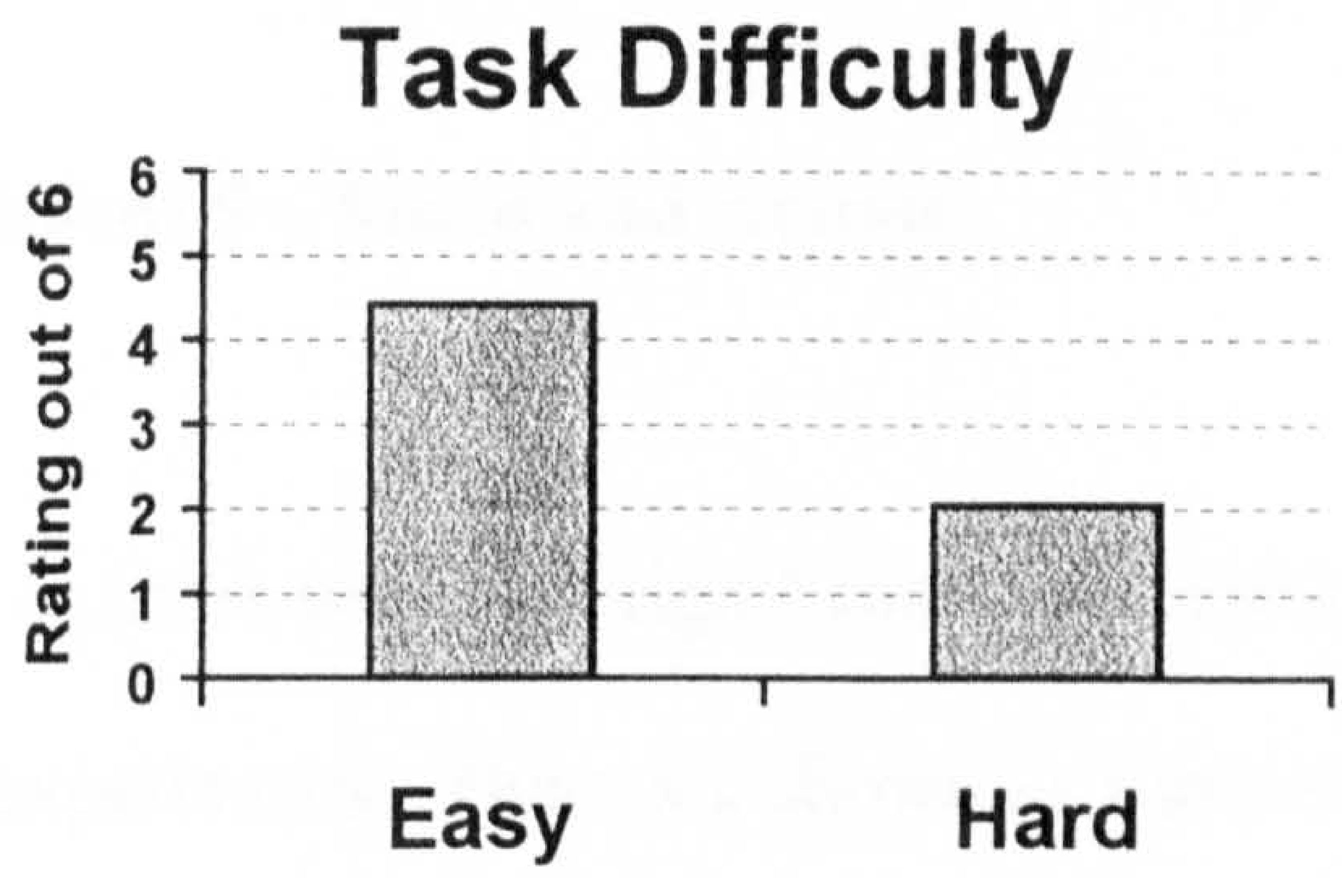
3.5.2.4 Experiment 5 – Fun, Enjoyment and Interest



In contrast to Experiments 1, 2 and 4 in this chapter, there were significant differences in participants ratings of fun, $t\,(22) = -2.25$, $p < .05$, enjoyment, $t\,(22) = -3.36$, $p < .01$

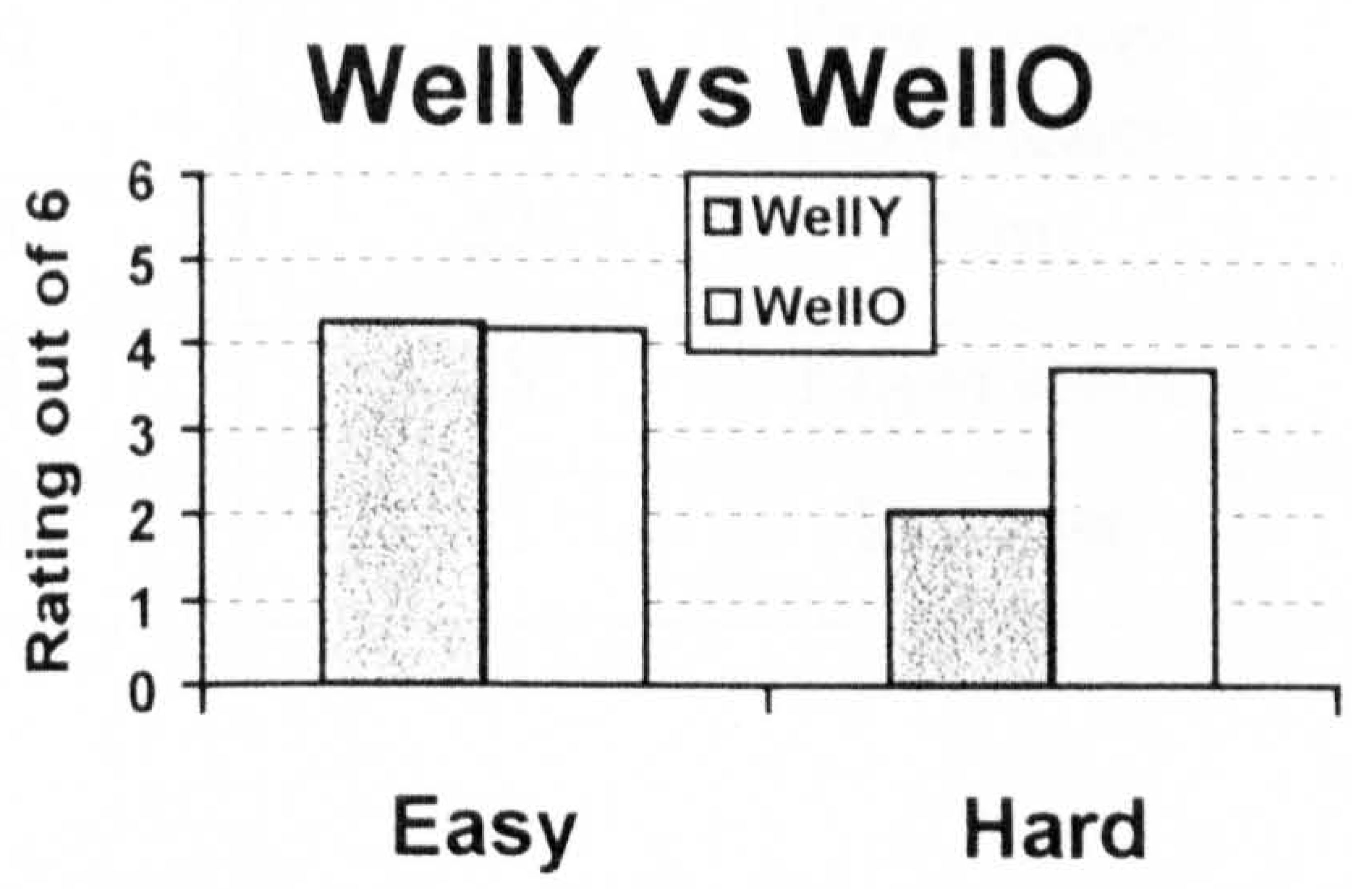
and interest, $t\,(22) = -2.73, p < .05$ whereby participants in the Easy group rated the task significantly more positively.

3.5.2.5 Experiment 5 – Task Difficulty



Participants in the Easy group reported the task as significantly easier than participants in the Hard group, $t\,(22) = -7.25, p < .001$.

3.5.2.6 Experiment 5 - WellY vs. WellO



<u>WellY</u> Participants in the Easy group ratings of their performance (WellY) was significantly higher compared to those in the Hard group, $t\,(22) = -5.69,\ p < .001$.

<u>WellO</u> There were no differences between the two groups for how well they thought others (WellO) would do at the task.

WYWO Participants in the Hard group rated WellO significantly higher than WellY, $t\,(22) = 4.99$, $p < .001$, that is, they thought that others would do better than they did. For participants in the Easy group, there were no differences in their ratings of WellY and WellO.

3.5.2.7 Experiment 5 – Mood and Activity

Unlike Experiments 1-3, there were two significant relationships between time spent on the task and reported experiences. These are shown in Table 8 below.

Table 8: Correlation table for the dependent measures of Initial persistence, free-choice persistence, fun, enjoyment and interest for Experiment 5.

Hard Anagrams	Initial persistence	Free-choice persistence	Easy Anagrams	Initial persistence	Free-choice persistence
Free-choice persistence	.01	--	Free-choice persistence	- .19	--
Fun	- .33	.22	Fun	.62*	- .18
Enjoyment	- .39	.28	Enjoyment	.23	.59*
Interest	.01	.01	Interest	.01	.33

* = $p < 0.05$

Table 8 above shows that for the Hard group, there were no significant correlations between the amount of time they spent on the task in the initial persistence phase and the their ratings of how much they enjoyed the task. This was also true for the free-choice phase. For the Easy group, there was a significant correlations between the amount of time they spent on the task in the initial persistence phase and the amount of fun the reported. For the free-choice phase, there was a significant (positive) correlation between this measure and the enjoyment ratings.

3.5.2.8 Experiment 5 - Attribution

No differences between the two groups were found in the ratings of whether their performance was due to their ability or the inherent difficulty of the anagrams.

3.5.2.9 Experiment 5 - Stop Questions

There were no significant differences between the groups for any of the questions relating to why they stopped persisting at the task.

3.5.3 Discussion

In Experiment 5, there were significant differences between the Easy and Hard groups across all the main dependent measures. The four main findings from experiment 5 were:

- Participants who tackled the Easy anagrams persisted *less* during the initial persistence phase relative to those attempting the Hard anagrams.
- Participants who attempted the Easy anagrams persisted *more* during the subsequent free-choice phase than those attempting the Hard anagrams
- Participants who attempted the Easy anagrams reported the task as more fun, interesting and enjoyable.

All of the trends from the previous 4 experiments were significant under the experimental conditions adopted in Experiment 5. The implications of these findings are now discussed.

Initial Persistence

In a similar fashion to all the previous experiments in this chapter, participants who attempted the Easy anagrams persisted less during the initial phase than participants who attempted the Hard anagrams. Given that this is now the fifth experiment to where participants have persisted with the task longer after failure relative to success, it seems that regardless of how the task was presented to participants, this form of measuring persistence seems to consistently lead to greater persistence after failure. The implications of the findings for the initial task persistence observed in all five experiments in this chapter are discussed in more detail in the next chapter.

Interestingly, compared to the anagrams and numbers task in Experiment 4, mean persistence times in Experiment 5 were lower for both the Easy (Exp. 4 = 1398 seconds vs. Exp. 5 = 925) and Hard (Exp. 4 = 2347 vs. Exp. 5 = 1399) groups. Recall that the claim was made earlier that success in a mastery-task should result in earlier disengagement from a task, because participants would recognise more quickly that they were competent. Following this logic, the prediction in terms of levels of success experienced in experiments 4 and 5, would be lower initial persistence in Experiment 5 <u>only</u> if participants actually experienced a higher level of success in Experiment 5 in both the Easy and Hard groups. Fortunately, there seems to be evidence to support this hypothesis. It is probably useful to examine the Easy and Hard groups separately to help make the logic clearer.

For the Easy group, recall that in Experiment 4, the numbers tasks took three times as long to solve compared to the anagrams. Therefore, it is reasonable to suggest that in Experiment 4, participants were experiencing <u>some</u> level of difficulty, even in the Easy condition. However, in Experiment 5, participants only solved the Easy anagrams which were made up from the same sample used in experiment 4 together with additional anagrams of similar difficulty. The removal of the numbers task meant that participants in Experiment 5 experienced virtually no failures, nor did they have to work hard to work out the anagrams as they seemed to have had to do with the numbers tasks. The worksheets for Experiment 5 show a mean success rate of 95.4% for the Easy anagrams. Also, participants in the Easy condition in Experiment 5 rated the task as easier and thought they had done better at the task (WellY rating) compared to participants in the Easy condition in experiment 4. Therefore, it seems reasonable to suggest that participants in the Easy group for Experiment 5 did indeed experience greater success compared to participants in the Easy group in Experiment 4.

Applying the same logic to the Hard group in Experiment 5 is more problematic, because the Hard tasks were <u>also</u> experienced as easier compared to the Hard tasks in Experiment 4. However, a tentative case can be made. Although the average number of anagrams solved in Experiment 4 was only 32.9%, analysis of the worksheets reveal that virtually all these successes came from the anagram tasks, not the number tasks. In Experiment 5, the average success rate was 52.6%. Using these percentages, a case can therefore be made that participants in the Hard group in Experiment 5 also experienced a greater level of success compared to participants in Experiment 4 and this is why they persisted less.

Free-choice persistence

Participants who attempted the easy anagrams persisted longer during the free-choice period compared to those who attempted the hard anagrams. This finding is similar to Experiment 4 except that this time, the difference was significant. Taken together, the initial and free-choice persistence patterns are intriguing because when the two groups are compared, they not only persisted in the opposite way from session to session, but their patterns of persistence were also opposite to one another. Participants in Hard group persisted more during the initial session but less during the free-choice session. There are several explanations that could be offered for these patterns of persistence but because they require reflection and further speculation as to what might have been happening in the previous four experiments in this chapter, these explanations are covered in the next section of this chapter.

Fun, Interest and Enjoyment

Participants who attempted the easy anagrams thought the task was significantly more fun, enjoyable and interesting compared to those who attempted the hard anagrams. For the first time in any of the experiments to date, all ratings exceeded an average of 4. It is also interesting to note that participants in the Easy group also rated the task as particularly easy ($M = 4.42$). Taken together, these findings suggest that by removing the numbers task, participants did indeed perceive the task to be easier and that this may have contributed to their enjoyment of the task.

The patterns for all the other dependent measures in Experiment 5 mirrored those observed in Experiment 4. Therefore it seems that it seems reasonable not to dwell on these findings as the same conclusions drawn for these dependent measures in Experiment 4 apply for Experiment 5.

Conclusions

Experiment 5 built on the findings of the previous 4 experiments to produce a methodology that resulted in clear differences between the Easy and Hard groups across all the main dependent measures. The small changes implemented in Experiment 5 (relative to Experiment 4) suggest that issues such as the choice of task and how this affects individuals' perceptions of their competence are important when attempting to manipulate feelings of success/failure.

However, the most intriguing result from Experiment 5 was the confirmation of the patterns of persistence suggested in Experiment 4 whereby participants persisted in one direction during the initial persistence phase but in the opposite direction in the free-choice phase. There are several possible explanations and speculations why this may have been so and these are addressed in the next section of this chapter.

3.6 Conclusions for Experiments 1-5

Summary of findings

Chapter 2 presented five experiments that examined how long individuals persisted with a task when they performed either poorly (failure) or well (success) under different types of task orientation. Experiments 1 and 2 showed that when participants were presented with the task with what was purportedly a mastery-focus, they persisted longer with the task after failure. When in Experiment 3, the same task was specifically presented with a performance-focus, participants still persisted for longer after failure compared to success. Experiments 1-3 therefore suggested that regardless of how the task was presented to participants, they always persisted longer after failure.

It was speculated that one of the reasons why participants persisted longer after failure in Experiments 1 and 2 was because of the method used to measure persistence. To investigate this further, Experiments 4 and 5 included an additional free-choice persistence phase. These experiments showed that during the initial persistence phase, when participants were presented with a task with a mastery-focus, they persisted for longer after failure, but that in the subsequent free-choice persistence phase, they showed the opposite persistence patterns, namely, greater persistence after success relative to failure.

Along with the persistence findings, Experiments 1-5 also showed that individuals generally reported more positive experiences of the task after success compared to failure. This effect occurred regardless of whether the task was presented with a mastery or a performance focus.

3.6.1 Motives for persistence

It may be useful to comment firstly on the patterns of persistence observed in the phases where participants were allowed to engage with the task for as long they wished. During these persistence phases, the evidence that participants persisted less after success compared to (relative) failure was almost overwhelming. Even in the experiments where participants' ratings of how well they thought they had done was similar between the Success and Failure groups, participants in the Failure groups *always* persisted longer. For example, in Experiment 2, participants in the Failure group who declined in performance to an average success rate of 10%, did not differ in how much they enjoyed the task compared to the success group. It was suggested that this may have been due to participants in the failure group actually experiencing a degree of success because they felt they had "cracked the code" - i.e., that the computer was influencing their results. However, despite this possible reason for experiencing success in the Failure group, there was still a significant difference between the two groups in terms of the amount of time they spent on the task.

Also, in Experiment 4 (the anagrams and numbers experiment), it was suggested that the lack of divergence in interest and enjoyment ratings may have been due to the fact that in the Easy group, participants found the numbers task difficult. However, as with all the other experiments, participants given the hard anagrams still persisted with the task for longer during the initial persistence session compared to participants who completed the easy anagrams. So even when the differences between the Success and

Failure groups for reported task enjoyment/interest were non-significant, this did not seem to be reflected in the amount of time participants persisted with the task initially.

Thus, the experiments in chapter 3 consistently showed that when participants failed at the task, they persisted longer than when they succeeded. In chapter 1, it was suggested that it was unclear whether or not participants who performed poorly at a task presented with a mastery-focus would persist for a similar amount of time as participants who did well, or would persist less. Dweck and Leggett (1988) had suggested that a mastery-orientation should lead to *"... the maintenance of effective striving under failure"* and *"... the generation of effective strategies in the face of obstacles"*, and Diener and Dweck (1978) had reported that mastery-orientated children appeared to maintain an *"unflagging optimism"* in the face of failure. However, not only did participants who performed poorly persist as long as those who did well, they actually persisted significantly longer. On the face of it, it would seem that the mastery-orientation did indeed result in participants striving to master the task. However, Diener and Dweck (1978; see also Diener and Dweck, 1980) reported that mastery-orientated children made positive verbalisations about the task they had just completed such as "I did it before, I can do it again" or "I'm sure I have it now", and that the mastery-orientated children were less anxious, depressed and bored than their helpless counterparts. These positive responses to poor performance were not observed for participants in Experiments 1-5. In all instances, there was a trend for participants who performed poorly to rate the task less positively than those who did well. In this sense, there seems to be little evidence that the (apparently) mastery-focused participants in Experiments 1-5 were experiencing the tasks positively.

In the discussion section for Experiment 1, it was suggested that the methodology used to measure initial persistence may have inadvertently resulted in participants experiencing a pressure to perform well at the task. The suggestion was that because participants were aware they were going to be questioned regarding their experiences of the task, this may have implied to them that one of the issues that they would be asked to comment on was their performance. If so, it would not be unreasonable to suspect that they would want to perform well at the task, not because they wanted to master the task, but because they were concerned about how their performance would appear to the experimenter. Thus if participants were in fact performance-focused, then this would explain their negative experiences when they performed poorly.

There are several possible explanations that may help to support the hypothesis that participants were performance-focused during the initial persistence phase. For example, recall that in chapter 1, when discussing issues regarding the ecological validity of psychology experiments, Tedeschi and Reis (1981) suggested that all behaviour is either consciously or sub-consciously an attempt to manage the impressions individuals give of themselves to others. These authors claimed that social contexts, such as experimental situations, serve as important causes for behaviour and that although participants may be motivated to behave in a one way prior to arriving for the experiment, the mere fact that they are engaging in a psychology experiment may cause them to behave in a way specific to the implicit social rules that concern the experimental situation (see also Alexander and Rudd, 1981). Intuitively, it would be reasonable to expect that participants arriving to take part in a Psychology experiment would be concerned about how they appeared to the experimenter. This is particularly relevant considering the population used in the experiments in chapter 3, as many of these students would have encountered myself as a tutor/lecturer. It seems reasonable to suspect that they would have been keen to create a favourable impression. In terms of Tedeschi and Reis's analysis, the argument could be phrased that in Experiments 1, 2, 4 and 5, although the purpose was to present the task with a minimum emphasis on evaluation, participants were already concerned about performing well even before they knew what the task was. Thus rather than the task being mastery-focused, participants were already performance-focussed because of their desire to create a good impression of themselves for the experimenter.

Similarly, Nicholls (1984) has suggested that in achievement contexts, individuals will seek to prove their ability, and that announcements that skill tasks are to be used should induce concerns about performance. It would be reasonable to suggest that the anagrams, numbers and stockmarket tasks could all be defined as skill tasks. If this is so, then it lends further support for the contention that participants may have already been performance-focused.

Whilst these explanations have merit, the breadth of evidence that has shown participants responding differently when tasks have been presented with different orientations suggest that this analysis may overstate the case. For example, in answer to Tedeschi and Reis's (1981) suggestion that psychology experiments cause participants to become concerned about performing well, when Harackiewicz et al (1989) gave participants a pinball task to perform, there were clear differences in the amount of time they persisted with that task depending on whether they were introduced to the task with a mastery or a performance-focus. If these participants were already performance-focused, they should have behaved similarly regardless of the task-orientation condition to which they were assigned to. A similar criticism can be levelled at Nicholls' suggestion that announcements that skill tasks are going to be used are likely to induce a performance-focus because, if this were so, then in the Harackiewicz et al (1989) study, there should have been no differences in persistence between the conditions.

Thus, as potential explanations for the behaviour observed in the initial persistence phase, it seems reasonable to rule out that participants were already performance-focused prior to arriving for the experiment or that the fact that the task was a skill one induced a concern to perform well.

Thus, for the experiments in chapter 3, the claim is that it was the methodology that induced the performance-focus. That is, because participants were required to meet and discuss their behaviour with the experimenter after they had completed the tasks, they would probably have been concerned about their performance. The similar patterns of persistence observed in Experiment 3 when a performance-focus was specifically induced supports the claim that Experiments 1 and 2 also induced a performance-focus, however unintentionally.

However, the position that participants were performance-focused becomes complicated when the results for Experiment 5 are reviewed. If participants were performance-focused, then according to the results observed in Experiment 3 of the Ryan et al (1991) study, participants should have persisted for longer during the subsequent free-choice session after failure, not after success. That is, if, as Ryan (1982) suggests, participants had experienced some sort of intra-psychic pressure to perform well during the initial persistence phase, then after failure, they should have also persisted for some time during the free-choice phase in an effort to recover lost self-esteem. This did not happen.

However, Ryan's explanation for why, after failure, participants persist for longer in the free-choice phase presumes that this behaviour was somehow a carry-over from the initial persistence phase. There is good reason to suspect that this might not have been the case. Indeed, it is not altogether clear that the behaviour in the subsequent free-choice persistence session tells us anything about orientation during the initial persistence phase.

Let us start by considering the implications if participants were mastery-focused during the initial persistence phase. For participants given the easy anagrams, it seems likely that their success would have encouraged them to feel that they had mastered the task or, at any rate, had made substantial progress towards doing so. It may have been that this sense of reaching, or nearly reaching their goal, led to their decision to stop. For the participants in the Hard group, on the other hand, their low levels of initial success may have motivated them to try harder, so that it is reasonable to suspect that they would have continued to work on the task, even after participants in the Easy group had stopped. In practice, this extra effort had little effect because of the extreme difficulty of the task—on average, participants in this group solved only five anagrams in twenty-three minutes —so eventually they concluded that they would not be able to improve on this level of performance and hence stopped trying. To phrase this explanation in a different way, it may have been that the Hard group persisted because it took them longer to realise that they were not likely to reach their goal of mastering the task.

Turning to the free-choice phase, it might at first appear that neither group should spend almost any time on the task because the Easy group would have already almost achieved their goal, while the Hard group would have concluded that further effort would be to no avail. However, even though the free-choice phase followed on almost immediately from the initial phase, it represented a potentially important shift in experimental context. When participants initially decided to stop working on the anagrams, they would have believed that they only needed to answer a few questions and then would be free to engage in other activities. When the experimenter informed them that he had to leave and asked them to wait for him, they would have found themselves effectively trapped in a room with a very limited range of activities. Thus, even if they had earlier exhausted (at least temporarily) their desire to work on the anagrams, when confronted with a suddenly restricted range of options, they might have concluded that the anagrams were still more interesting than the other available activities. In the paradigm employed in Experiment 5, the term "free-choice period" may be a misnomer, in that participants actually only had a restricted range of options.

If we assume that participants still had some desire to master the task, it becomes easier to understand why those in the Easy group now persisted longer. These participants might have believed that they could still improve further, and hence returned to working on the anagrams. Participants in the Hard group, on the other hand, had already spent a substantial period of time trying to improve—recall that the average persistence time for participants in the initial phase was twenty-three minutes. Believing that further effort would be to no avail, it would be unreasonable to expect them to choose to spend further time working on the anagrams.

This interpretation of the results seems plausible, but there is still one aspect that is troubling. Specifically, it is not entirely convincing that participants in the Easy group, having already worked on the anagrams for a long period during the first phase and then decided to stop, would have returned to working on them moments later in order to further prove their ability. It seems doubtful that their motive for working on the anagrams during the free-choice period was to demonstrate competence or to improve their ability. Deci (1975) has suggested that intrinsic motivation gives rise to two classes of behaviours one of which is related to conquering challenging situations. For example, he suggests that .. *people are motivated to "reduce uncertainty" or "reduce dissonance" or "reduce incongruity" when they encounter it"* (Deci, 1975, p. 57). Applying this hypothesis to behaviour during the free-choice persistence phase, once participants concluded that there was no point in further effort during the initial persistence phase, why did they decide to try again only moments later? What exactly were they trying to prove or reduce that they had not already proved or reduced during the initial persistence phase? Is it really plausible to believe that after persisting for so long during the initial persistence phase, they had not reached some sort of resolution with regards to their competence at the anagram task? Deci (1975) has also suggested that individuals engage with tasks because they desire an optimal challenge, one that is neither too easy nor too hard for them. Again, if this is so, why re-engage with a task when one has presumably already decided whether it is challenging or not?

Similar problems arise if we assume that participants in the Easy group had a performance orientation rather than a mastery orientation. Recall that Ryan (1982) suggested that the reason that participants persisted during the free-choice period after failure was because they had put themselves under an internal (or *intrapsychic*) pressure to prove to themselves that they could perform the task well so that they could recover lost self-esteem. This version of the performance hypothesis, faces the same problem as the preceding version: What exactly were they trying to resolve that they had not already resolved during the initial unlimited persistence phase?

The problem with understanding why participants persisted as they did during the free-choice period seems to stem from the same ambiguity that is inherent in both intrinsic motivation and achievement motivation theory, namely, insofar as mastery is a goal, how do individuals operationally define mastery for themselves? For example, must they achieve perfect scores; if not, what constitutes an acceptable approximation? How exactly do individuals decide when they have or have not mastered a task? In terms of a performance-focus, in the absence of verbal confirmation or explicit normative standards, how exactly do individuals decide whether or not they have exhibited high ability? I have speculated that participants resolved their mastery or performance issues during the initial persistence period but these speculations are only valid if participants were indeed trying to resolve or prove something during the free-choice period.

However, if participants were not trying to prove competence or trying to recover lost self-esteem, then an alternative explanation is required to answer why they behaved as they did during the free-choice period in Experiment 5? One possibility is that they weren't trying to prove anything; they returned to working with the anagrams because solving them was fun. Phrasing this point with somewhat greater precision, participants' successes during the first phase could have left them feeling that they were indeed competent at solving anagrams, and they could have resumed working on them for the anticipated pleasure of doing so, rather than to further improve their ability. In a series of studies, Feather has shown that when participants perceive they have been successful or believe that they will be successful in the future, they report low levels of anxiety Feather (1963a, 1968, 1969). Similarly, both intrinsic motivation and achievement goal-theorists have acknowledged that success on a task can be enjoyable. Deci and Ryan (1980), for example, have noted the "*intrinsic satisfaction*" that flows from "*taking on new challenges and working to master them*". Dweck and Elliott (1988) have also suggested that "*intrinsic rewards from... high-effort progress*" can sustain work in the face of difficulties (p. 262). Direct evidence that participants in Experiment 5 experienced the task positively after performing well is available from participants' questionnaire responses. Relative to the Hard group, participants in the Easy group rated the task as significantly more fun, interesting and enjoyable.

However, in both Deci's and Dweck and Elliot's theory, the concept of intrinsic satisfaction or reward is rarely used to account for behaviour, and in some cases the wording of the theories seems to imply that this pleasure is only an adjunct to making progress. For example, Deci (1975) states that "... *intrinsically motivated behaviours are behaviours aimed at bringing about certain internal rewarding consequences .. specifically, they are intended to bring about the feeling of competence and self-determination*" (p. 59). In other words, the assumption seems to be that the primary goal is task mastery, with pleasure from doing well a small (though fortunate) by-product. Applied to the situation in Experiment 5, the implication would seem to be that if participants were no longer hopeful of improving their ability by the time they reached the free-choice period, they would have derived little pleasure from solving further anagrams. In contrast, it may have been that rather than trying to improve, the pleasure or satisfaction derived simply from having done well at the anagrams during the initial persistence phase, played a major role in motivating behaviour during this period. Participants in the Easy group experienced substantially more success during the initial phase; this would have led them to expect similar success during the free-choice period, and it could have been this expectation of success, rather than a desire for further improvement, that motivated them to return to the anagrams. There is tentative support for this hypothesis if we consider the mood and activity ratings in Experiment 4 (see Table 6). Table 6 shows that it was primarily in the free-choice phase of the Experiment that mood and activity scores were correlated and this was observed for both the Hard and Easy groups. This finding suggests that participants persisted in line with how much they were enjoying the tasks. Indeed, for Experiments 1-5, mood-activity only positively correlated once during the initial persistence phase (see Table 8) Unfortunately, the mood-activity patterns observed in Experiment 4 were not replicated in Experiment 5, but the mood-activity correlations for the free-choice phase in Experiment 4 at least suggest that participants may have been persisting in line with how much pleasure they were getting from the task.

Two possible accounts have been suggested for the opposite patterns of persistence observed during the initial and free-choice persistence phases, and each has implications for future research. If the first account is correct, and the different results simply reflect different levels of intrinsic motivation at the conclusion of the first phase, the results may be taken as a pointer to the need for further elaboration of current theories of achievement motivation. As mentioned earlier, it seems that further specification is needed of how people decide when to stop working on a task. To reiterate, how exactly do individuals decide when they have or have not mastered a task? How exactly do individuals decide whether they have exhibited high ability?

If the second account is correct, existing theories may need to give greater emphasis to the role of the pleasure or satisfaction that flows from success. It has become standard practice to refer to two possible motives for achievement-related tasks, a performance motive that focuses on impressing others and a mastery motive that focuses on attaining mastery—in Dweck and Leggett's (1988) succinct summary, the difference between *proving* ability and *improving* ability. However, the pleasure derived from success may also be a powerful motive for working on tasks, whether crossword puzzles, sports, or psychological research. It is difficult to speculate about the source of this motive—it could be a derivative of the desire for mastery, a form of secondary reinforcement derived from experiences of social approval, or something quite different—but the findings from Experiment 5 tentatively suggest that the role of this motive may have been underestimated and overshadowed in the now widely-accepted division of achievement motives into performance and mastery. There may be a third important motive: The pleasure derived from doing well.

3.6.2 Comments on WellY vs. WellO

There are two issues that seem worth commenting on with regards to the WellY vs. WellO analyses for the experiments in this chapter. The first of these is how consistently participants believed that they did as well as others would after performing well, but that after performing poorly, they tended to think that others would do better. This latter finding seems an interesting one because it suggests that after performing poorly, participants suffered some type of loss of confidence in their ability.

To elaborate, for those participants who performed poorly, they did not report their performance as good but others as even better, they reported their performance as poor and others as mediocre. For example, if participants in the Failure groups had reported that they had done quite well (rating > 3.5), then if they subsequently rated others as doing much better than this, then this would not really represent a loss of confidence. In this case, participants would be saying something like "..well, I did OK but others would do much better than me". Instead, participants seemed to be making a statement something like, " I performed poorly compared to others whose performance was only mediocre". Brown and Dutton (1995) have suggested that when individuals experience failure, they can suffer losses in what they have called "feelings of self worth" (FOSW) or a temporary loss in self-confidence. In keeping with Brown and Dutton's analysis, the WellY vs WellO data seems to indicate that participants did indeed experience some loss of confidence after performing poorly at the task. These negative experiences of the task also seem to be similar to the types of verbalisations made by the children classed as helpless in the Diener and Dweck (1978, 1980) studies. For example, the self-deprecation of participants who performed poorly in the experiments in chapter 3 seem to be similar to types of verbalisations made children in the Diener and Dweck studies (e.g., attributing failure to personal inadequacy).

It seems that rather than simply saying that participants lost motivation for the task or were disappointed to have performed poorly, it seems that their experiences were more devastating than a simple recognition that they had done poorly at the task. In the Anderson and Rodin (1989) study, they reported that participants who were told that they had performed in the 55th percentile reported being "mildly disappointed but not devastated". This did not seem to be the case for participants in the experiments in chapter 3. These participants seemed to have experienced something more dramatic.

At a broader level of analysis, it is particularly interesting to note that participants' believability about how well others would have performed at the task (WellO) were the same whether they themselves had done well or poorly, that is, the magnitude of this rating was similar across all experiments. It seems that regardless of how well participants did at the task, they seemed to have a collective and agreed estimate of the type of standard that should be achieved for the task. But why should this be so and moreover, how do participants arrive at these seemingly arbitrary standards of acceptable performance? Recall that at no time in Experiments 1-5 were participants given any indication of how well they were performing relative to others, yet when asked, they seemed to report that others would perform to a certain standard. One argument for this could simply be that participants did not know how to respond to this question and simply chose a rating that was essentially nebulous (i.e., 3.5). Collapsing all of the WellO data for all the groups in all the experiments and comparing this to sample mean of 3.5 revealed no significant differences. This suggests that when asked to rate how others would have done at the task, participants were not really comparing their performance to others at all, they were just assigning a rating for a question that was ambiguous. Whilst this might sound plausible, I suspect that this was not the case and that the WellY vs. WellO analysis gives some indication of how individuals compare themselves to others in task situations. It is felt premature to elaborate further at this point in the thesis because Experiment 6-10 also used a WellY vs. WellO analysis. Thus, this issue of the validity of the WellY vs. WellO analysis will be discussed in the final chapter of this thesis and the discussion will be extended to speculate as to how individuals set performance standards in situations where normative standards are either ambiguous or unavailable.

3.6.3 Methodological issues

Throughout the course of the experiments in chapter 3, several methodological issues arose which seemed to critically affect the results. Many of these have been commented on in the discussion sections for each of the experiments - e.g. levels of failure and believability; type of task used. However, there are several broader issues that are of interest.

3.6.3.1 Unlimited persistence

In considering the different patterns of behaviour observed in the initial and free-choice phases, it was speculated that because participants were allowed to persist as long as the wanted to in the initial persistence phase, this may have given them the chance to overcome any negative performance-focus effects caused by the methodology. If so, this could have implications for many other experiments in the field of achievement motivation. Most experiments in this area give participants only a restricted amount of time or number of tasks to complete. The results from the experiments in this chapter suggest that if participants were given unlimited time, this might in some cases significantly affect their reactions to the task. For example, in studies that have claimed to have researched the phenomenon of "ego-involved" persistence, this ego-involvement might only be an issue if participants are given the chance to resolve their competence issues (e.g., Ryan, 1982; Ryan, Koestner and Deci, 1991). Thus, ego-involved persistence may only relate to situations where tasks are interrupted or when participants are not allowed to progress through that task to completion. Such a conclusion does not discount the possibility that the phenomenon of ego-involvement is a real one, it merely suggests that the phenomenon may be restricted to limited range of task situations.

3.6.3.2 Demand characteristics

Whilst the phenomenon of greater persistence after failure was a consistent and robust one, similar significant differences in self-report ratings did not emerge until Experiment 5. The suggestion therefore is that the evidence from the first 4 experiments suggests that the self-report measures were less sensitive relative to the purer behavioural measures of persistence. When participants experienced differing levels of success, they persisted for a significantly different amount of time, but when it came to reporting on their experiences, it was only in Experiments 3 and 5 that significant differences emerged for the task enjoyment measures.

So despite acting differently, participants were less likely to report their experiences differently. One suggestion for this pattern of behaviour is that when individuals are asked to report on their experiences of a task, they are more likely to search for reasons for their behaviour rather than reporting their actual experiences. For example, in Experiments 4 and 5, when participants were asked to say why they stopped persisting with the task, although all the other dependent measures (e.g. task enjoyment, task persistence, WellY vs WellO) indicated that participants were persisting because of the success or failure they were experiencing, this was not reflected in their reported reasons for desisting with the task. Instead, participants reported their reason for desisting with the task in line with the task instructions (e.g. they reported stopping because they had enough information to comment on the task).

There are several interesting implications of these findings. For example, it was mentioned earlier that demand characteristics might have a large role to play in determining participants' behaviour in psychology experiments. Recall that it was suggested by Tedeschi and Reis (1981) that experimenters can give cues to participants for how they should behave and respond to the experimental manipulations (see also Page,1981; Orne, 1969, 1971). However, when participants were asked why they stopped engaging with the task in Experiments 4 and 5, they claimed that it was because "they had enough information to be able to comment on it" rather than their disengagement having anything to do with how well they were doing at the task. Recall that participants were also asked whether they stopped persisting with the task because they were doing well or poorly, but that they rated the truth of these statements as relatively low. If participants were aware of the experimental manipulations, when given the chance, they were either actually unaware of the experimental manipulation affecting their behaviour or they were strongly resisting admitting it. A further examination of the debriefing notes reveals no particular comment on participants having stated that their performance caused them to desist, further suggesting that participants were indeed unaware of the experimental manipulations.

Researchers such as Page (1981) are right to insist that other researchers adopt a sceptical approach to their experimental work. Indeed, the above comments on whether or not participants were aware of the manipulations in Experiments 1-5 could also be criticised because it may have been that participants were, as Tedeschi and Reis (1981) put it, "trying to be good subjects" and thus knew that I wanted them not to be aware of the success/failure manipulations. If this were so, this would also explain why the question they rated most highly was "I stopped because I felt I had enough information to comment on the task". Also, despite them not commenting on the success manipulations during the debriefing sessions, it should be recalled that some participants in Experiment 2 did suggest that it was the computer that was determining their performance. Whether this was a reaction to their poor performance or that their poor performance caused them to reflect on the purpose of the experiment with a more sceptical approach is probably moot. However, such arguments suggest that although there seems to be reasonable evidence that participants were not aware of the manipulations in Experiments 1-5, the opposite may be the more appropriate conclusion to draw, but less easy to support.

The ecological validity of the experiments in this and the next chapter are discussed in more detail in the final conclusions chapter so further comment in this regard is deferred until then.

3.6.3.3 Objective success vs. experiencing success

Another feature of the set of experiments in this chapter was the difficulty in getting participants to experience success. Even when participants (in Experiment 2) achieved a 90% success rate, the reported WellY average was still only 3.63. Participants reported feeling most successful in Experiment 5 where those in the Easy group achieved an average of over 95% anagrams correct. Here the reported WellY average was 4.25, hardly an overwhelming endorsement that they were experiencing high levels of success.

The methodological point to be made is that when researchers define the conditions of an experiment, care must be taken when claiming that certain conditions produce feelings of "success". This may be particularly so in experiments where participants are not explicitly told whether they are doing well or not. For example, recall that other researchers who have adopted a similar approach augmented performance outcomes by explicitly advising participants how well they did (e.g., Brown and Gallagher,1992; Brown and Dutton, 1995; McFarlin and Blascovich, 1984) suggesting that these researchers were aware that allowing participants to simply experience success/failure may not have been sufficient for them to experience the amounts of success the experimenter desired. The experiments in this chapter seem to support the concern that some researchers have had about inducing feelings of success. It seems that the independent variable is not any single objective level of success that the experimenter has manipulated, but whether or not participants actually experienced success.

As mentioned earlier, explaining why participants seem so reluctant to report having performed well when objectively, it would seem that they have, may be bound up with the types of standards they set themselves. This issue is discussed further in the final conclusions chapter of this thesis.

3.6.4 Summary of findings in chapter 3

Chapter 3 set out to investigate the length of time participants who performed poorly at a task would persist when that task was presented to them with a minimal emphasis on evaluation (i.e., mastery-focus orientation), compared a group who did well. It was shown that they consistently persisted for longer after performing poorly. This phenomenon was observed in Experiments 1 and 2. In Experiment 3, participants were introduced to the same task with a performance-focus and again, persisted longer after failure. Speculating that this might have been due to the method used to measure persistence, Experiments 4 and 5 included two types of persistence, unlimited and free-choice. In these experiments, participants persisted longer during the initial unlimited persistence phase but for less time during the free-choice phase.

Although the key finding for the experiments in chapter 3 was the opposite persistence patterns in the two different phases, the issue that has proved to be the most controversial was trying to account for why participants persisted at all during the free-choice period. Ryan's (1982) suggestion that participants may have experienced some sort of intra-psychic pressure to persist with the task was discarded on two counts; firstly because if this theory was applicable to the persistence observed in the free-choice period, participants should have persisted longer after failure not success. Secondly, it was not entirely clear why participants should have experienced any pressure to perform well during the second persistence phase. In a similar fashion, Dweck and Leggett's (1988) and Nicholls' (1984) suggestion that individuals may have been trying to master the task or prove ability is also difficult to apply. This is because participants persisted for so long during the initial persistence phase, it is questionable whether they had anything further to prove, in terms of competence or ability, during the free-choice persistence phase.

It was suggested that a possible explanation for why participants persisted during the free-choice period could simply be phrased in terms of "pleasure". That is, rather than participants pursuing any particular goal, the pleasure derived from performing well at the task during the initial persistence phase served as an explanation for their behaviour. This account is very similar to Deci's (1975) suggestion that "*individuals seek out opportunities to behave in ways that allow them to feel competent*" (pp.61). The problem is that whilst such a suggestion gives a broad explanation for behaviour in apparently free-choice situations, it does not help to define the conditions under which individuals will feel competent. Indeed, Experiments 4 and 5 in chapter two suggest that even when participants may have resolved a goal, this will not necessarily predict their subsequent behaviour. Achievement-goal theorists such as Dweck and Leggett (1988) and Nicholls (1984) have offered useful theories, but it has been difficult to apply these to the results observed in chapter 3, and particularly Experiments 4 and 5. It might be that further research is required to more precisely identify models that can incorporate factors that help to predict when individuals are, and are not, likely to experience competence.

4 The effects of improvement on task enjoyment

In chapter 3, the change of experimental design from Experiment 1 to Experiment 2 involved manipulating task outcomes so that individuals improved at the task. The rationale for the change in design hinged on the speculation that improvement was an ecologically more valid type of experience compared to consistently achieving the same level of performance. For example, in a novel task situation, it would be unlikely that performance over trials would not change to some degree. In light of this speculation, two changes were made to the Success condition in Experiment 2. Firstly, the final success rate was amended from 70% to 90%. Secondly, instead of participants achieving a consistent success rate of 70%, participants improved from an initial success rate of 50% to the new final success rate of 90%. A comparison of the interest and enjoyment ratings between Experiment 1 and Experiment 2 revealed that there was a marginal positive (but non-significant) increase in ratings. However, given that two variables had been manipulated, it was not clear whether this increase was due to a) the improvement schedule or, b) the fact that participants reached a higher overall performance level (i.e. 90%).

4.1 Improvement as an independent variable

The purpose of the experiments to be reported in this chapter was to investigate whether improvement is an important factor in determining people's experiences of a task. For example, does achieving a consistent 70% success rate result in the same types of experiences of a task as improving from an initially low success rate, up to a final success rate of 70%. Is it the process towards success that influences subsequent experiences, or is it just the final level of success that determines subsequent experiences?

A recent review of the literature suggests surprisingly little research on improvement as an independent variable. A recent Psyclit search using the Psyclit databases from 1988-1999 for journal articles using "improvement" in the title produced 532 hits but none related to improvement as an independent variable. Instead, improvement has been used as an adverb for a dependent variable -- for example, performance improvement (e.g. Scott, Scott and Goldwater, 1997); quality improvement (Elliott, 1994) – and, most frequently, in terms of patient rehabilitation outcomes e.g. improvement after treatment (e.g. Markowitz et al, 1996); clinical improvement (e.g. Coplan et al, 1997) and cognitive improvement (e.g. Lorusso, Poli and Casmiro, 1994; Nathanson and deFaria; 1993).

The purpose of the experiments in chapter 4 was therefore to investigate whether individuals who gradually improved at a task would enjoy it more than a group who did well throughout the task.

4.1.1 The role of success-expectations in mediating experiences of tasks.

Because improvement, as a phenomenon in its own right, does not appear to have been researched in any systematic fashion, it may be useful to speculate on its likely effects on task experiences. One possible analysis starts from the assumption that when an individual performs at some level at a task, this sets the level of their expectation for future success. Thus, when they perform at a higher level on their next attempt, it seems reasonable to assume that they will have surpassed their prior expectation. For example, if an individual initially performs at a success-rate of 90%, but then continues to perform to this level, they probably will be meeting their expectations of success, but not surpassing them. However, if they initially perform at a success rate of 70% but then, on their next trial (or set of trials), they achieve a success rate of 80%, they probably will have not only met their expectations, but surpassed them. Although this hypothesis might seem speculative, the considerable work started by Tolman (1932); extended by Atkinson and colleagues (e.g., Atkinson, 1957; 1964; Atkinson and Litwin, 1960; see also Weiner, 1992, pp.180-201 for a review) and Feather (see later for references), seem to provide good evidence that current levels of performance do indeed influence expectations for future success. A review of a typical experiment in this line of research might help to clarify the argument.

In a study by Feather (1966), participants were split into four groups and all four groups were given fifteen anagrams to solve. Half of the participants were told that 70% of students were able to solve all the anagrams (high expectation-of-success condition), whilst the other two groups were told that only 30% of students were able to solve all the anagrams (low expectation-of-success condition). For two of these groups, the first five of these anagrams were unsolvable (initial-failure condition); for the other two groups, the first five anagrams were extremely easy (initial-success condition). The subsequent 10 anagrams were of average difficulty (pilot testing has suggested that on average, students solved 50% of these anagrams). This meant that on the subsequent ten anagrams, participants were expected to (and did) solve approximately half of them (i.e., they experienced an actual average success-rate of 50%). Prior to attempting each anagram, participants were asked to estimate the probability that they would solve it. The average estimated probability for solving each anagram was then broken down by condition.

A within-subjects ANOVA that included "trials" as a factor revealed a highly significant main effect for trials that was qualified by a trials x initial-experience interaction. This interaction revealed that participants who initially experienced failure began to increase their estimates of success as they progressed through the trials, whilst participants who initially experienced success did the opposite. Presumably, this was because after the initial success/failure, their subsequent success rates on the next 10 anagrams were similar (i.e., 50%). Thus, for participants who experienced initial failure, subsequent estimates moved up, whereas for those who experienced initial success, estimates went down. The above findings have received support from several of Feather's other studies - that is, when individuals experience success before a problem (e.g. Feather, 1966, 1968) or are told to expect success (Feather, 1959, 1961, 1963a), they are more likely to report that they will do well on the next problem.

But does surpassing expectations result in participants experiencing tasks more positively? Some evidence suggests that it does. For example, (Feather, 1963b; 1968) has shown that confidence in future success correlated positively with measures such as low anxiety. (Unfortunately, Feather did not include a mood/enjoyment measure in his 1965 study). Similarly, Remedios, Lieberman and Benton (*in press*) have shown that when individuals surpass their outcome-expectations (in this case, the grade they expected to receive on a course), they report more positive experiences of the course (see also Greenwald and Gilmore, 1997a; 1997b; Holmes, 1982). In this study, participants were asked to rate their enjoyment of their course at four different points, pre-course, after they received the mid-semester test results, after they received their grades for their projects, and then again the following semester after they had received their final grades. The results from this study revealed that the degree to which participants surpassed their initial grade-expectations positively correlated with their reports of interest and enjoyment at the different points in the course.

The above analyses suggest that improving at a task should result in increasing expectations of success as individuals progress, compared to when they simply achieve a consistent performance level. Also, because it has been shown that when individuals surpass their prior expectations, this leads to more positive experiences of that task (e.g. Remedios et al, *in press*;. Feather, 1963b), it is also reasonable to expect that individuals who improve should rate the task as more enjoyable, relative to a group who achieves a consistent level of performance. It is this hypothesis which was investigated in the 5 experiments in this chapter.

4.1.2 The experimental task

The key requirement of the task for the experiments in this chapter was that it be possible to precisely control levels of success as the experiment progressed. The anagram task was not suitable for this purpose, because it was not possible to control success levels with the required degree of precision. However, the first stockmarket experiment used in Experiments 1-3 did allow precise control of outcomes, and so a task very similar to the stockmarket task was devised for the improvement experiments.

For Experiments 6-10, participants were introduced to the task as one that was investigating the relationship between personality and implicit learning. The sign-up posters explained to them that implicit learning was the phenomena whereby individuals were able to learn rules and apply them without being aware of the rule they had learned. The implication was that their personality influenced how well they learned tasks. They were told that their task was to detect the pattern in a series of events. They were presented with a sequence of 30 outcomes (e.g. left or right) and then asked to predict the next 10 outcomes in the sequence. The task was devised with the intention that the solution should be conceived of as possible (i.e. getting all 10 predictions right), but that when participants failed to achieve the complete correct sequence, they would not think that the computer was controlling their success. Pilot testing of this task with various levels of success revealed that participants did not suspect the outcomes to be manipulated by the computer/experimenter except when the overall success rate dropped to 20%. Therefore, although some experimental phases in Experiments 6-10 contained situations where participants' performance dropped to as low as 10% (e.g., Experiment 10), participants never ended up scoring less than 40%.

4.1.3 Dependent measures

The purpose of the experiments in this chapter was to investigate the effects of improvement compared to achieving a consistent performance. Originally, the intention was to apply the same suite of dependent measures that were used for Experiments 1-5. However, the use of the persistence measures proved to be impractical for the following reasons.

Firstly, if participants were allowed to persist for as long as they wanted during an initial persistence phase and were allowed to improve to a pre-determined limit, say 70%, then in the latter trials of this initial persistence phase, they would essentially be achieving the same rate of success as the comparison group who were to achieve a consistent 70% success rate. One way to resolve this problem would be to ensure that participants just kept improving. But in an unlimited persistence situation, by definition, participants can just persist for as long as they want to. Therefore, it would be impossible to ensure that participants kept improving because if they persisted for long enough, they would eventually reach 100% success. If this were to happen, it would be unclear whether task enjoyment was due to participants having fully completed the task or because they had improved.

For the free-choice period, the concern was whether performance during the free-choice period would contaminate the tightly controlled performance outcomes from the experimental phase. In the improvement group, what level of success should be set for any persistence after the experimental phase? Should they continue to improve, in which case they would not really be experiencing the same level of success as the consistent 70% group? Or should they achieve the same level of success throughout the free-choice period - e.g. 70%, - in which case, they would be failing to meet improvement expectations?

In either scenario for the improvement group during the free-choice period, future persistence with the task would violate the criteria for each of the experimental conditions. It was therefore decided that it would be impractical to use persistence as a dependent measure for the improvement experiments.

However, the answers to the questionnaire items used in the stockmarket (Experiments 1-3) and anagram (Experiments 4 and 5) experiments eventually proved useful in demonstrating differences in participants' experiences between the different groups. These questionnaires were therefore used in experiments 6-10. Two questions were excluded, the attribution question and the stop questions regarding persistence. The attribution was excluded because it was felt that this question had been too ambiguous for participants and had not yielded any useful information.

4.1.4 Change of Experimenter

There were potential problems in having me as the experimenter running Experiments 6-10. One was that by the time Experiments 6-10 were conducted, I had completed several other experiments similar to those in chapter 3 in which participants had subtly received different levels of success information. The population pool from which participants were taken was the same for the improvement experiments as it was for the previous experiments. This meant that there was a good chance that participants who signed up for the improvement experiments would have participated in my earlier experiments. Moreover, since all participants were debriefed in the earlier experiments and told that the experiments were about individuals' reactions to success and failure, participants might well have been suspicious of possible deception when volunteering for other experiments of mine. Additionally, having had a wardennial role in the student halls of residence, many students already knew the nature of my research.

Indeed, for some of the experiments that do not appear in this thesis, several participants' had specifically made comments such as "I hope this is not going to be another experiment where you make me do badly".

A second factor in changing to a different experimenter was that by having an experimenter blind to condition, this would help to control for any experimenter effects. As outlined in chapter 1 (see section 1.8), Page (1981) has highlighted the problems of demand characteristics influencing participants' responses to tasks. For example, Orne (1971) has suggested that even when experimenters try to control their responses to participants, they still cannot control for the subtle, unintended cues they give to participants by tone or inflection of voice, or by posture. Although communication with participants was kept to a minimum in the experiments in chapter 3, it is possible that my actions provided participants with subtle cues as to whether they had done well or badly, especially at the moment when I administered the questionnaire. Also, as I had to set up the computer for the stockmarket experiments, and would have done so for the improvement experiments, I always knew the condition that the participants were in. In short, I was never blind to experimental hypotheses or conditions.

For these reasons, an experimenter was recruited who was blind to the purposes and hypotheses of the experiment. She was given the minimum information to enable her to run the experiment. When participants completed the experiment, they were told that if they wished to know the purposes of the experiment, then they were to come and see me. The purpose of the experiment was only revealed to the experimenter after the final experiment had been run.

4.2 Experiment 6: Establishing the internal validity of the experimental task

Given that the only dependent measures that were to be used in the experiments in this chapter were the answers to the questionnaire items, it was felt important to check that these dependent measures were sensitive to the experimental manipulations. Therefore, the first experiment in the improvement series did not actually include an improvement condition. Instead, as with the first stockmarket experiment (Experiment 1), participants achieved a constant success rate of either 30% or 70%. It was felt that if Experiment 6 failed to show differences in enjoyment between the groups when the success rates were so different, then it was unlikely that differences would emerge in later experiments when the final success rates between the groups were the same.

4.6.1 Experiment 6 - Methods

Overview This was a two-condition experiment in which participants were shown a sequence of 30 outcomes and were asked to predict what the next 10 items in the sequence would be. At the end of each prediction sequence, participants were told how successful they had been in their previous 10 predictions. Outcomes were manipulated so that participants achieved either a 70% or 30% success rate. Participants were allowed 6 attempts to predict the correct sequence. When participants had completed these 6 trials, they were given a questionnaire to assess their experiences of the task.

Participants 24 participants (4 males, 20 females) were recruited via two sources, either the Stirling University Psychology department's subject panel or a financial inducement of £2.

Apparatus The experiment was presented on a Viglen ENVP15P PC with software program for the prediction program written in Pascal Turbo.

There were two identical testing rooms. Each was a small windowless room with two tables and two chairs. On one table was a computer and on the other was a box marked "completed questionnaires". Next to the "completed questionnaires" box was the Rosenberg (1965) short-form self-esteem questionnaire (see Appendix 6). This questionnaire was intended to corroborate the cover story of the experiment being about "personality and implicit learning" (see procedure section below) and participants' responses were not analysed.

Procedure Participants were recruited for the experiment via a sign-up poster which advertised the study as an investigation of the effects of *"Implicit learning and Personality"*. The poster read as follows:

> Implicit learning is the phenomena whereby people are able to learn patterns without being able to say what the actual patterns they have learned are. In this experiment, you will be asked to look at a sequence of letters and suggest what patterns are occurring.

All participants for this study were tested individually in separate rooms, though two people could be tested at the same time. Prior to participants arriving for their appointment, the experimenter randomly assigned the participant to either the 70% or the 30% group. As the experimenter was naive to the software program, she was given a series of written instructions as to how to load the program that ran the experiment. Also, as the experimenter was blind to the experimental conditions, the conditions were coded as CP (constant 70%, P = positive feedback condition) or CN (constant 30%, N = negative feedback condition) so as not to give any clues to the experimenter as to the purpose of the experiment.

When the participant arrived for their appointment, they were greeted by the experimenter and taken into the room where they were to be tested. The participant was asked by the experimenter to complete the self-esteem questionnaire, told to put the questionnaire in the box when they had finished, and then let the experimenter know that they were ready to start the experiment. When the participant indicated that they were ready to start the experiment, the experimenter took them back into the room and seated them by the computer. The experimenter then gave the participant a set of written task instructions, told them to read the instructions very carefully, and then to complete the task. The participant was told to come and get the experimenter if they were unsure what to do or ran into any problems. The experimenter then left the room. The task instructions read as follows:

Task Instructions

Thank you for agreeing to take part in this experiment

Implicit learning is the phenomena whereby people are able to learn patterns without being able to say what the actual patterns they have learned are. The task you are about to complete is one such implicit learning task where you have to try and predict patterns in a sequence.

There are two possible outcomes, LEFT or RIGHT. You will firstly be presented with a sequence of 30 outcomes, one after the other. Just sit back and try to look for patterns in the sequence.

Then, the computer will ask you to try and predict what you think the next 10 items in the sequence will be.

YOU then make your predictions as to what you think the next 10 items will be. Don't worry if you feel you have no idea, the computer will tell you how you did.

In total, you will be given the chance to make SIX sets of predictions.

If this is all clear to you, then press the spacebar to start the program.

When the participant pressed the spacebar, the computer presented 30 Left or Right outcomes to the participant. Each outcome was displayed for 2.5 seconds. The sequence of 30 outcomes were pre-programmed and consisted of an equal number of left and right outcomes in a random sequence. All participants saw the same 30 outcome sequence. At the end of the sequence, the computer generated the following message to prompt the participant to make their predictions.

You should now make your predictions for the next 10 items

in the sequence.

Press "L" for a LEFT prediction and "R" for a RIGHT prediction.

Press the spacebar if you are ready to make your predictions.

The sequence of predictions made by the participant were displayed on the computer screen as the participants made their predictions. When the participant made their 10th prediction, the computer displayed how successful they had been.

You got *n* predictions correct

Press the spacebar and make another 10 predictions

The number of predictions that participants were told they got right was the independent variable. Although both groups were to achieve a consistent success rate, it was felt that if participants achieved a constant 7 or 3 out of 10 success rate, they would become suspicious of the validity of the feedback. Unlike the stockmarket experiments, participants in these sets of experiments were <u>explicitly</u> told their overall performance levels. For example, rather than receiving trial by trial feedback and having to work out their average performance, as they had done in the stockmarket experiments, participants actually had their performance displayed to them after each prediction sequence. The 70% success rate group achieved the following outcome feedback over the 6 sets of trials: 767877. The 30% success rate group were given the following outcome feedback over the 6 sets of trials: 343233.

When the participant had been given the final set of outcome feedback, the computer presented the message that the experiment had finished and that they were to go and get the experimenter.

The experimenter then took the participant back into the room and gave them the task experiences questionnaire to complete. The participant was told to complete the questionnaire, to place it in the completed questionnaires box and then to come and see the experimenter. The questionnaire read as follows:

Your experiences

How much 'FUN' would you say that task was?

Lots of Fun	6	5	4	3	2	1	Not Fun

How much did you Enjoy solving the problems?

A lot	6	5	4	3	2	1	Not at all

How interesting was the task as a whole?

Very interesting	6	5	4	3	2	1	Uninteresting

On the whole, how difficult was the task?

Difficult	6	5	4	3	2	1	Easy

How well did you think you did at the task?

Well	6	5	4	3	2	1	Badly

How well do you think others would do at the task?

Well	6	5	4	3	2	1	Badly

Please note down any particular rules or patterns you think you saw emerging in the sequences. Don't worry if you didn't notice any, just put *none*

When participants had completed the questionnaire, the experimenter verbally asked them about their experiences of the task, especially about how well they thought they had done and how much they had enjoyed doing the task. The experimenter made a mental note of the answers given by the participant and recorded these on a separate sheet when the participant had left the testing area. After giving the participant the appropriate reward for participating in the experiment (e.g. money or Psychology Department participation credit), the participant was thanked for their time and told that if they wanted to know any more about the experiment, to come and see me.

4.2.2 Experiment 6 - Results

In total, 26 (4 males, 22 females) participants were tested. On the advice of the experimenter, two participants' data was excluded from the analysis as the experimenter thought that these participants had not understood the task instructions.

All analyses were completed using two-sample t-tests except for the WellY vs. WellO analysis which was analysed using a within-subjects t-test.

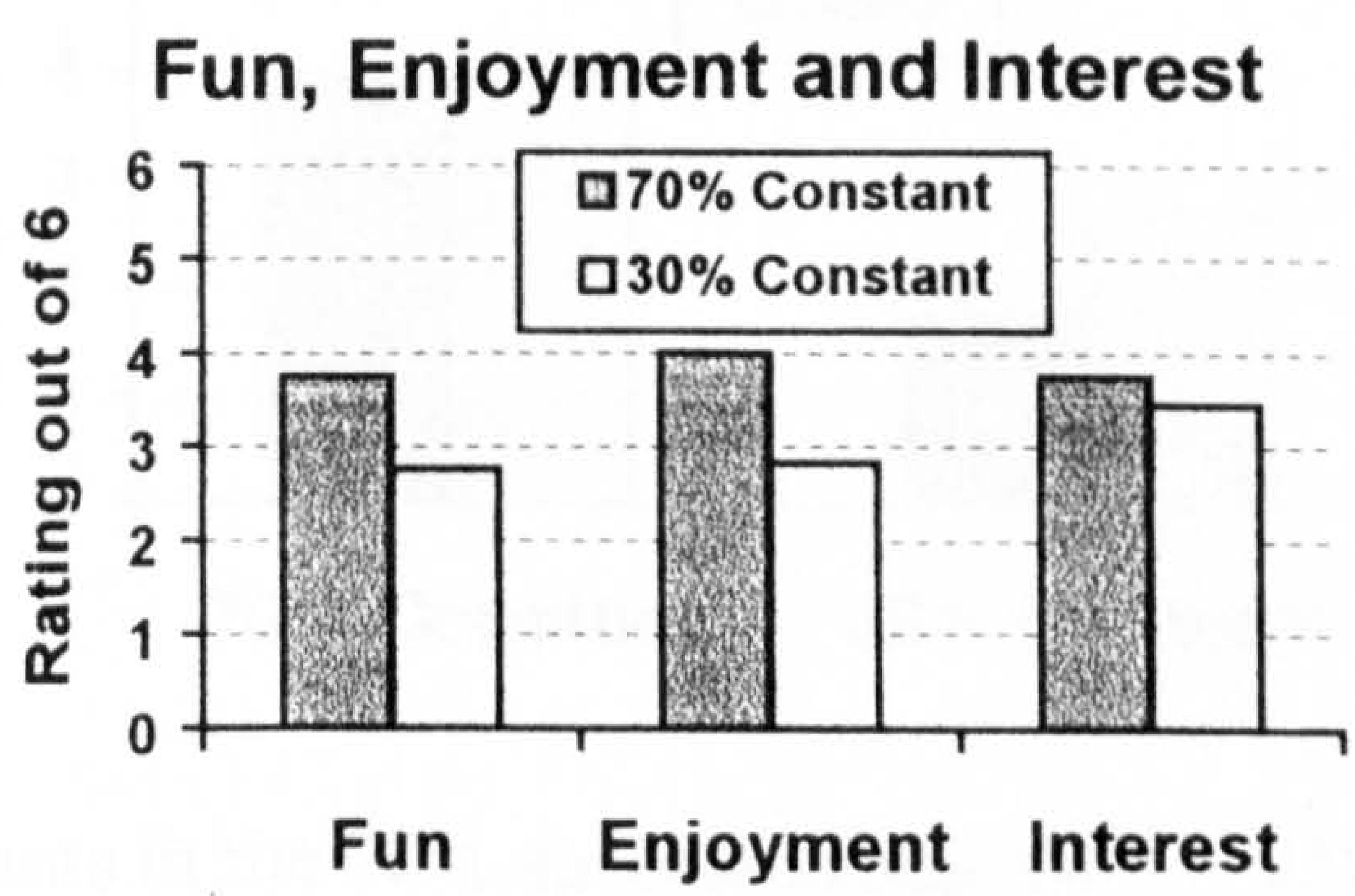
Table 9 overleaf shows that for all measures, the 70% success rate group rated the task more positively compared to the 30% success rate group. Significant differences were observed for the dependent measures for fun, enjoyment, task difficulty, WellY and WellO. Also, participants in the 30% success group rated WellO as significantly higher than WellY. The results for each dependent measure are dealt with individually in the following sections.

Table 9: Means and standard deviations (in italics) for all dependent measures for Experiment 6 (n=24).

	70% success		30% success	
	Mean	***s.d.***	**mean**	***s.d.***
Fun	3.7*	*1.06*	2.7*	*0.87*
Interest	3.7	*1.22*	3.4	*1.0*
Enjoyment	4.0*	*1.13*	2.8*	*1.19*
Task Difficulty	2.7*	*0.88*	1.7*	*0.87*
WellY	3.9**	*0.79*	1.7**	*0.75*
WellO	4.0*	*0.95*	2.7*	*0.88*

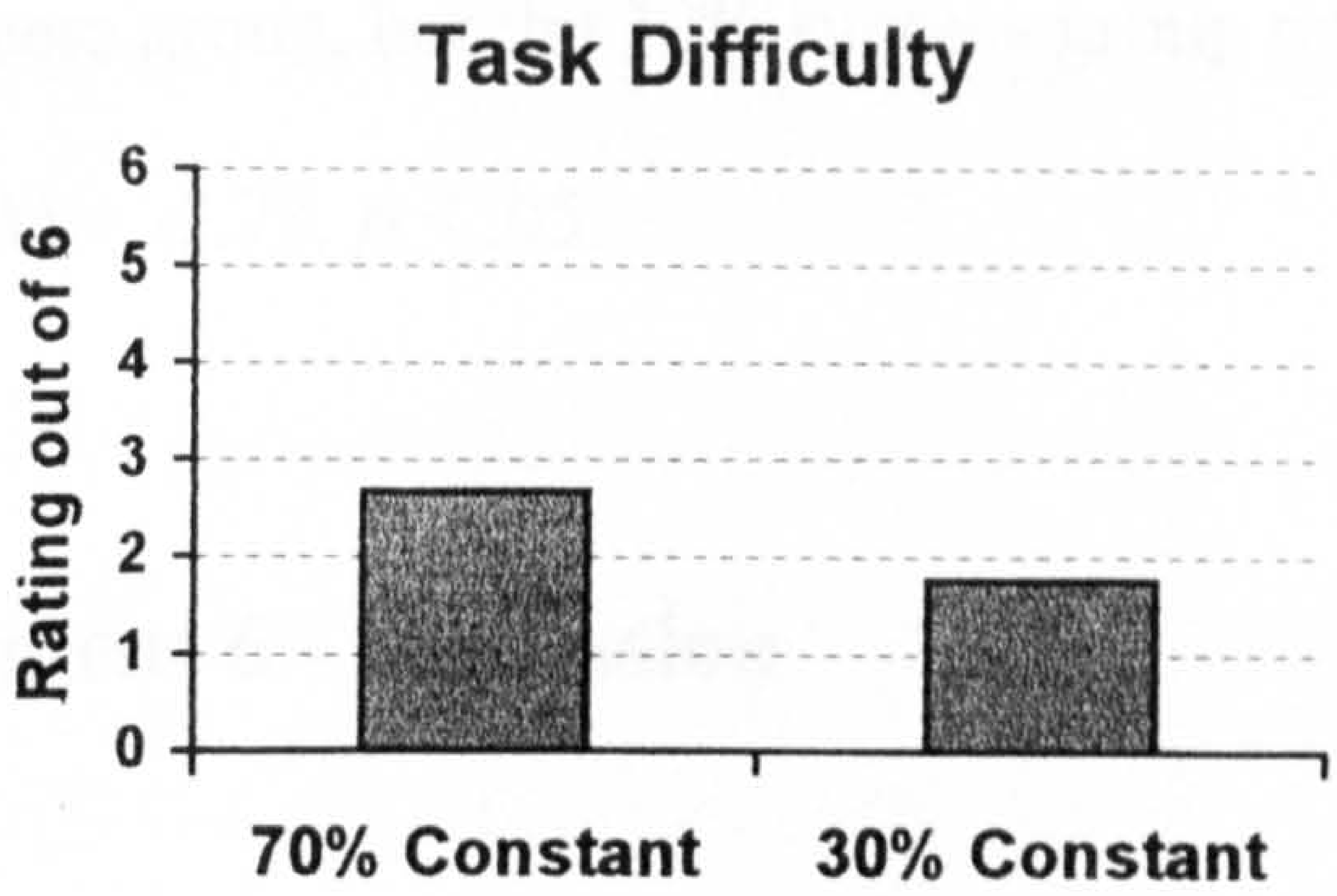
* $p < .05$ ** $p < .001$

4.2.2.1 Experiment 6 - Fun, Enjoyment and Interest



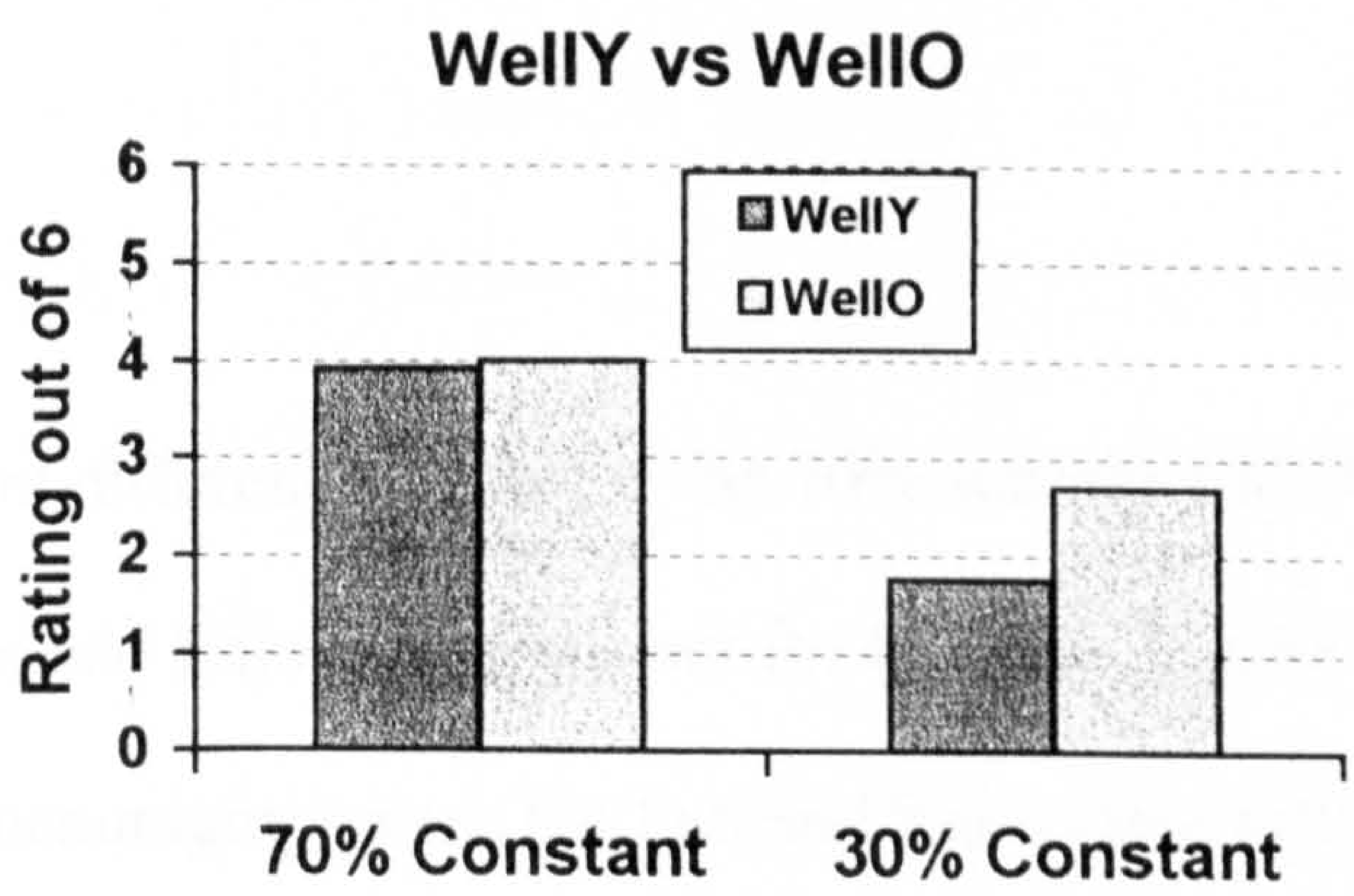
Analysis showed that there were differences across all three measures with the 70% success group rating the task more positively compared to the 30% success group. These differences were significant for the dependent measures of Fun, t (22) = 2.54, $p < .05$ and Enjoyment, t (22) = 2.46, $p < .05$, but not for Interest.

4.2.2.2 Experiment 6 – Task Difficulty



Participants in the 70% success group reported the task as significantly easier, $t\,(22) = 2.56$, $p < .05$, compared to those in the 30% success group.

4.2.2.3 Experiment 6 – WellY vs. WellO



<u>WellY</u> Participants in the 70% success group (M = 3.92) ratings of their performance (WellY) were significantly higher, $t\,(22) = 6.86$, $p < .001$ compared to those in the 30% success group (M = 1.75).

<u>WellO</u> Participants in the 70% success group (M = 4.0) rated how they thought others would do at the task (WellO) significantly higher, $t\,(22) = 3.55$, $p < .001$, than those in the 30% success group (M = 2.67).

WYWO There was no significant within-group difference between WellY and WellO for the 70% success group, but the 30% success group rated WellO significantly higher than WellY, $t\,(11) = -2.73, p < .05$.

4.2.3 Experiment 6 - Discussion

Apart from the lack of significance between the 70% success group and the 30% success group for the dependent measure of Interest, significant differences between the two groups emerged for all the other dependent variables. A brief discussion of these results is now included.

Task experiences

There were significant differences between the 70% success and 30% success groups for the ratings on Fun and Enjoyment, but not for Interest. Unlike the first stockmarket experiment, it was encouraging to see the Fun and Enjoyment ratings separating out in Experiment 6 under essentially the same success rates as the first stockmarket experiment.

There are several speculations that could be offered why this was the case. Firstly, although there was no specific emphasis for individuals to do well at the task, it might have been that the task had a performance goal. In the previous chapter, several arguments were made to determine whether or not participants were mastery or performance-focused. Recall that according to Nicholls (1984), the mere fact that a task is introduced as a skill one should be sufficient to induce participants to be concerned about their performance. In Experiment 6, the task was presented to participants as one where they had a performance target to attain, namely, to try and work out the underlying pattern of behaviour. So, even though participants were told that the experiment was one that examined implicit learning, and that they might not even be aware of the patterns they were learning, it would seem reasonable to suspect that these participants would have been concerned with their performance. Unlike the stockmarket and anagram experiments where participants were told the experimenter was only interested in their experiences of the task, participants in Experiment 6, (and the subsequent experiments in this chapter), were not distracted away from performance, or given any reason to suspect that it was solely their performance the experimenter was interested in. Thus, it would seem reasonable to suspect that participants in the improvement experiments were performance-focused. But why should simply being performance-focused have caused the differences in enjoyment between the two groups? Recall that in section 3.3.3, the argument was made that because a performance-focus is thought to heighten the awareness of competence cues, both success and failure are likely to be experienced more acutely. This would therefore explain why those participants who performed well rated their fun and enjoyment experiences significantly more positively relative to those who performed poorly.

Secondly, a comparison of how the 70% and 30% groups rated task difficulty and the WellY questions might also help explain why the enjoyment and fun ratings were significantly different. Although there was a significant difference between the 70% and 30% success groups for how difficult they thought the task was, the actual ratings were very low (i.e., M for 70% = 2.67 vs. M for 30% = 1.75). This suggests that participants in the 70% success group, although performing relatively well, recognised that the task they were doing was a difficult one. When subsequently asked how well they thought they had done at the task, the difference in mean ratings between the two groups was again significant, but this time, participants in the 70% success group ratings were much higher than the 30% group (i.e., M for 70% = 3.92 vs. M for 30% = 1.75). A summary of these results is therefore that participants in the 70% success group, although rating the task as difficult, nevertheless thought they had done quite well at the task. However, for the 30% group, they not only rated the task as very difficult, they also thought they had done very poorly at the task.

A third possible explanation for why there were significant differences in Fun and Enjoyment ratings between the two groups is that, unlike the stockmarket and the anagram tasks, participants were explicitly told the level of success they were achieving. Thus, participants only had to assess whether they thought performing at a 70% success rate was good performance or poor performance, they did not have to ascertain what level of performance they were achieving. With the knowledge that they were actually performing at a 70% success rate, then they only had to decide whether this was good or poor performance. Recall that in the Anderson and Rodin (1989) study, participants who were told they had scored in the 55th percentile reported being "disappointed but not devastated". If it is reasonable to suspect that achieving a success rate of 70% would be experienced as being above the 55th percentile, then it is unsurprising that participants would be pleased with their performance. Similarly, if it is reasonable to suspect that achieving a success rate of 30% would be experienced as being below the 55th percentile, then again, it would be unsurprising if participants would have been unhappy with their performance. Thus, the fact that participants were explicitly told how successful they had been might have removed any ambiguity about their performance levels and contributed to their subsequent ratings of their task experiences.

Within group-comparison of WellY vs WellO

In keeping with Experiments 1-5, there was no significant difference between the WellY and WellO ratings for the 70% success group but there was for the 30% success group. This data provides further support for the suggestion that when participants fail at a task, they experience a loss in self-confidence. Indeed, the particularly low task difficulty ratings for this experiment indicated that all participants recognised the task to be a difficult one, but it was still only participants in the lower success group who seemed to experience losses in self-confidence.

Conclusions

The purpose of Experiment 6 was to establish whether the experimental design for the improvement experiments was sensitive enough that participants would rate the task differently across the different levels of success. This was especially important as the only dependent measures used in these set of experiments were the questionnaire items, and when the 70% success rate vs. a 30% success rate was run in Experiment 1, although differences in persistence emerged, differences in task enjoyment did not. The results from Experiment 6 suggested that the experimental design was an appropriate one for testing whether improving at a task affected individuals' enjoyment.

4.2 Experiment 7 – Controlling for average success rates

In Experiment 6, participants who achieved a success rate of 70% reported significantly more positive experiences of the task than participants who achieved an average success rate of 30%. This finding suggested that the methodology employed in Experiment 6 was appropriate for participants to experience the task in different ways when they achieved differing levels of success. Having established that the methodology employed was appropriate, Experiment 7 was the first experiment that compared a group who improved to a group who achieved a constant success rate.

The main purpose of the experiments in this chapter 4 was to investigate whether or not improving to a level of success was more enjoyable than constantly attaining the same level of success. However, in endeavouring to manipulate the variables of a constant level of success and improvement, a third variable needed to be controlled for, namely, average success. For example, say over six blocks of trials, participants improved to 70% by attaining the following sequence of scores; 234567. In this case, the average success rate would be 45% (group A). Now if the constant 70% success group (group B) achieved the following sequence of success, 767787, the average success for this group would be, of course, 70%. If subsequently, group B rated the task more positively than group A, then a potential explanation for this finding could be that the constant success group achieved a higher average success rate. Similarly, if there were no differences between the groups, any improvement effects might have been masked by the fact that this group achieved a lower average success rate. That is, the lower success rate might have lead them to rate the task less positively, but that the improvement meant that some of the negative effects of not achieving a high average success rate were neutralised. In this case, there would be the potential of making a Type I error by concluding that the improvement had no effect. Either way, the average success rate would clearly be a potential confound when trying to interpret the subsequent results.

The purpose of Experiment 7, therefore, was to establish whether average success rates mediated task enjoyment when participants improved at a task. In this experiment, all participants achieved the same average success rate (i.e. 50% success) but one group improved to 70%, whilst the other group achieved a constant 50% success rate.

4.3.1 Experiment 7 - Methods

Participants 26 participants (8 males, 18 females) were recruited via two sources, either the Stirling University Psychology department's participants' panel or a financial inducement of £2.

Procedure Experiment 7 was run in identical fashion to Experiment 6. The only change was in the feedback that was displayed to participants at the end of each block of 10 predictions. In Experiment 7, the constant 50% success group achieved the following sequence of explicit success feedback, 545655, whereas a 70% improvement group received feedback of 345567.

4.3.2 Experiment 7 - Results

In total, 26 participants were tested. On the advice of the experimenter, two participant's data were excluded from the analysis. In one case, the participant reported that they were not enjoying the experiment and left the testing room before completing the 6 trials. Another participant, after completing the experiment accidentally pressed the key that took them to the screen that displayed the Pascal program code. Within this code, there was a brief narrative explanation of the different conditions. Although it was not clear whether the participant had read this screen (or had made sense of it), it was felt appropriate to exclude this participant's data. Both of these participants were in the constant success group.

All analyses were completed using two-sample t-tests except for the WellY vs. WellO analysis which was analysed using a within-subjects t-test.

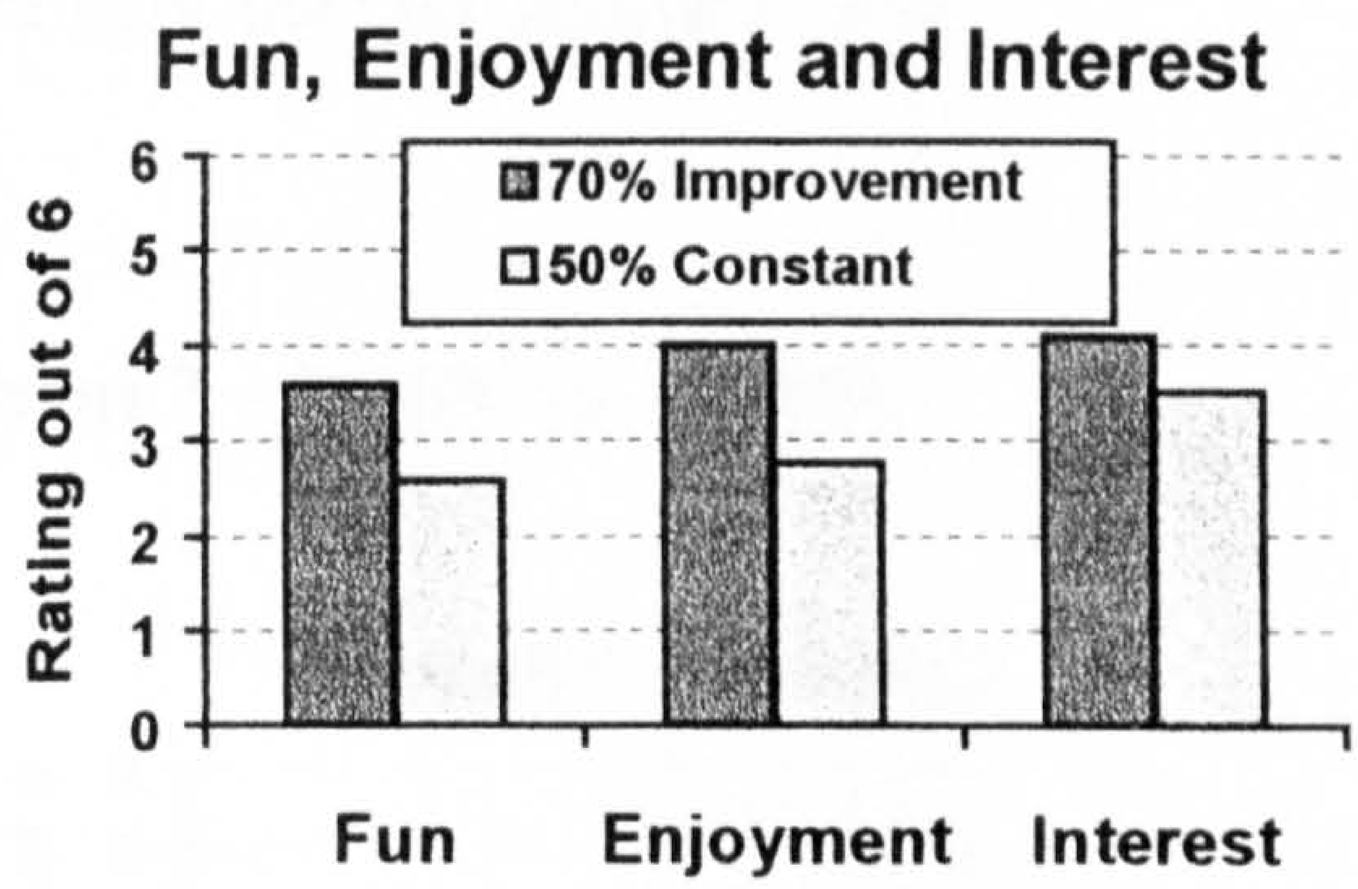
Table 10: Means and standard deviations (in italics) for all dependent measures for Experiment 6 (n=24).

	50% constant		**70% improvement**	
	mean	***s.d.***	**mean**	***s.d.***
Fun	2.6*	*1.16*	3.6*	*0.90*
Interest	2.7*	*1.42*	4.0*	*1.35*
Enjoyment	3.0*	*1.41*	4.1*	*0.99*
Task Difficulty	2.2	*1.03*	2.5	*0.91*
WellY	1.8*	*0.72*	3.6*	*1.51*
WellO	2.7	*1.15*	3.6	*0.91*

* $p < .05$ ** $p < .001$

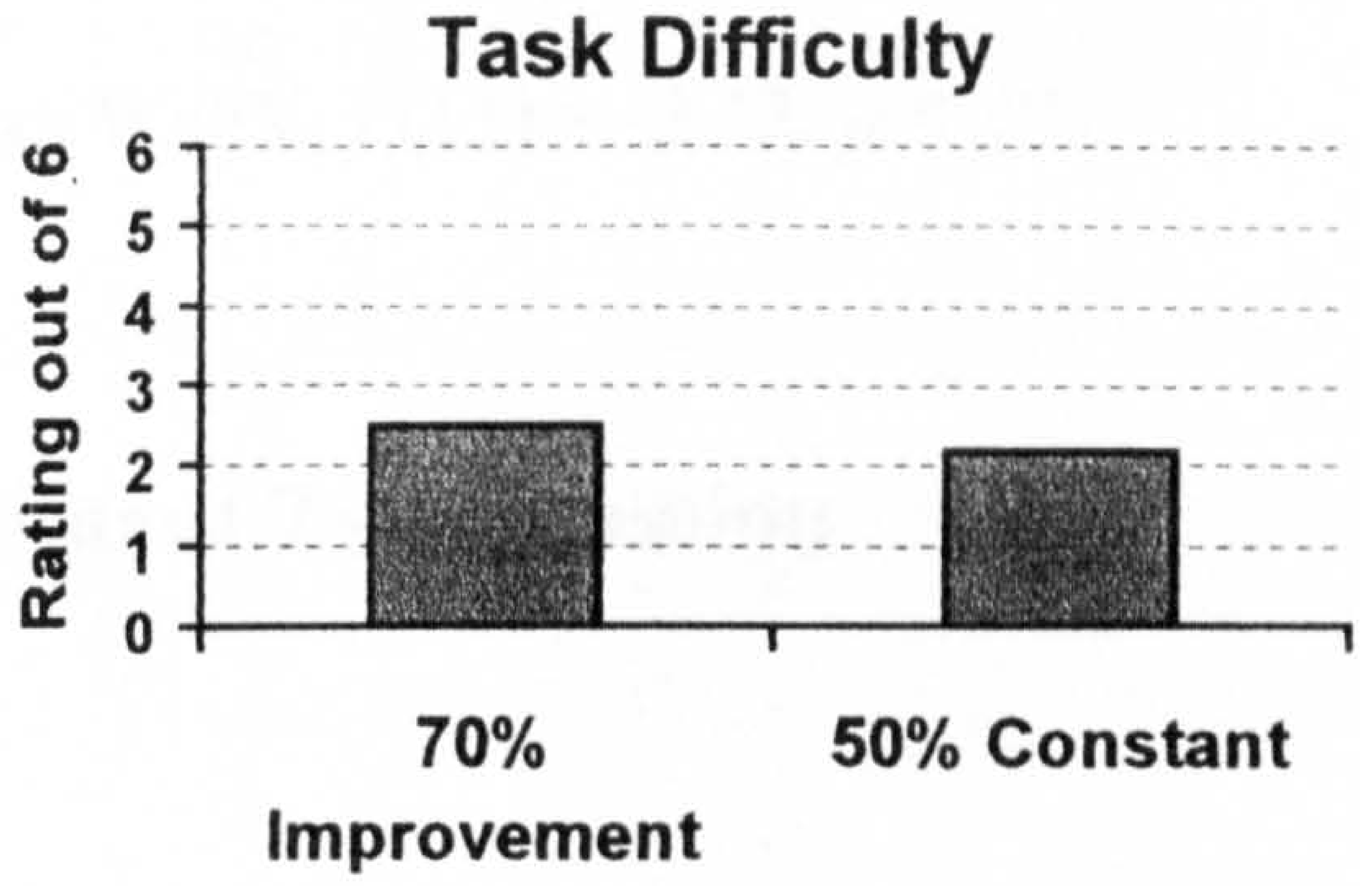
Table 10 shows that participants in the 70% improvement group rated the task more positively compared to those in the 50% constant success group. There were significant differences in the ratings for the dependent measures of fun, enjoyment, interest, and how well participants thought they had done at the task (WellY). The difference between WellY and WellO was significantly different for the 50% success group, but not for the 70% improvement group.

4.3.2.1 Experiment 7 - Fun, interest and Enjoyment



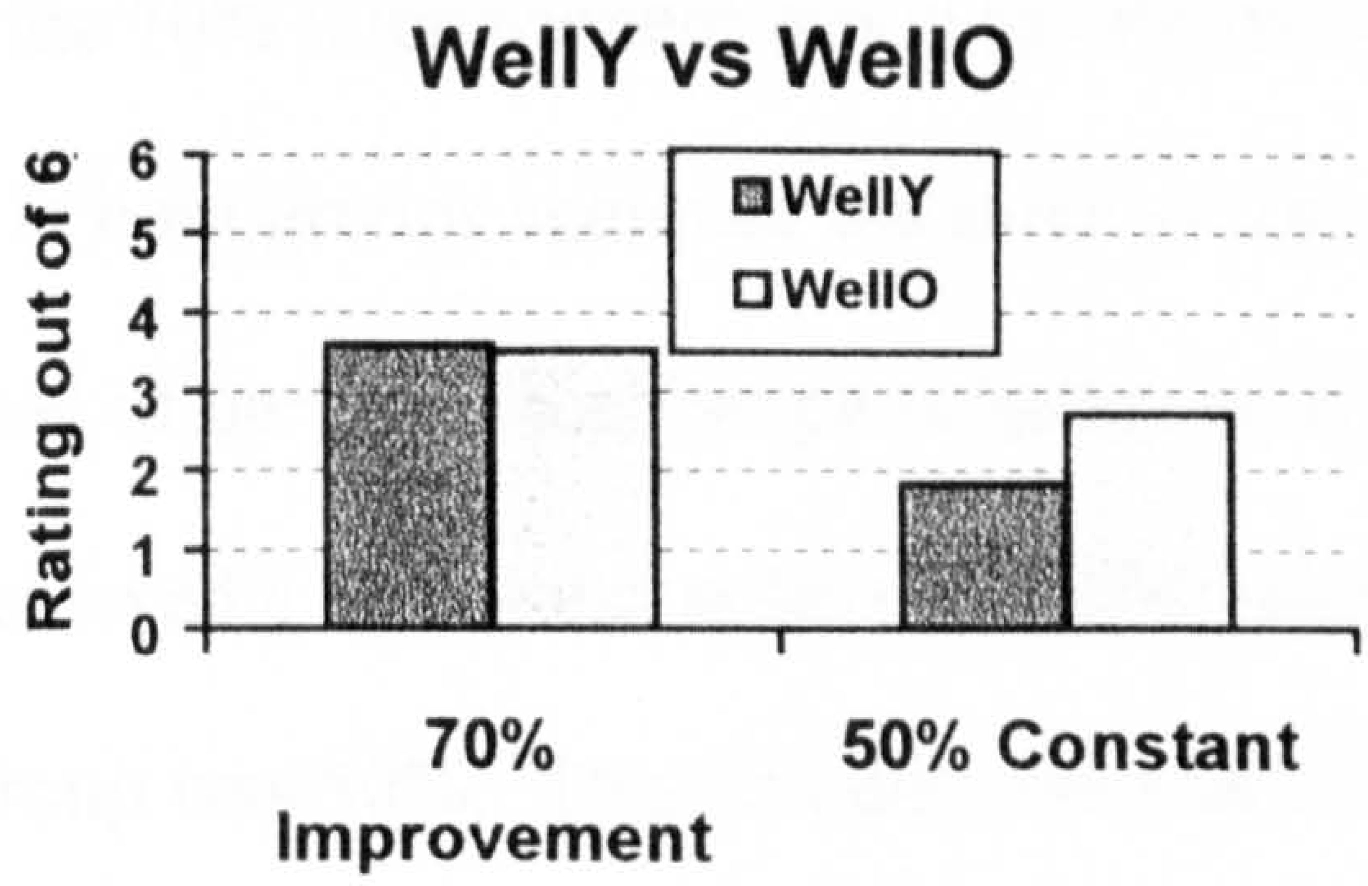
There were significant differences between the 70% improvement group and the 50% improvement group for the dependent measures of Fun, t (22) = 2.35, $p < .05$, Enjoyment, t (22) = 2.21, $p < .05$ and Interest, t (22) = 2.17, $p < .05$, whereby the improvement group who improved reported significantly higher mean ratings of the task across all the dependent measures.

4.3.2.2 Experiment 7 – Task Difficulty



There were no differences between the improvement group and the constant 50% group for how difficult they thought the task was (p = .41).

4.3.2.3 Experiment 7 – WellY vs. WellO



WellY Participants in the 70% improvement group rated how well they had done at the task (WellY) significantly higher, t (22) = 3.64, $p < .01$, than those in the 50% constant group.

WellO There was no significant difference between the 70% improvement group and the 50% constant group for the mean ratings of how well participants thought others would have done at the task (WellO).

WYWO There was no significant within-group difference between WellY and WellO for the 50% constant group but the 70% improvement group rated WellO significantly higher than WellY, t (11) = -2.12, $p < .05$.

4.3.4 Experiment 7 - Discussion

The results from Experiment 7 revealed that participants who improved to a final success rate of 70% rated the task more positively than participants who achieved a constant success rate of 50%. Thus, it would seem that achieving the same average success rate does not necessarily result in participants experiencing tasks in similar ways. However, it would be premature to conclude that it was the improvement that caused participants in the 70% improvement group to rate the task more positively. Although participants in both groups achieved the same average success rate, they differed in the final level of performance that participants achieved. The 70% improvement group achieved a final success level of 70%, whereas the final level of success for the 50% group was 50%. Thus, it could have been the final level of performance that individuals achieved that actually determined their task experiences.

There is support for this suggestion when the results for the remaining dependent measures are reviewed. For example, although there were no differences in the task difficulty ratings for the two groups, participants in the 70% final-outcome group rated their performance significantly higher than those in the 50% final-outcome group (M = 1.83). Also, in a similar fashion to Experiment 6, participants in Experiment 7 who experienced the higher final outcome of 70% did not rate WellY significantly different to WellO (WellY = 3.58 vs WellO = 3.50), but participants who experienced the 50%

final outcome did (WellY = 1.83 vs WellO = 2.67). So, at the 50% final success rate, participants thought they had done badly compared to others, whereas at the 70% success rate, they did not. Although this was not predicted, when viewed in terms of high success vs. low success, it is reasonable to see the 50% success group as a low success group, and so far in this thesis, low success groups have generally evidenced losses in self-confidence. Again, this supports the suggestion that differences in task experiences might have been due to the different final levels of success, even though the difference in final outcome was only 20%.

Conclusions

Although Experiment 7 provided evidence that participants rated tasks more positively when they improved, this finding was confounded by the final level of success that participants achieved. The question therefore remains as to whether participants enjoyed the task more because they improved or because they achieved a higher overall total? One way to test this hypothesis would be to compare a group who achieved a constant success rate of 70% with a group who improved to 70%.

4.4 Experiment 8 - Experiment 8: Improvement to a success rate of 70% vs. achieving a constant 70% success rate

Although Experiment 7 provided evidence that participants rated the task more positively when they improved during a task, it was unclear whether the higher ratings were due to participants improvement, or the fact that they had achieved a higher final success rate compared to the comparison group.

Experiment 8 was designed to control for the final level of success by ensuring that both groups achieved the same final success rate. This final success rate was set at 70%. This level of success was chosen because in Experiment 6 and 7, whenever participants achieved a success rate of 70%, there were no differences between WellY and WellO, suggesting that at this level of success, participants seemed to be satisfied with their performance. In Experiment 8, participants either achieved a consistent 70% success rate (average = 70%) or improved to a 70% success rate (average = 50%). A potential confound still remained in that participants in the constant success group would experience a higher average success relative to the improvement group, so that if the constant group were to report greater enjoyment of the task, this could be attributed to this higher average success rate. However, if the opposite were found, then this could be seen as powerful evidence for the effects of improvement, because these effects would have been seen to occur even when participants achieved a lower average success rate. Indeed, an argument could be made that support for the null hypothesis would also be marginal support for the effects of improvement. This is because the difference in average success rate between the groups was 20% and so if the improvement group still rated the task as positively than the group who achieved the consistent 70% success rate, they would be doing so despite experiencing a much lower average success rate.

4.4.1 Experiment 8 - Methods

Participants 28 participants (8 males, 20 females) were recruited via two sources, either the Stirling University Psychology department's participants' panel or a financial inducement of £2.

Procedure Experiment 8 was run in identical fashion to Experiments 6 and 7. The only change was the manipulation to the feedback that was displayed to participants at the end of each block of 10 predictions. In Experiment 8, the constant 70% group achieved the same sequence of explicit success feedback (i.e., 767877) as was used in Experiment 6. For the improvement group, participants received the same sequence of explicit success feedback that was used in Experiment 7 (i.e., 345567).

4.4.2 Experiment 8 - Results

In total, 28 participants were tested. On the advice of the experimenter, four participants' data were excluded from the analysis. In two cases, the participants reported not believing the computer outcomes and in another two cases, in a similar fashion to Experiment 6, the experimenter thought that these participants had not understood the task instructions. All of these participants were in the improvement group.

All analyses were completed using two-sample t-tests except for the WellY vs. WellO analysis which was analysed using a within-subjects t-test.

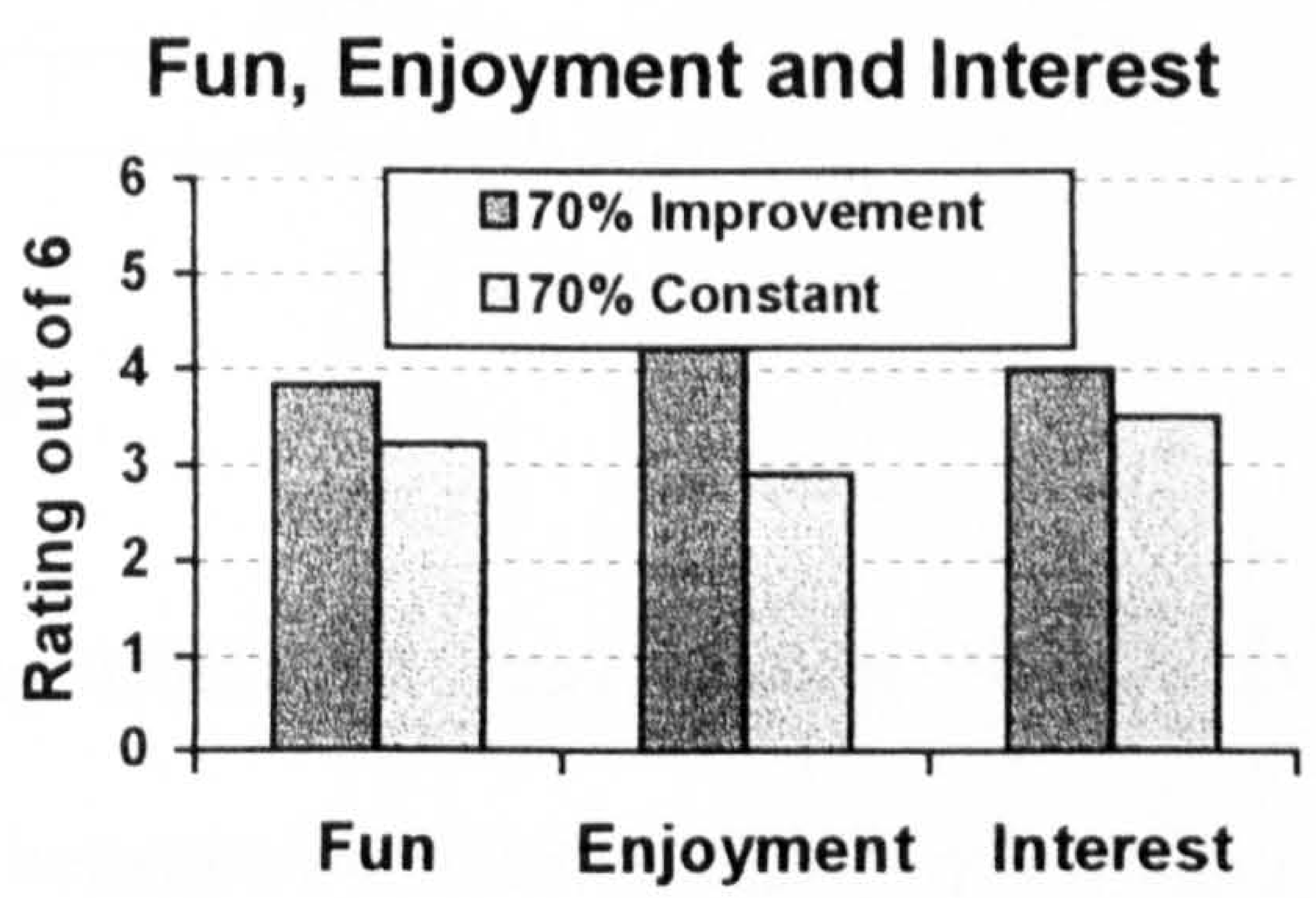
Table 11: Means and standard deviations (in italics) for all dependent measures for Experiment 7 (n=24).

	70% constant		70% improvement	
	Mean	***s.d.***	**Mean**	***s.d.***
Fun	3.2	*1.11*	3.8	*1.03*
Interest	3.5	*1.62*	4.2	*0.87*
Enjoyment	2.9*	*1.08*	4.0*	*0.85*
Task Difficulty	2.2*	*0.96*	3.2*	*1.22*
WellY	3.3	*1.15*	3.9	*1.0*
WellO	3.7	*1.3*	4.1	*0.67*

* $p < .05$

Table 11 above shows that participants in the 70% improvement group rated the task more positively compared to the 70% constant success group. However, the only differences to reach significance were for Enjoyment, $t\ (22) = 2.72, p < 0.5$, and Task Difficulty, $t\ (22) = 2.23, p < 0.5$.

4.4.2.1 Experiment 8 - Fun, Enjoyment and Interest



There were no significant differences between the 70% constant success group and the 70% improvement group for the dependent measures of Fun, Enjoyment and Interest.

As the trends were that the 70% improvement group ratings on all three of these measures were higher than the 70% constant group, I examined whether the lack of significant results might have been due to the insensitivity of the individual dependent measures to reflect genuine differences in the experiences of the groups. For example, Tables 12 and 13 below shows that there were highly significant correlations between the ratings for these three measures.

Table 12: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 70% constant group in Experiment 8 (n =12).

70% Constant	Fun	Enjoyment	Interest
Fun	--	.54*	.80**
Enjoyment	.54*	--	.54*
Interest	.80**	.54*	--

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

Table 13: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 70% improvement group in Experiment 8 (n =12).

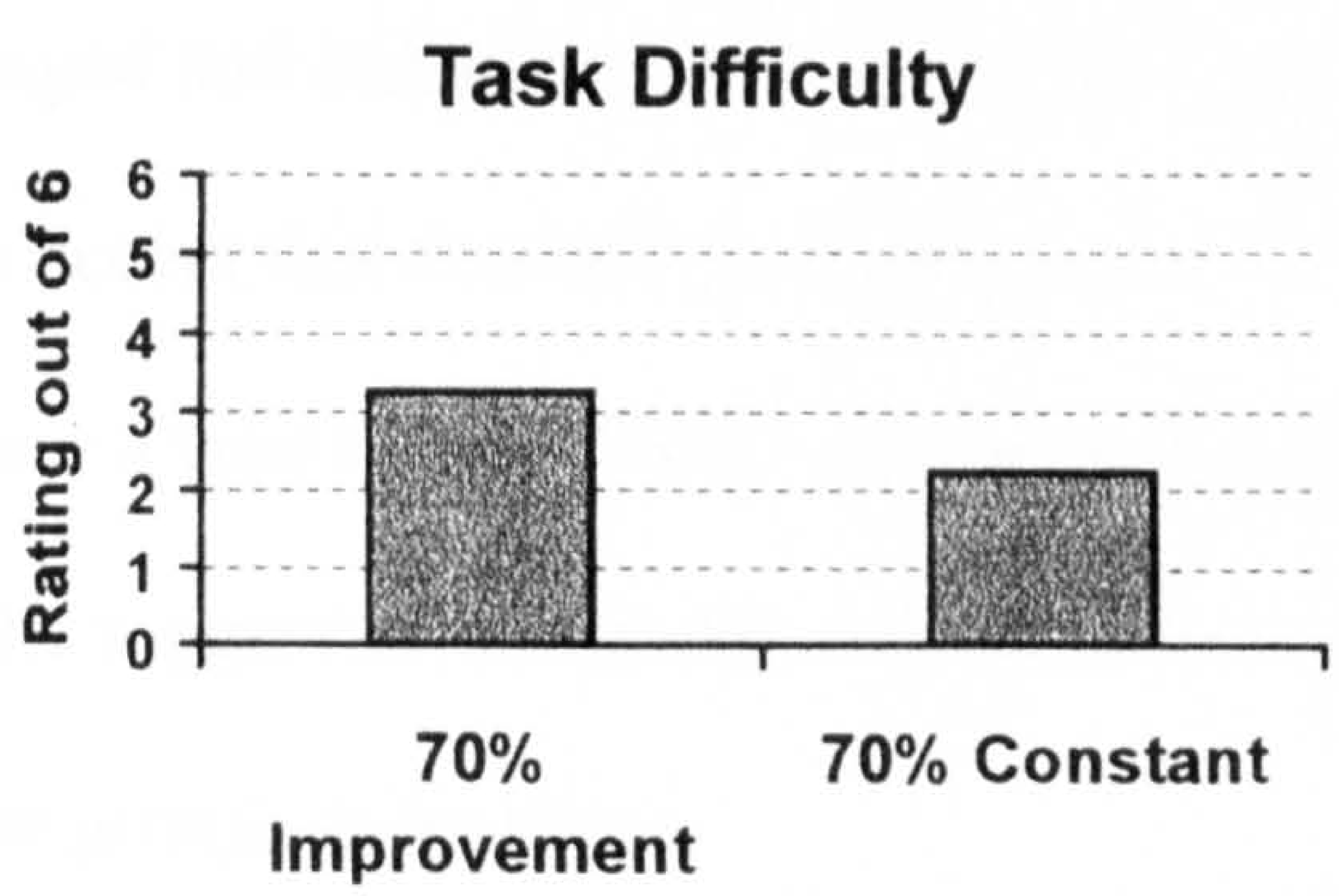
70% Improvement	Fun	Enjoyment	Interest
Fun	--	.93**	.76**
Enjoyment	.93**	--	.74**
Interest	.76**	.74**	--

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

Additionally, a one-way ANOVA run against these three measures revealed no significant differences between the three ratings in either group (70% constant, $p = 0.54$;

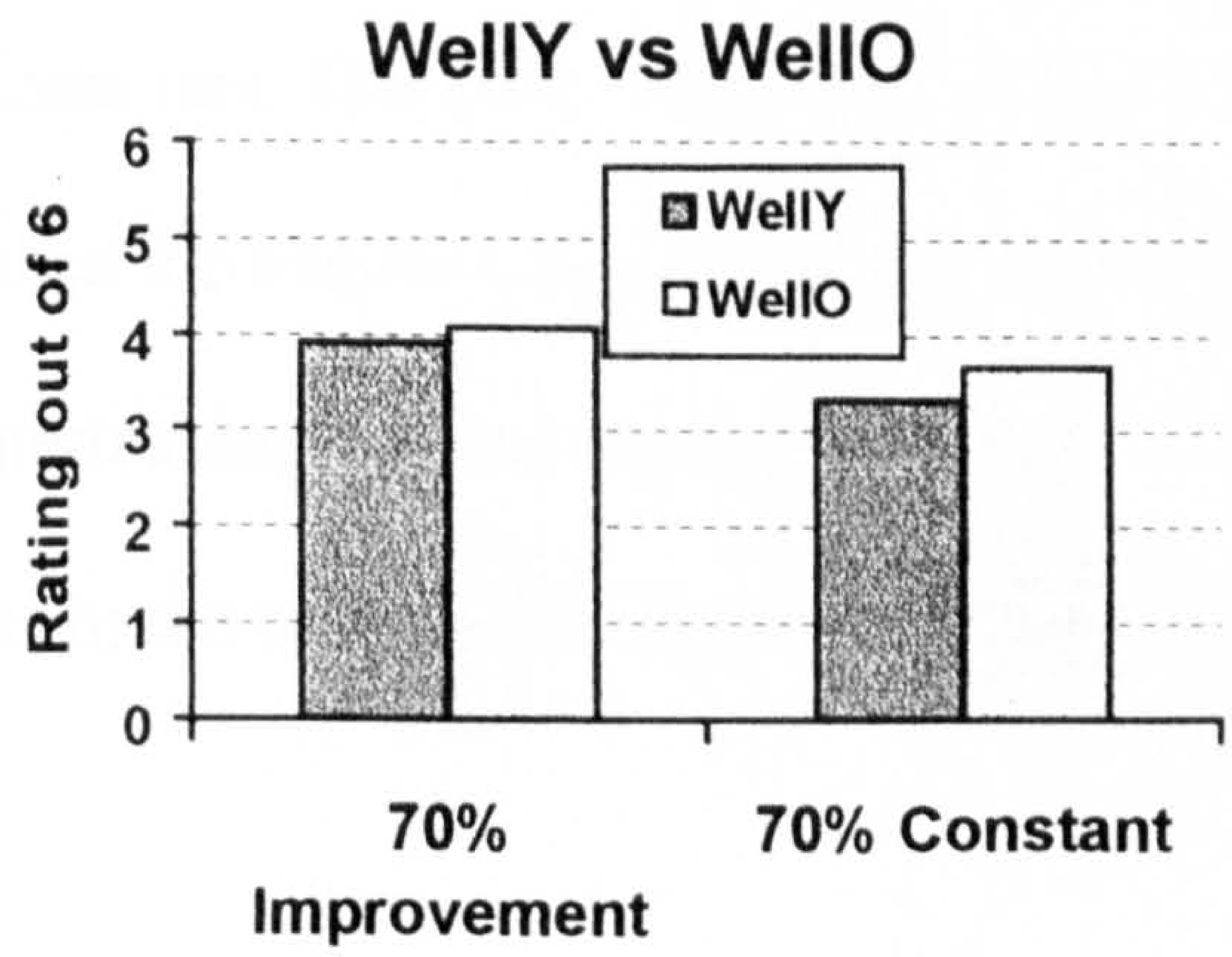
70% improvement, $p = 0.55$). Such a finding suggests that these three measures essentially measure the same underlying construct that could be labelled "Enjoyment". Using this logic, the three ratings were averaged for each individual and then a t-test compared the averages for the improvement and the constant group. Using this procedure, significant differences now emerged, with the 70% improvement group (M = 2.8, s.d. 1.32) rating the task significantly more positively than the 70% constant group (M = 3.2, s.d.1.11), $t\,(11) = 2.05, p < .05$.

4.4.2.2 Experiment 8 – Task Difficulty



Participants in the 70% improvement group reported the task as significantly easier, $t\,(22) = 2.23,\ p < .05$, compared to participants in the 70% constant group.

4.4.2.3 Experiment 8 – WellY vs. WellO



There were no significant differences between the groups for WellY and WellO ratings. There were also no within-group differences between WellY and WellO for either the constant 70% or the 70% improvement group.

4.4.3 Discussion

The hypothesis that participants who improved would report more positive experiences of the task compared to a group who achieved a consistent success rate was not supported significantly until the ratings for the dependent measures for Fun, Enjoyment and Interest were averaged and compared between the two groups. The question therefore is whether to accept this data and potentially commit a Type II error, or reject it and potentially commit a type I error. There are two arguments that could be made that give credence to the suggestion that it might be more reasonable to accept the differences between the groups as genuine.

Firstly, as outlined in the introduction to Experiment 8, there was a potential confound in this experiment in that the improvement group experienced a lower overall success rate. Thus, despite experiencing a lower average success rate, participants in the improvement group still rated the task at least as positively as the group who achieved the consistent 70% success rate. Secondly, the improvement group experienced the task as significantly easier than the constant success group. Again, this finding occurred despite participants experiencing a lower average success rate. These findings strongly suggest that participants in the two groups genuinely experienced the task in different ways.

Additionally, the findings for task difficulty also help to make a case that the improvement group really did enjoy the task more than the constant group. For example, despite both groups achieving the same final level of success, the improvement group thought the task was (significantly) easier. One reason for this might have been that for the improvement group, the gradual improvement gave them confidence that they were doing well at the task, and had they had more trials at the task, then they would have solved the problem. For the consistent success group, their only realisation was that they were performing well at the task, but that they had reached a performance ceiling. Thus, for the improvement group, the task might have seemed easier because they had had more reason to believe they would eventually solve it.

Another possible consequence of the gradual improvement was that participants (in the improvement condition) might have attributed their original poor performance not to the task's difficulty, but to the fact that they had not learned a certain set of rules. It might have seemed to them that as they tried out new strategies, these strategies were having some positive effect on their performance. This might have suggested to them that they were solving rule-based problems. And because they were continually improving, this might have given them confidence that they could solve future rules. For the constant group, this would not have been the case. Although experiencing a relatively high level of performance initially, their lack of improvement gave them no reason to believe they would perform any better in the future. After their initial success, it is reasonable to suspect that they attempted several strategies to improve their performance in the next five blocks of trials. Continual failure to improve might have suggested to them that the next rule that had to be assimilated was a difficult one and clearly one that was outwith their current level of skill.

To summarise, participants in the improvement group might have enjoyed the task more because their continual improvement gave them confidence that they could solve future problems. On the other hand, participants in the consistent group would have had less confidence in their ability to solve future problems because their attempts to do so had not been successful. The differences in terms of task difficulty tentatively suggest that for participants in the improvement group, their expectations of future success were high, but this was not so for participants who initially achieved a high level of performance.

Conclusions

Although in Experiment 8, significant differences emerged between the improvement and constant groups' ratings, these only occurred when the dependent measures of Fun, Enjoyment and Interest were averaged. The fact differences emerged only when the measures were combined suggests a need for caution before concluding that participants who improved experienced the task more positively.

One reason why differences did not emerge more clearly between the groups for the key dependent measures of fun, enjoyment and interest could have been that the degree of improvement was not sufficiently dramatic for participants to really feel that they had improved. In Experiment 8, participants firstly achieved 3 out of 10, then 4, then 5, then 5, then 6 and finally 7 out of 10. For them, improvement was gradual rather than dramatic. Their small incremental gains in performance might have indicated to participants that they were performing well, but not that they were making dramatic gains. If an individual performed at 20% on one block of trials and then 40% on the next block of trials, such a jump in performance might be more likely to increase expectations of future success. This would also suggest that changing the difference between initial success rate and final success rate might also influence the degree of participants' success expectations. For example, rather than starting at a success level of 30%, it might be useful to start at one success level lower (i.e., 20%). This would increase the difference between initial success and final success from 40% to 50%.

4.5 Experiment 9 – Improvement to 70% vs. Constant 70% using a different improvement schedule from Experiment 8

Experiment 9 aimed to replicate Experiment 8 except that this time, the increase in overall success rate from block 1 to block 6 was larger, and at two points in the improvement sequence, the increase was greater than one unit. In Experiment 8, the success sequence for the improvement group was 345567 for the 6 blocks; in Experiment 9, it was 245577. Thus, the increase in performance between block 1 and block 2 was 20%; between block 4 and 5 was 20% and the overall increase between block 1 and block 6 was 50%. This compares against the graded increase of 10% and overall increase of 40% for Experiment 8.

4.5.1 Experiment 9 - Methods

Participants 30 participants (12 males, 18 females) were recruited via two sources, either the Stirling University Psychology department's participants panel or a financial inducement of £2.

Procedure Experiment 9 was run in identical fashion to the previous experiments in this chapter. The change this time was the manipulation of the feedback that was displayed to participants at the end of each block of 10 trials/predictions. In Experiment 9, the constant 70% success group achieved the same sequence of success (i.e. 767877) as was used in Experiment 8. For the improvement group, participants received the sequence 235577.

4.5.2 Experiment 9 - Results

In total, 30 participants were tested. On the advice of the experimenter, 6 participants' data were excluded from the analysis. In all cases, the experimenter reported that participants were suspicious about the validity of the computer outcomes. All of these participants were in the improvement group.

All analyses were completed using two-sample t-tests except for the WellY vs. WellO analysis which was analysed using a within-subjects t-test.

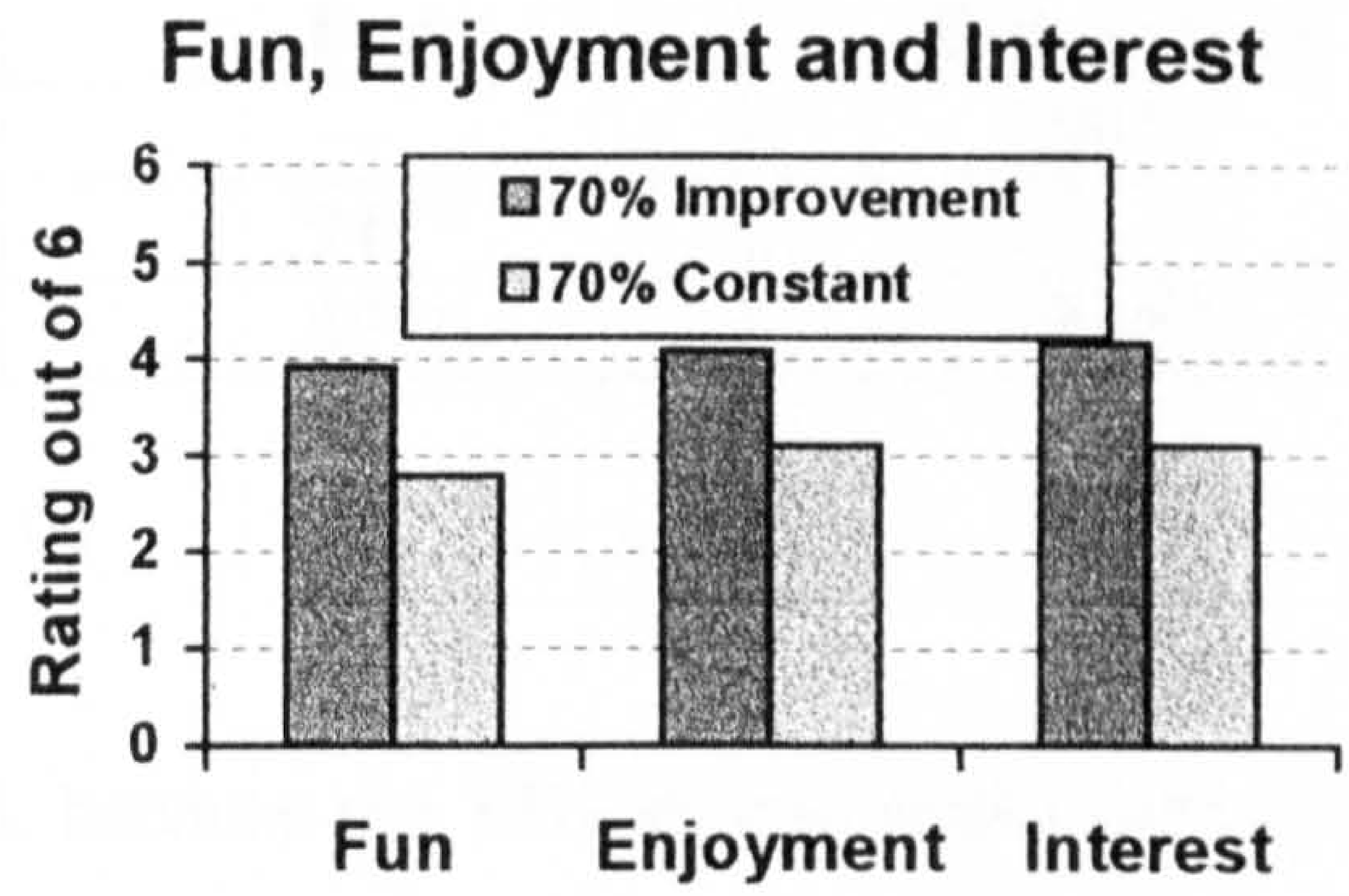
Table14: Means and standard deviations (in italics) for all dependent measures for Experiment 9 (n=24).

	70% constant		70% improvement	
	Mean	***s.d.***	**mean**	***s.d.***
Fun	2.8	*1.4*	3.9	*1.44*
Interest	3.1**	*0.79*	4.1**	*1.62*
Enjoyment	3.1	*1.11*	4.2	*1.03*
Task Difficulty	2.4	*1.16*	3.2	*1.85*
WellY	2.9	*0.51*	3.5	*1.17*
WellO	3.2	*0.97*	3.4	*0.79*

** $p < .01$

Table 14 above shows that participants in the 70% improvement group rated the task more positively compared to the 70% constant success group. The only difference to reach significance was for the measure of Interest, $t\ (22) = 2.89, p < .01$.

4.5.2.1 Experiment 9 - Fun, Interest and Enjoyment



There were no significant differences between the 70% constant success group and the amended improvement group for the dependent measures of Fun, Enjoyment and Interest. However, in a similar fashion to Experiment 8, Tables 15 and 16 below shows that there were highly significant correlations between the ratings for these three measures.

Table 15: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 70% constant group in Experiment 9 (n = 12).

70% Constant	Fun	Enjoyment	Interest
Fun	--	.79**	.56**
Enjoyment	.79**	--	.60**
Interest	.56**	.60**	--

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

Table 16: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 70% improvement group in Experiment 9 (n =12).

70% Improvement	Fun	Enjoyment	Interest
Fun	--	.94**	.87**
Enjoyment	.94**	--	.86**
Interest	.87**	.86**	--

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

As with Experiment 8, because the 70% improvement group ratings on all three of these measures were higher than those of the 70% constant group, and a one-way ANOVA run against these three measures revealed no significant differences between the three different ratings (e.g. 70% constant, $p = 0.73$; 70% improvement, $p = 0.90$), the three ratings were averaged for each individual and then a t-test compared the average scores in for the improvement and the constant group. In a similar fashion to Experiment 8 which used this procedure, significant differences emerged, as the 70% improvement group (M = 4.1, s.d. 1.32), rated the task more positively than the 70% constant group"(M = 3.0, s.d. 0.94), $t(11) = 2.20, p < .05$.

4.5.2.2 All other dependent measures

There were no significant differences between the groups for task difficulty, WellY and WellO ratings. There were also no within-group differences between WellY and WellO for either group.

4.5.3 Discussion

Once again, the hypothesis that improvement should enhance individuals' experiences of a task compared to a group who achieved a consistent level of success, was not supported until the ratings for the dependent measures for fun, enjoyment and interest were averaged and compared between the two groups. However, the fact that a significant difference was again found for the combined ratings strongly suggests that this result needs to be accepted as genuine. The magnitude of the difference was not large, but it now appears to be both significant and reliable. When participants experience improvement while working on a task, there is clearly a tendency for them to enjoy it more.

One further speculation regarding why there were no significant differences between the key dependent measures in the first instance lies in the final level of success that participants achieved. The common factor for both Experiments 8 and 9 is that participants in the improvement group achieved a final success rate of 70%. Thus, one explanation for why the individual dependent measures were not significantly different may be because participants reached a final level of success that was acceptable to them. If participants were already performing at a standard that was acceptable to them, then maybe the fact they were improving at the task as well did not increase their enjoyment of the task.

One way to test the above hypothesis would be to implement the improvement schedule at low levels of success. At the 70% success rate, Experiments 1, 6, 7 and 8 have all shown that participants generally report positive ratings of the task and do not report losses in self-confidence. Experiments 1, 3, 4, 5 and 6 have shown that at low levels of success (e.g. 30% - 40%), participants rate tasks less positively and show losses in self-confidence. It therefore seems reasonable to hypothesise that when participants do particularly poorly at a task, the improvement might help them to believe that they will be successful in future trials, which in turn will help them to experience the task more positively. In contrast, a group who consistently performs poorly may not only have lower expectations of success, but their poor performance is likely to be sufficiently low for them not to experience the task positively at all.

4.6 Experiment 10 – Improvement to 40% vs. constant 40%

In Experiments 8 and 9, there was suggestive evidence that participants who improved at a task enjoyed that task more than participants who simply attained a constant level of success. The statistical technique employed, whilst valid, still suggested that any differences in ratings between the two groups had to be treated with some caution.

One possibility why clear differences did not emerge between the two groups might have been because the level of final performance was so high, this alone was accounting for a large proportion of their enjoyment ratings. In the discussion section of Experiment 9, it was suggested that the effects of improvement might be better observed when participants did not perform so well. Experiment 10 therefore tested whether improving to a lower level of final success, — specifically, achieving a final outcome of 4 out of 10 predictions correct — would cause participants who improved to rate the task more positively relative to a group who achieved a constant 4 out of 10 predictions correct.

4.6.1 Experiment 10 - Methods

Participants 24 participants (11 males, 13 females) were recruited via two sources, either the Stirling University Psychology department's participants' panel or a financial inducement of £2.

Procedure Experiment 10 was run in identical fashion to the other four experiments in this chapter. The only change was in the feedback that was displayed to participants at the end of each block of 10 trials/predictions. In Experiment 10, the sequence of success feedback for the constant 40% success group was 434544, whereas for the 40% improvement group, it was 122344.

4.6.2 Experiment 10 - Results

In total, 24 participants were tested. Unlike the previous experiments, no data needed to be excluded from the final analysis.

All analyses were completed using two-sample t-tests except for the WellY vs. WellO analysis which was analysed using a within-subjects t-test.

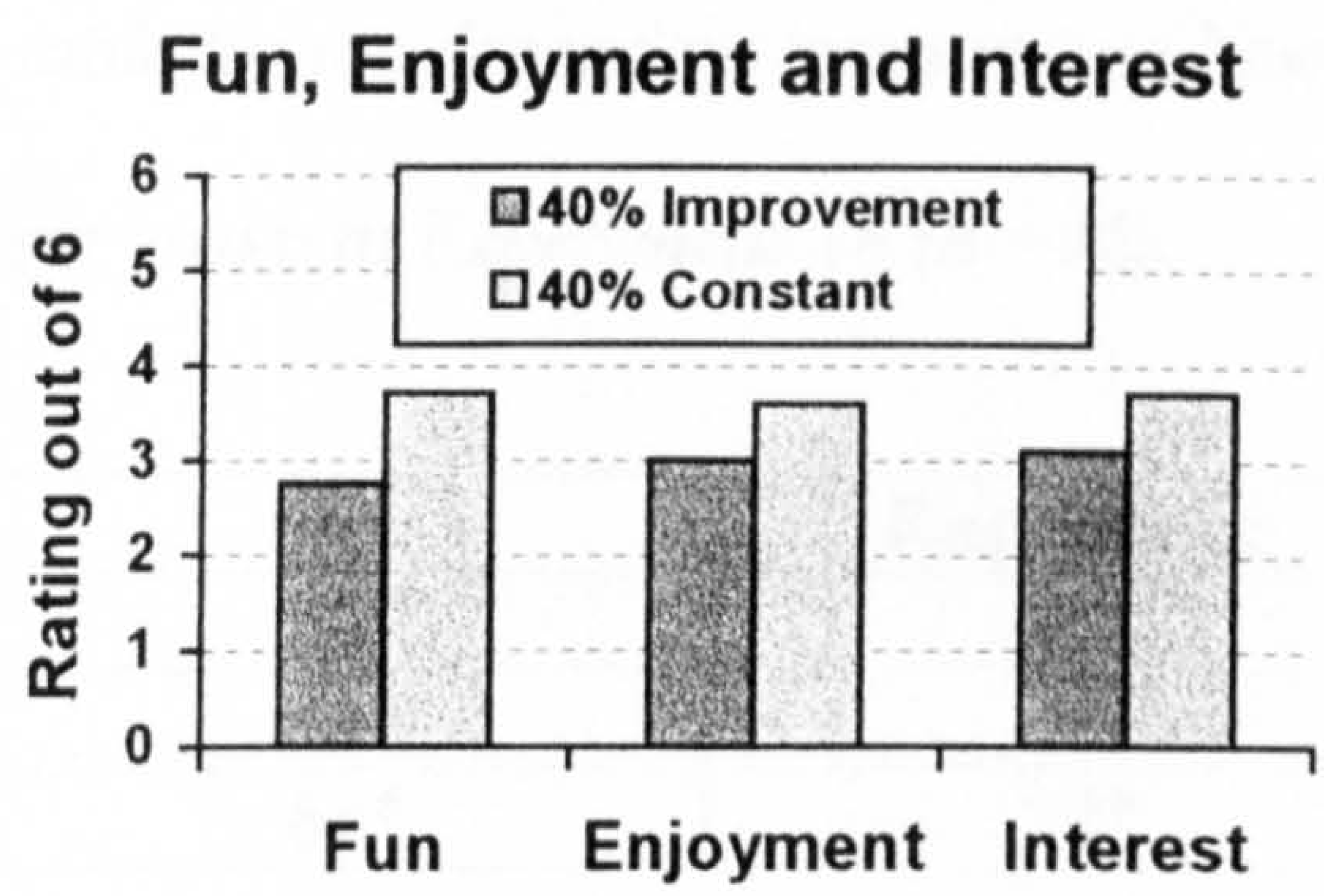
Table17: Means and standard deviations (in italics) for all dependent measures for Experiment 10 (n=24).

	40% constant		40% improvement	
	Mean	***s.d.***	**Mean**	***s.d.***
Fun	3.7*	*0.78*	2.7*	*1.14*
Interest	3.7	*0.98*	3.0	*1.13*
Enjoyment	3.6	*0.51*	3.1	*0.90*
Task Difficulty	2.2	*1.29*	2.4	*1.38*
WellY	2.2	*0.97*	2.3	*0.65*
WellO	3.2	*0.75*	2.7	*0.65*

$*p < .05$

Table 17 above shows that, in direct contrast to Experiments 8 and 9, participants in the 40% constant group rated the task more positively compared to the 40% improvement group. The only difference to reach significance was for the dependent measure of Fun, $t\,(22) = 2.30, p < .05$.

4.6.2.1 Fun, interest and Enjoyment



The only significant difference between the 40% improvement group and the 40% constant group was for the dependent measure of Fun. Indeed, contrary to Experiments 8 and 9, the trend was that it was now the constant group who rated their experiences of the task more positively, not the improvement group.

In a similar fashion to Experiments 7 and 8, there were highly significant correlations between the ratings for the fun, enjoyment and interest measures for both the improvement and the constant group.

Table 18: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 40% constant group in Experiment 10 (n =12).

40% Constant	Fun	Enjoyment	Interest
Fun	--	.53*	.55*
Enjoyment	.53*	--	.77**
Interest	.55*	.77**	--

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

Table 19: Correlation table for the dependent measures of Fun, Enjoyment and Interest for the 40% improvement group in Experiment 10 (n =12).

40% Improvement	Fun	Enjoyment	Interest
Fun	--	.64*	.64*
Enjoyment	.64*	--	.72*
Interest	.64*	.72*	--

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

To determine whether there were significant differences between these three measures, the scores were compared using one-way ANOVA's. There were no significant differences between these measures in either the constant group ($p = 0.41$) and the improvement group ($p = 0.73$). The three ratings were therefore averaged for each individual and a t-test compared the average ratings for the 40% improvement vs. the 40% constant group. Using this procedure, significant differences emerged whereby, in contrast to Experiments 8 and 9, the constant group (M = 3.6, s.d. 0.66), rated the task more positively than the improvement group (M = 2.9, s.d. 0.93), $t(11) = 2.11$, $p < .05$.

4.6.2.2 All other dependent measures

There were no significant differences between the groups for task difficulty, WellY and WellO ratings. There were no within-group differences between WellY and WellO for the 40% constant group, but there was for the 40% constant group, $t(11) = -2.83, p < .05$.

4.6.3 Discussion

Contrary to expectations, it was the constant 40% group who rated their experiences of the task more positively compared to participants in the 40% improvement group. This finding was the exact opposite to the findings from Experiments 8 and 9. Rather than improvement resulting in participants rating the task more positively, it seems that the improvement schedule had the opposite effect. Indeed, originally, the hypotheses have been (unwititngly) set up as one-tailed, because there was only one experimental hypothesis, namely, that the improvement group would rate the task more positively than the constant group. In the introduction to this chapter, no speculations were offered why the constant group should rate the task more positively. The fact that this phenomenon occurred in Experiment 10 clearly requires explanation.

One aspect of the data seems inconsistent with the conclusion that the constant group had more positive experiences of the task. This is that for participants in the constant group, their average WellY rating was significantly lower than their average WellO rating. In the improvement group, on the other hand, the two ratings did not differ significantly. This could be taken as an indication that participants in the 40% constant group did suffer losses in their self-confidence whilst those in the improvement group did not. On the other hand, given that when the key dependent measures of Fun, Enjoyment and Interest were collapsed, it was the 40% constant group who rated the task significantly more positively.

On the basis of the available evidence, it is difficult to unequivocally resolve the contradictory implications of these two measures. Suppose, however, that we were to accept the Fun-Interest-Enjoyment ratings as more likely to be a valid indication of enjoyment – how could we then explain why the constant group enjoyed the task more than the improvement group, despite obtaining the opposite findings in Experiments 8 and 9?

One possible explanation can be derived from a study by Feather (1963a). In this study, participants were split into three groups and achieved success rates of either 80%, 50% or 20%. They were given a total of 120 individual problems and after completing each problem, they were told whether or not their answer was correct. They were then asked to estimate whether they thought they were likely to get the next problem correct. Feather provides data for the number of problems participants in the various groups expected to get right (e.g. tables 1 and 2, p. 295 and 297). Over the 120 trials, participants in all three success groups (i.e. 80%, 50% and 20%), tended to overpredict the number of problems that they would get right. For example, over the full 120 trials, participants in the 80% success group were successful on 8 out of 10 trials. So, over 120 trials, they were told that they had predicted correctly on a total of 96 trials. When the number of estimations for being correct on the next trial were totalled, the average total for the 80% success group was 101.95. The mean amount of overprediction for the

80% success group was 5.95 predictions. This tendency to overpredict also occurred for the 50% success group (13.85 overpredictions) and for the 20% success group (21.1 overpredictions). This data therefore shows that although all groups tended to overpredict the number of successes they were likely to achieve, the lower the actual number of successes individuals experienced, the greater their overprediction. If the term overprediction can be interpreted in terms of success-expectation, then there might be a case that at low levels of success, individuals resist believing that they will continue to do so poorly, and as a result exaggerate their chances of success.

Feather (p.296) suggests that the participants who achieved the success rates of 20% were more "optimistic" about their future performances. However, it is particularly interesting to note that he puts the word "optimistic" in quotation marks. In retrospect, perhaps he did not mean optimistic in the sense that individuals were expecting to do well in any pleasant way. Rather, it could have been that Feather meant that participants were disappointed at achieving a low level of success and felt that they should be doing far better. In this sense, they were still "optimistic" about future success, but not because they were confident of doing better, but more in the sense that they were disturbed by the levels of success they were achieving.

Also, Feather's (1963a) data shows that when participants achieved a success level of 20%, over 120 trials, they expected to get an average of 35% of predictions correct. In Experiment 10, participants achieved an initial success level of 10%. If participants in Feather's study who achieved an initial success rate of 20% expected to achieve an average of 35% on the next block of trials, would it be reasonable for participants who achieved a success rate of 10% after the first block of trials (in Experiment 10) to be content with an improvement to 2 out of 10, and then from 2 to 3 out of 10 etc? It seems unlikely.

What I am suggesting here is that participants in the improvement group may have expected to do better on future trials, but even though they were improving, they were not improving to the levels they considered to be reflective of adequate performance. It was as if the improvement they were achieving set them up to believe they had turned the corner and were about to do well, only to find that they were still performing poorly. On the other hand, participants in the 40% constant group, although not improving, perhaps achieved a less disappointing discrepancy between their actual success and their expectations. In Feather's (1963a) study, participants in the 50% success group expected to get about 55% of future predictions correct, that is, they expected to perform only slightly better than their current level of performance. Thus for participants in the 40% constant group in Experiment 10, it may have been that although they were expecting to do better on subsequent trials, they were missing their future expectations by less than participants in the improvement group.

Of course it is recognised that the above suggestions are speculative, because Feather did not test the success rates of 10%, 30% and 40%, and also because Feather's study used a different task and a different number of trials. The point to take from the above analysis is that Feather's findings suggest a possible explanation for why participants in the improvement group enjoyed the task less compared to participants in the constant group, namely, that participants in the improvement group missed their expectations to a greater degree compared to participants in the constant group. In other words, the reason that Experiment 10 produced a different outcome from Experiment 8 and 9 could lie in Feather's finding that participants who do poorly at a task are sometimes more likely to overpredict success than those who do well. Participants in the improvement groups who started at 10% (Experiment 10) would have been more likely to be disappointed by gradual improvement than those who started at 20% and 30% in Experiments 8 and 9.

Thus, there seems to be several reasons to suspect that the findings in Experiment 10 were not anomalous. Indeed, if the speculations are valid, the findings may have identified an important feature of improvement, namely, when it is likely to help and when it is likely to hurt. These issues are discussed in more detail in the next section of this chapter.

4.7 Experiments 6-10: Discussion and conclusions

The five experiments in this chapter examined whether or not improving at a task would result in participants rating a task more positively, compared to a group who achieved a constant rate of success. The results suggested that improvement increased task enjoyment when participants were already performing well at a task, but actually undermined experiences when they performed poorly. A summary of the findings from all five experiments is presented below.

Summary of findings

- Experiment 6 provided evidence that the experimental task had internal validity, as participants who achieved a success rate of 70% rated the task significantly more positively than those who achieved a success rate of 30%.
- Experiment 7 showed that when the two groups achieved the same average success rate (50%), the group that improved rated the task more positively compared to the group who achieved a constant success rate. This finding was confounded by the fact that the improvement group also achieved a higher final success rate compared to the constant group.
- Experiment 8 showed that when two groups achieved the same final success rate (70%), the group that improved again rated the task more positively compared to the group who achieved a constant success. These differences were not significant. However, when the ratings for the dependent measures of fun, enjoyment and interest were averaged for each individual and the two groups were compared,

significant differences emerged whereby the group that improved rated the task more positively.

- Experiment 9 replicated experiment 8 but this time with a different improvement schedule. The findings were identical to those of Experiment 8.
- Experiment 10 tested improvement to 40% vs. a constant 40% success rate. Using the same comparison procedures employed in Experiments 8 and 9, the finding was that contrary to the results observed in Experiments 8 and 9, it was now participants in the constant group who rated the task more positively.

4.7.1 The role of the final outcome in determining task experiences

Given that improvement caused participants to rate the task more positively in Experiments 8 and 9 but had the opposite effect in Experiment 10, it seems that how improvement influences task experiences depends on the exact conditions. So how can the pattern of effects, whereby improvement seems to helping under one set of conditions but hurting in another, best be interpreted?

One possibility is that the final outcome that individuals achieve may be an important determinant of whether or not improvement effects will occur. That is, when participants do well, improvement helps, but when they do poorly, it does not. To elaborate, the most robust finding in this chapter is that whenever one group has ended up doing better than other, that group always reported the task significantly more positively. This effect even occurred in Experiment 7 when the difference in final success rate between the two groups was only 20%. Thus, it might be that when individuals do well at a task, the satisfaction this produces overshadows any possible contribution from a feeling of having improved. Some tentative support for this theory can be taken from a study by Redelmeier and Kahneman (1996). In this study, participants were asked to rate their real-time pain experiences whilst they were undergoing a medical procedure. They rated their pain every sixty seconds via a hand-held device which was used to control a marker on a computer screen. The marker was anchored by the terms "no pain" and "extreme pain". This methodology allowed Redelmeier and Kahneman to calculate the total amount of pain that participants experienced throughout their operation. One hour after the operation, participants were asked to rate the total amount of pain they had experienced throughout by indicating their level of discomfort on a 10-point scale. Given that the patients had reported on their own levels of pain throughout the procedure, it would be reasonable to suspect levels of discomfort would have been positively correlated with the total amount of pain reported throughout their operation. However, the results revealed that although the patients experienced different amounts of total pain, their judgements of total pain were correlated with two measures: The peak intensity of pain and the intensity of pain recorded during the last three minutes of the procedure. In other words, patients seemed to base their overall judgements of their experience on two significant moments, namely, the highest amount of pain and the terminal level of pain. This finding would suggest that regardless of the sequence of outcomes that individuals experience, it is the final and the peak rates of experience that most strongly determine their experiences. It might be that simply performing well by the time a task is completed is sufficient for individuals to experience that task positively.

Other lines of research also support the contention that final outcomes are powerful predictors of participants' experiences of ongoing tasks. For example, in the literature examining the effects of the grade that students attain on a course and their subsequent ratings of that course, there is considerable evidence that suggests that higher grades result in higher ratings (see Stumpf and Freedman, 1979 and Feldman, 1976 for reviews; also Chako, 1983; Blunt, 1981; Vasta and Sarmiento, 1979; Worthington and Wong, 1979; Powell, 1977; Holmes, 1972). In these studies, students' end of semester course evaluations were examined to investigate which variables related to positive course ratings. The variable that consistently emerged as a key predictor was the overall grading leniency of the course, that is, those course where students attained, on average, higher grades, were those that were rated most positively.

4.7.2 The key role of surpassing expectations

But do high outcomes alone predict task experiences? There is considerable debate whether or not they do (e.g., Greenwald and Gillmore, 1997a; 1997b). For example, Marsh and Roche (1997) have reviewed a number of studies in this area and have suggested that the best estimate of the proportion of variance accounted for by grades is probably around 4 percent. One recent criticism of the grade evaluation studies is that because the evaluations are anonymous, it has not been possible to assess whether students who do well in a course rate it more postively than students who do less well in that same course (Remedios, Lieberman and Benton, *in press*). Additionally, in all these previous studies, grades have been used as an indirect measure of grade expectations on the assumption that expected grades must be reflective of the current grades that students are currently achieving on the course (e.g., Greenwald and Gillmore, 1997a; 1997b). Where it has been possible to identify students individually, these studies have been lab-based where grades have been artificially manipulated (e.g., Holmes, 1972; Worthington and Wong, 1979; Blunt, 1981). The results in these lab-based studies have been criticised for the inappropriateness use of statistics (e.g., Marsh and Roche, 1997). To overcome these problems, Remedios et al. used individual student identification numbers and monitored students as they progressed through their course to assess the relationship between their ongoing grade expectation, the absolute grades they were attaining and their enjoyment of the course. Using course enjoyment as the dependent variable, a stepwise regression analysis revealed final grades to be a predictor of course enjoyment, but not the best predictor. Instead, it was the difference between students' final grade and their initial grade expectation that best predicted their final enjoyment ratings ($p < .01$). The implication of this finding is that high final outcomes might not be sufficient to cause participants to enjoy tasks, the important component might be whether or not participants surpass their initial grade expectations.

Whilst the foregoing evidence seems relevant, it could be argued that studies such as Remedios et al. are not analogous to the experiments conducted in chapter 4. One difference is that participants in these experiments were not given an explicit normative standard that they could match their own performance against, whereas in the Remedios et al. study participants were able to compare their final grades with other students, because all final grades and names were displayed on a central noticeboard. This meant that they had knowledge of the average class grade that other participants achieved. Thus, a key predictor of task enjoyment in the Remedios et al. study might not have been the degree to which participants surpassed their expectations, but rather the degree to which they performed better than their peers. To see whether or not passing or failing the class average had an effect on enjoyment, Remedios et al. entered the degree to which final grade differed from average class grade into the regression model but it did not emerge as a significant predictor. This suggests that students' knowledge they had done better or worse than their peers was not as important in determining their enjoyment as how well they had done relative to their own expectations.

A similar conclusion has been suggested by research on learning in animals. For example, Crespi (1942) has also shown that it is not the final outcome that determines responses but expectations relative to past experiences. In his study, rats were trained to run down a maze to a goal-box containing 1, 16 or 256 pellets of food. As expected the larger the reinforcement, the faster the rats ran down the alley. After 20 trials, reinforcement was standardised so that all the rats were now rewarded with 16 pellets of food. The speed of running to the goal-box was again monitored over a further 8 trials. The results revealed that the group who were originally accustomed to the single-pellet reward not only began to run towards the goal-box faster, but they ran there faster than the group who had previously been rewarded with 16 pellets. That is, the group who had previously received only one pellet showed what Crespi referred to as an "elation effect". In a similar fashion, the rats who originally received 256 pellets of food but now only received 16 ran to the goal-box slower than the rats who had originally received 16 pellets, what Crespi referred to an as a "depression effect". What Crespi's study demonstrated was that it was not the absolute level of a reward that determines responses, but the level of reward in contrast to prior rewards -- that is, prior rewards set the expectation level for future rewards. When these future rewards are either greater or less than before, then it is the level of contrast between these different levels of rewards that determine responses.

The findings from the Crespi and the Remedios et al study therefore suggest that the relationship between final outcomes and initial expectations have an important effect on participants' responses. Applying these analyses to the findings from the experiments in chapter 4, there seems to be considerable overlap. Recall that the argument was made that participants who improved to 70% may have surpassed their expectations because, as they progressed through the task, their current level of performance became their new expectation for success, a level that they then exceeded on their next trial. For participants in the constant 70% conditions, this initial level of success might have set the standard they hoped to achieve on future trials but because they continued to achieve the same level of success throughout, they did not repeatedly surpass current expectations of success. The relationship between outcomes and expectations also explains the findings in Experiment 10, where participants in the 40% constant group enjoyed the task more than those who improved to 40%. Based on the findings of Feather (1963), it was speculated that those in the 40% improvement group would have failed to surpass success expectations as they progressed through the task because their expectations would have been higher due to their poor performance. That is, performing poorly caused them to expect to do much better on subsequent trials, an expectation that they consistently did not meet.

There are probably many factors that contributed to participants' experiences of the tasks in this chapter but the role of improvement, or, as it has been interpreted in this chapter, the effects of surpassing expectations, appear to be important. If improvement at task really does cause participants to enjoy those tasks, then clearly there are important practical implications. For example, if a teacher wants pupils to enjoy tasks that may be difficult, then it might be helpful to prime students with low expectations of success, and also to organise the task so that the pupils experience gradual improvement. Indeed, and somewhat counter-intuitively, it might also be helpful to allow students to fail to begin with so that they can experience gradual improvement with future engagements at the task. That is, performing to a high standard immediately might not engender the most positive experiences towards a task. Of course, as demonstrated in Experiment 10, allowing participants to fail to too low a level might be detrimental, but there seems at least to be a theoretical case that performing averagely to begin with and improving thereafter may be a profitable policy.

Whilst there is considerable evidence that simply knowing one has performed well is sufficient for participants to experience tasks positively, the experiments in this chapter suggest that the process that they arrive at these high levels of performance is also important. If educators and parents want their charges to enjoy tasks, then it seems that improvement may be a useful training tool. In this sense, improvement is not just an interesting research variable, but one with potentially important practical implications.

5 Conclusions

In Chapter 1, the argument was made that although several theories had focused on the positive effects of engaging with a task that had a mastery-focus orientation, it was unclear how participants would respond, in terms of task persistence and task enjoyment, to performing well or poorly.

In Chapter 3, five experiments examined the question raised in Chapter 1. In these experiments, the method used to measure persistence was to allow participants to persist with a task for as long as they wished. To examine whether or not the use of the interview at the end of the initial persistence phase may have explained why participants persisted as they did, an alternative method of measuring persistence – i.e., free-choice persistence - was employed, thus allowing participants to persist with the task under two different forms of persistence. If participants persisted in the same way under the two different modes of persistence, then this would have suggested that the method of persistence was probably not a contributory factor to the patterns of persistence. However, it was shown in Experiments 4 and 5 that participants persisted in opposite ways depending on how their persistence was measured. Moreover, existing theories did not appear to account for participants' behaviour in the free-choice period. For example, according to Dweck (1986) in achievement settings, individuals are constantly trying to resolve ability-related issues. However, given that participants were allowed to persist for as long as they wanted to during the initial persistence phase, it was unclear what exactly they were trying to resolve in the subsequent free-choice phase. Instead, the argument was made that participants might have persisted during the free-choice period simply because of the pleasure they were getting from either doing well or poorly at the tasks.

In chapter 4, the concept of improvement as an independent variable was examined. In these experiments, two groups were compared, one which improved to a certain level of success and the other which experienced a constant level of success. The results suggested that at high levels of success, there appeared to be a positive effect of improving, but at lower levels of success, the opposite effect occurred, namely, more positive experiences after achieving a constant level of success. It was suggested that these effects occurred because individuals seemed to have set themselves performance targets. It was noted that further speculation was required for where these targets may have come from. I shall comment on the issue of the types of standards individuals might have been setting themselves later in this chapter, but firstly, there are two possible explanations for task persistence in achievement settings that require addressing.

5.1 Alternative explanations for persistence after failure in the initial persistence phase in Experiments 1-5.

In chapter 3, there were several reasons offered for why participants persisted so long after failure during the initial persistence phase. However, there are several other possible explanations that it may be useful to examine.

One possible explanation could be that participants persisted with the task because when they performed poorly, the task remained a challenge to them, but that when they performed well, the task was no longer a challenge. According to Deci (1975), a feature of intrinsic motivation is the seeking out of challenges. Several researchers had earlier identified exploration as a fundamental motive for task persistence. For example, Montgomery (1953) demonstrated that rats spontaneously explore novel stimuli and environments, findings that were supported by Butler and Harlow (1957) who, in addition, found that monkeys would work for the reward of visual stimulation, even when alternative food rewards were available. These exploration hypotheses have been framed within an optimal stimulation account of behaviour. For example, according to Hunt (cited in Deci, 1975), organisms require a certain level of stimulation, so when they are understimulated, they seek out stimulation to the point when that stimulus becomes boring. His argument for why individuals desist with a task is that as stimuli become uninteresting, so organisms will move away from those tasks (see also Leavitt, 1962).

Thus, one possible reason why participants persisted longer after failure during the initial persistence phases could have been because the task was still challenging to them. This argument seems particularly relevant in Experiments 4 and 5 because in those experiments, unlike Experiments 1-3, participants in the two groups were performing two different tasks, a hard anagram task and an easy one. It may be that the difference in persistence in Experiments 4 and 5 was simply due to the tasks affording different levels of challenge.

There are two possible counter-arguments to the suggestion that the challenge of the tasks was a significant factor in determining persistence. Firstly, in Experiments 4 and 5, if participants who attempted the easy anagrams stopped persisting during the initial persistence phase because they had reached some sort of optimal level of challenge, then it is difficult to explain why they returned to the anagrams during the free-choice period. Secondly, if participants in the Failure/Hard groups in Experiments 1-5 did indeed find the task challenging, then if, as Deci (1975) suggests, challenging tasks are intrinsically motivating, then why did participants rate their experiences of the tasks so negatively? It therefore seems unlikely that the challenge of the task had any significant bearing on participants' persistence behaviour.

Another possible explanation reason for why participants persisted so long with the task after failure concerns the Zeigarnik effect (Zeigarnik, 1927). The Zeigarnik effect is described as the tendency for participants to return to tasks which are uncompleted (see Deci, 1975, pg. 38-39). That is, whenever a participant is interrupted during a task, they will choose to return to that task rather than moving on to a different task. The problem with this explanation in terms of the behaviours of participants in chapter 3 is that none of the participants fully completed any task. In the stockmarket experiments, the best performance any participant achieved was a 90% success rate. For the anagram tasks, no participant solved every anagram in the anagrams box. So, during the initial persistence period, because neither group fully completed the problem, we would expect both groups to persist for the same amount of time.

However, interpreting the persistence patterns in terms of the Zeigarnik effect does allow another explanation that may account for why participants in the Failure/Hard groups persisted for so long at the task. Maybe participants had a personal standard that they hoped to achieve and that once they either achieved this, they stopped persisting. Seen in this way, the propensity for participants to persist longer when they were failing makes more sense. If participants were trying to reach a certain standard, then they would persist with the task longer if they were doing poorly. For example, if participants personal standard was, say, 70%, then it makes sense why those who experienced a constant 70% success rate would desist with the task once they achieved their target and why participants in the 30% success group would continue with the task. However, whilst this account seems plausible, it become difficult to explain why participants who achieved certain levels of success still persisted with the tasks. For example, in Experiment 1, participants in the constant 70% group achieved a success rate of 70% after 30 trials, yet the average number of trials for participants in this group was 93. Similarly, in Experiment 3, participants in the 90% success group achieved this target by trial 60 yet the average number of trials for participants in this group was 105. Of course, it could be argued that participants' targets were higher than 90% but if this were so, then in Experiments 1-3, all participants should have failed to meet their target and thus there should have been no differences in the amount of time they persisted with the tasks.

So, the conclusion regarding the Zeigarnik effect is that it does not easily account for the patterns observed during the initial persistence phase in the experiments from chapter 3. However, it does bring to light an issue which may be useful to examine further, namely, the types of performance standards that individuals set themselves when tasks are novel and no comparison standard is available. It is to this analysis that we now turn.

5.2 How do participants set standards?

In all ten experiments in this thesis, participants have had to determine for themselves whether or not they have been performing well. This begs the questions as to how exactly did they set standards for themselves, that is, what targets were they aiming for? In no experiment did the experimenter verbally guide them; nor were they given any normative target to which they could compare their performance. However, despite this lack of external feedback, participants seemed to be clear when they were doing well and when they were doing poorly. But how exactly did they know they were doing well or poorly? For example, suppose you were given a task you had not encountered before, how would you decide what constituted good performance at that task? It is perhaps worth investigating some of the literature in this field to see whether or not it may help throw light on explaining either or both the patterns of persistence observed in Experiments 1-5 and participants' experiences of the task throughout this thesis.

Festinger (1954) suggests that in the absence of external objective competence cues, individuals will look to others' performance (or opinions) to judge how well they themselves are doing. Festinger (1954) gives the example of poetry, where individuals can only assess how good their poetry is by looking to the opinions of others. According to Festinger, tasks with ambiguous standards require some sort of external validation before individuals can be sure of how they are doing.

There seems to be reasonable evidence to suggest that the stockmarket, anagram and improvement tasks were ambiguous. For example, a common way for experimenters to keep standards ambiguous has been to present tasks that are novel. The work of achievement motivation and intrinsic motivation theorists is replete with such tasks. Two examples already alluded to are Deci's soma cube task (1972) and the more-often used NINA task (e.g. Ryan, Mims and Koestner, 1983; Plant and Ryan, 1985; Koestner, Zuckerman and Olsson, 1990; Ryan, Koestner and Deci, 1991). The question is, how do participants, in the absence of normative feedback, "know" how well they are doing? For the NINA task, participants are expected to find as many instances of the word NINA in a cartoon drawing but they are not told how many actual instances there are. Therefore, finding eight instances might be very good or very bad. For the stockmarket and sequence prediction tasks used in this thesis to date, on what basis did participants believe that a 70% success rate represented good performance?

Kruglanski's (1989) theory of lay epistemics suggests that when individuals are placed in situation where they have no obvious comparison for their ability, participants will construct a concept of what it is to be good or bad at that task, and that these conceptions will be based on their existing knowledge structures (see also Kruglanski and Mayseless, 1990). Kruglanski suggests that in individuals' everyday interaction with others, they store information in their long-term memory which they then use to generate and validate hypotheses when situations are novel or ambiguous. Kruglanski elaborates on this suggestion by citing the work of schema theorists.

Schema theorists suggest that what is recalled from long-term memory will depend on how information was encoded to begin with, and how relevant the information is in helping to resolve a problem. They suggest that individuals develop typical scenarios for what they expect situations to be like, based on their cultural background. For example, Bartlett (1932) found that when individuals tried to remember features of a story, cultural differences between Western Americans and North-American Indians resulted in the groups recalling incidents in the story with a bias towards their own cultural expectations. So although all the participants received the same information (i.e., the story), recall of the story depended on what was salient to the different populations.

In a similar vein, Schank and Abelson's (1977) script theory suggests that when people try to comprehend texts or prose, they do so with a preconception of the events and people in the script. For example, in a restaurant script, there are typical events (e.g., ordering a meal, paying the bill) and people (e.g., the waiter, other customers). People in the script have "role slots" (see Eysenck and Keyne, 1993, p. 279) which determine the type of roles they are likely to play (e.g., the waiter is likely to take the order, serve the food and present the bill).

In terms of the experiments in this thesis, schema and script theory provide a possible explanation for the types of standards participants expect to surpass when faced with ambiguous normative standards. For example, when one performs a task, schema theory proposes that individuals look to previous interactions with that (or similar) tasks (see Fiske and Taylor, 1984, for a review). Following this logic through, it may be that when faced with a task where the measure of good performance is ambiguous, the script that individuals work with may read something like "Do task; compare performance with others; if performance is good, continue, if poor, consider stopping or redoing task". If there is no other objective criterion for evaluating performance, individuals may globalise, or approximate the task to a more general ability that the task represents (e.g. is it a spatial, verbal reasoning, a mathematical task), and compare performance on those dimensions (e.g. "I am not very good at solving mathematical problems, therefore I do not expect to do well at this task; I am particularly good at spatial tasks therefore I expect to do well at this task"). Where tasks are known but rarely personally encountered (e.g. anagrams), individuals may match their performance to how they perceive others would do at the task (see Kruglanski, 1989, p.47).

5.3 What types of standards did participants set themselves?

If the foregoing analysis is a useful way to understand how individuals construct standards, then it may be useful to speculate on the types of ability comparisons individuals might have made for the tasks in this thesis. For example, consider the anagram task. The speculation offered here is that, in line with Festinger's (1954) hypothesis II, individuals may have matched their performance with a referential group[1] (e.g., friends who do anagrams, or contestants on quiz shows). However, it is suggested here that mapping onto categories may mean that the individuals or concepts that become associated with the task may be atypical of the skill generally. For example, when individuals think of anagram solvers, the category is likely to be a small one, and it will probably consist of positive examples, that is people who are good at anagrams. Therefore, in comparing one's performance to others', it is suggested that when faced

---

[1] This is actually an extension of Festinger's (1954, hypothesis II) which suggests that individuals match their opinions and abilities with a referential group. I am adding the speculation with regards to the level of ability.

with tasks such as the anagram one, in the absence of normative information, individuals may set themselves too high a target to achieve.

For the stockmarket and improvement experiments, it is more difficult to speculate what types of targets participants might have set themselves as these tasks were probably entirely novel to them. One possibility is that they expected to achieve 100% success, that is, they expected to get all their predictions right in both the stockmarket and improvement tasks. This would seem to be a rather optimistic target to set oneself for a task that essentially had no specific standard. However, there is some evidence that participants did set themselves high targets. For example, when we consider how participants viewed their performance relative to others —i.e., the WellY vs. WellO analyses —in these analyses, participants who performed well did not rate WellY any higher than WellO. Now recall that in Experiments 2 and 3, participants in the success groups achieved success rates of 90%, and in Experiments 4 and 5, participants' who attempted the easy anagrams achieved success rates that were in excess of 90%. Thus, there was a marked reluctance for participants to rate their performance as better than others, even when they were told that they performed to what might be considered a very high standard. It seems that although they performed well, they did not consider this to be any better than others would have done at the task.

So why did individuals fail to see themselves as particularly competent when they were achieving what seemed to be extremely high levels of success? One possibility is that for some reason, participants had a pre-existing low opinion of themselves, that is they might have had low self-esteem. As personality traits have not been examined specifically in this thesis, we can only speculate on possible levels of such personality measures. It has been shown that when individuals are low on self-esteem (e.g. Baumeister and Tice, 1985), need for achievement (e.g. Feather, 1961), internal locus of control (Rotter, 1966) or need for cognition (Thompson, Chaiken and Aizlewood, 1996), they do not respond positively to success. That is, after success, participants tend to attribute their success to external factors such as luck rather than internal factors such as ability. Contrariwise, individuals high on these traits respond positively to both success and failure, that is, success causes them to rate tasks positively whilst failure causes them to become motivated to resolve the tasks. The example given earlier in this thesis about the positive verbalisations given by children who were mastery-focused but continued to fail at a task is perhaps a good example of this phenomenon (e.g. Diener and Dweck, 1978, 1980). Given that the general finding in this thesis was that participants tended to report negative experiences of tasks when they performed poorly, this would in turn suggest that may have been low on one, or a combination of all the above mentioned traits.

This in turn would suggest that participants in this thesis who failed must have had, on average, low opinions regarding their ability to perform the tasks they were given. The question is whether it is plausible to believe this? The participant population used in this thesis were either undergraduates at Stirling University or Open University students. Would it really be reasonable to claim that on average, this population would have low opinions of themselves, or at least, lower opinions than the general public? It is difficult to see that this would be true.

A more plausible interpretation may be that individuals preferred to be conservative about their performance. The findings from this thesis certainly suggest that individuals were reluctant to report to others that they thought that they had done well, and this might have been because they had a fear of being wrong. This may have been because participants knew their answers were going to be analysed by the experimenter and that it was maybe better to report that they were average and be told they were better than average, then to say they had performed well and then run the risk of being told they actually performed averagely.

Thus, in speculating what types of standards individuals set themselves, the suspicion is that participants may have set themselves too high standards resulting in them never really experiencing high levels of competence. This conclusion would both account for the relatively low ratings of task experiences in both the various success and failure conditions. Thus, to summarise, the conclusion is that in the absence of specific normative targets, participants must have constructed some sort of target to attain, and that in terms of this target, they only ever perceived themselves to have performed averagely.

This position is also supported when we consider the interpretations of the findings for the improvement experiments in chapter 4. Here it was suggested that the different effects of improvement could potentially be explained by participants not meeting their expectations of success. Within that argument, it was suggested that the failure to find the improvement effects in Experiment 10, when participants improved to 40% was because although participants were improving, they expected to perform even better and so were missing their expectations by large amounts.

We can also tie this analysis to another feature of Festinger's (1954) theory of social comparisons. For example, *corollary IIA* states that when standards are ambiguous, individuals judgements about their ability will be "unstable". That is, without an objective norm or a subjective standard to refer to, individuals' perception of their performance will depend solely on their rate of success. However, Festinger does not state what levels of success will lead to high perceptions about one's ability. In the improvement experiments in chapter 4, when individuals improved to success rates of 70%, they perceived themselves to be relatively competent, but not when they improved to success rates of 40%. If perceptions of ability are solely determined by prior success, then we would expect the improvement group in Experiment 10 to have had more positive experiences of the task relative to the constant 40% group. This was not the case. So, where standards are ambiguous, the evidence from chapter 4 suggests that Festinger's suggestion is only supported when participants perform at a relatively high level of success. When individuals perform at a low level (i.e., 40%), they seem to have a larger pre-determined target in mind because even when they improved, they seemed to be disappointed at not achieving a certain level of success.

The implication of the above analysis is that when individuals are placed in task situations, then experimenters need to be aware that the experimentally-manipulated objective values of success and failure need to be considered in terms of whether or not these values will be experienced by participants as success and failure. This is because it seems that whenever individuals are placed in achievement settings, they construct some target of achievement for themselves. It seems therefore that simply manipulating task outcomes may not be sufficient to create the appropriate feelings of success because it assumes too much about the targets that individuals have set for themselves.

Indeed, it seems that if researchers want to be certain that participants believe they have done well (or poorly), the most effective option might just be to tell them.

5.4 The role of perceived ability in task persistence

Perceived ability or perceived competence has been cited as an important determinant of behaviour in task situations by several researchers. For example, perceived ability plays an important role in Dweck's (1986) achievement-goal theory. She suggests that perceived ability determines how individuals will engage with tasks depending on whether they are mastery-focused or performance-focused (see also Dweck and Elliott, 1988; Dweck and Leggett, 1988; Dweck, 1996). She suggests that when individuals have high levels of perceived ability and are mastery-focused, this high confidence leads to them persisting with tasks, even when they perform poorly. One the other hand, when participants are performance-focused, Dweck suggests that high perceived-ability is more prone to instability, so that after failure, participants are more likely to disengage from tasks. Thus, for Dweck, perceived ability influences how individuals will react to failure when they are either mastery or performance-focused. This position is similar to that of Bandura (1990), who suggests that self-efficacy is the key determinant of behaviour in task situations. His position is that as individuals progress through tasks, their goals change because of their levels of self-efficacy. Nicholls (1984) too offers a very similar account of the role of perceived competence in achievement situations. His analysis not only places an emphasis on individuals' levels of perceived ability, but also their perception of task difficulty. According to Nicholls, in achievement settings, individuals are constantly trying to prove to others that they have high ability. According to his analysis, if a task had a performance-focus, he would expect participants to desist after failure, but only when tasks were perceived to be easy. This is because failure at a task that is easy, even when individuals feel they have a high probability of solving the task, would never indicate to others that the participants had high ability. Thus, participants' initial perception of the level of difficulty of the task, their perceived ability to solve that task, and their orientation to the task, jointly determine how long they will persist.

Given the prominence of perceived ability as an explanation for behaviour in achievement settings, it may be useful to examine how this variable may have influenced behaviour in the experiments in this thesis.

In Experiments 1-5, the overwhelming finding was that during the initial persistence phase, participants persisted for longer after failure compared to success. What was particularly surprising was the amount of time that participants seemed to persist after failure. Let us concentrate on the persistence after failure during the initial persistence phase because it is the responses to failure that figure prominently in Dweck's analyses (e.g., Diener and Dweck, 1978; 1980; Dweck, 1986; Dweck and Elliott, 1988; Dweck and Leggett, 1988; Dweck, 1996). Let us also assume that the task was performance-focused, that is, let us accept the suggestion made earlier in this thesis that the fact that participants knew they were going to be interviewed after the task caused them to become concerned about their performance. If the task was performance-focused, then how might the theories explain why participants persisted so long after failure?

Dweck's (1986) analysis does not seem to be able to account for the greater persistence after failure. For example, let us assume participants had high perceived ability. If participants were performance-focused, then according to Dweck, failure should have caused participants to lose confidence and desist with the task. If they had low perceived ability, then they should have desisted with that task very early on. The only situation that predicts high persistence after failure is if participants were mastery-focused. However, if they were mastery-focused, then it is surprising why their subsequent experiences of the task were so negative - in all cases, participants who performed poorly rated WellO higher than WellY. If participants had low perceived ability, then if the task was performance-focused, they would desist very early with the task. If the task was mastery-focused, Dweck suggests that individuals would still take on the challenge of trying to improve at the task. However, in terms of task experiences, Dweck would predict that participants would experience the task positively, or at least, not negatively. Again, this was not borne out by the findings in Experiments 1-5.

Assessing the position in terms of Nicholls' (1984) theory requires an assumption to be made about how difficult participants perceived the task to be. Let's look at the cases when participants felt the task was very difficult, moderately difficult and very easy. Let's first consider the position where individuals had high perceived ability. If the task was perceived as very difficult and participants had high perceived ability, then according to Nicholls, participants would continue to persist with the task because they would have high expectations of future success. However, they would have to have a level of perceived ability that was greater than the difficulty of the task, and if they were failing, they would probably desist from the task. This is because failure would cause them to lower their perceptions of their ability and thus they would realise that they could no longer exhibit high ability. The same goes if the task was very easy. In this case, participants, even if they had high perceived ability would have no opportunity to exhibit high ability because they would know that others with less ability than themselves would also do well at the task. However, if the task was perceived as moderately difficult, then according to Nicholls, participants would persist for that task for considerable amounts of time because they would always have the confidence of resolving it. Thus, in terms of Nicholls' analysis, in the case where task difficulty was perceived to be moderate (i.e., within capabilities), high persistence would be predicted, but in the other two cases, where task difficulty was very high or very low, low levels of persistence would be expected.

However, there is also another scenario within Nicholls analysis that might predict high persistence after failure. According to Nicholls, in task which has a performance-focus, if participants thought the task was extremely difficult, then if they had low ability-perceptions, persisting at that task could not indicate to others that they had low ability. Although they would clearly not be resolving the task, they would perceive that even individuals with high ability would perform as poorly as them, that is, even individuals who were highly skilled at the task would also fail at the task. Thus, persisting at the task would not lead to others considering their performance to be poor.

Given that Nicholls' analysis potentially predicts persistence after failure under two conditions when tasks are performance-focused, it may be useful to speculate on whether the task was perceived as difficult or easy for those participants who failed. It seems that not much speculation is required because when asked how difficult they thought the task was, apart from Experiment 2, participants always rated the task more difficult after they had performed poorly at it. So, relating this back to Nicholls' analysis, the greater persistence could have been because participants had low levels of perceived ability and they thought the task was very difficult. On the other hand, when interpreting the data in terms of Dweck's analysis, the persistence could only be explained if participants were mastery-focused. Given that we have already established that it is unlikely that participants had low levels of perceived-ability, and that during the initial persistence phase, there is good reason to suspect that participants were performance-focused, then neither Dweck or Nicholls' analysis seems plausible. So when we examine the behaviour of participants in Experiments 1-5, it seems that in trying to work from the data backwards, this does not enable us to speculate clearly on how perceived ability may have influenced behaviour, at least in terms of Dweck's and Nicholls' theories. It seems that even if we had asked participants a priori how well they thought they could do at the task and how difficult they perceived the task to be, it would have still been difficult to predict their patterns of persistence.

However, it would be inappropriate to suggest that perceived ability played no role in determining task persistence. The problem lies in determining *how* perceived ability had a role. For this analysis, it may be useful to turn to Bandura's (1990) account of behaviour in task situations.

Bandura (1990) believes that it is incorrect to see goals as the driving force behind task persistence. Instead, he claims that individuals can hold different goals at different times throughout the same task depending on how efficacious they feel. He suggests that individuals readjust their personal goals in light of their current attainments (Bandura, 1990). This position has received support from other researchers who suggest that as individuals progress through tasks, so their perceived ability changes (e.g., Feather, 1963a, 1963b, 1967, 1969; Norem and Cantor, 1990; Harter, 1985). Nicholls too has shifted his emphasis away from just seeing individuals as striving to attain ability judgements and towards a perspective that also recognises individuals' ongoing experiences during a task (see Thorkildsen and Nicholls, 1998 for review).

From the patterns of persistence observed in Experiments 1-5, it seems difficult to be able to consider the role of perceived ability in any terms other than as Bandura suggests. Of course, this is in part because in Experiments 1-5, perceived ability was never measured, either before or during the task. However, in examining the role of perceived competence as a thought experiment, it is difficult to understand how either Dweck's or Nicholls' theory would predict the patterns of persistence that were observed.

It is therefore my contention that perceived ability probably did play a crucial role in determining why participants behaved as they did, both in terms of task persistence and task experiences. I suspect that as individuals progressed through the tasks, their conceptions of their ability changed as they became more or less confident in the likelihood of solving the problems and so their persistence changed. For participants who performed poorly, as they continued to perform below expectations, so they began to strive harder to solve what they thought was a solvable task. However, because they had set their targets high, they continued to fail to met their targets, causing them to continue to persist with the task. When they finally left the task, they were disappointed at failing to meet their targets, causing them to rate the task negatively.

5.5 Ecological validity

The experiments in this thesis have examined how individuals respond to performing well and performing badly in specific achievement settings. However, as already mentioned, it might be questionable whether the behaviour observed in the experiments actually reflect how these same participants might behave in non-experimental situations. For example, is it really reasonable to conclude on the basis of the findings in experiments 1-5 that participants, in their everyday lives, would persist with tasks for so long after performing poorly? Is it reasonable to conclude that improvement really was the independent variable in experiments 6-10? These two issues are now examined.

5.5.1 Initial task persistence after failure

Recall that the claim was that participants might have persisted for a long time during the initial persistence phase because they felt that their performance was going to be evaluated in some way by the experimenter, after they had finished engaging with the task. Looked at in this way, it may not have been just the poor performance that caused participants to persist with the task, but simply the fact that they felt they were being watched. Lepper and Greene (1973) have shown that when children knew they were being watched, this affected their persistence at a task. Indeed, intrinsic motivation theorists believe that when individuals perceive they are going to be evaluated, this causes them to become externally-regulated, that is, they feel that they are behaving because they are forced to rather than because they want to (see Deci, 1975; DeCharms, 1968). For example, according to DeCharms (1968), when individuals feel they are being watched, this causes them to take an external perspective on themselves and view their own behaviour in terms of how others would see them. Lepper and Greene's (1975) overjustification hypothesis suggests that when individuals are given a reason to believe that the cause of their behaviour is external – e.g., I am doing this behaviour because someone is watching me, ordinarily, I would not do this behaviour on my own volition – then they overjustify this cognition as a reason for their behaviour. This cognition then affects their behaviour in subsequent interactions with the task. According to researchers in the field of intrinsic motivation, when individuals feel externally regulated, this undermines their (intrinsic) motivation (see Lepper and Greene, 1973), that is, concern about being evaluated by others causes individuals to <u>desist</u>, not to persist with tasks. This is the opposite of what was observed in the persistence experiments. If surveillance really did have an undermining effect, it may have served to cause participants to become aware of the task as a performance one, but it did not seem to undermine persistence. On the contrary, it may have served to increase it.

This leads us on to speculate about the type of persistence that was observed and whether this persistence is a phenomenon observed outside of laboratory settings. In this regard, it seems to me that Ryan's (1982) suggestion that individuals might have been persisting because they were trying to prove competence to themselves seems a reasonable account of participants' behaviour in the initial persistence phase, especially for those who performed poorly. For example, it was particularly marked how, during the post-experimental briefing sessions, participants commented on their performance and seemed to be, at times, distraught at their poor performance. Also, recall that the task instructions in Experiment 4 had to be slightly amended because the first three participants tested all persisted with the task for over an hour and had to be collected from their testing rooms as the rooms were required for another class. This event should not be taken lightly. Although these participants had signed-up for the experiment as part of a course requirement, the typical length of time for other experiments on the sign-up board was between 10-15 minutes. If participants had a temporal template for how long they should have committed to the task, it is unlikely this would have been much different to this average. Indeed, two of the participants had classes to go to in the following hour and both of these participants were wearing watches. Unfortunately, because of the amount of time they spent at the task, their de-brief was short so I did not get a chance to ask them why they persisted for so long. Several authors, mostly in the intrinsic motivation tradition, have suggested that individuals can become so immersed in a task that they lose track of time (e.g. Csikszentmihalyi, Larson and Prescott, 1977; Csikszentmihalyi and Larson, 1987; Csikszentmihalyi, 1988, pgs. 33-34; DeCharms, 1968, pgs. 269-270; Deci, 1975, p.23). However, in all of these examples, the argument is that individuals become immersed in task because they are enjoying them. Participants who performed poorly in experiments 1-5 certainly did not enjoy their experiences of the tasks. In this sense, they seemed to be exhibiting the type of ego-involved persistence espoused by Ryan (1982).

The question is whether or not the initial persistence participants exhibited in Experiments 1-5 reflected how they would have behaved outside of the experimental situation. I suspect it is. It was markedly noticeable that whenever I asked volunteers to help pilot the tasks for Experiments 1-5, they seemed to be concerned about their performance. This was true even for friends and colleagues in the department, even though I told them that in case of the stockmarket task, I only wanted to test the program. Indeed, some individuals refused to help me test the anagrams claiming that their data would be useless anyway as they were very poor at solving anagrams. It should be noted that those who helped with the pilot studies were largely fellow postgraduates (in the first instance – subsequent pilot studies were completed with the undergraduate population), hardly a population who would be thought to shirk tests of ability. This concern about performance seems therefore to be pervasive and, as mentioned before, it is strongly suspected that the method of measuring persistence highlighted these real concerns. Thus, in terms of Ryan's (1982) analysis, it may be that this analysis is not solely applicable to free-choice persistence paradigm situations, maybe the phenomenon he identified is one that is applicable to individuals everyday interactions with tasks. It seems that individuals are concerned about their performance and that when they feel they are going to be evaluated by others, however ambiguous the form of that subsequent evaluation, they will strive to perform well causing them to persist in the face of failure, even when they are not enjoying themselves. In short, it seems that the persistence observed in Experiments 1-5 was not just an artefact of the experimental situation, it seems that individuals may indeed suffer ego-involved types of persistence in all types of tasks they engage with, especially when those tasks are seen as ones that question their ability.

5.5.2 Did participants really experience improvement?

It is difficult to compare the experimental manipulations adopted in Experiments 6-10 with similar research that has specifically improvement as an independent variable for the simple reason that there appears not to be any. However, this does not prevent us from speculating on the validity of the manipulations used in the experiments in chapter 4.

Knowing whether participants believed they had improved is difficult because unfortunately, they were not asked about this in their post-task questionnaire. However, it was encouraging to note that the experimenter who ran the studies did report that participants felt they were getting better, even though they were not sure why. Typical comments by participants in the improvement conditions were "I thought I did OK and seemed to get better towards the end" and "I was pleased that my marks were getting better". Some participants commented on the implicit learning cover-story and incorporated this into their perceptions of their performance. For example, several participants reported not being aware of how they were getting better but said that they were pleased that they were. Given these accounts, it seems reasonable to suspect that participants did indeed recognise that they had improved.

There are several additional methodological criticisms that can be levelled at the experiments in chapter 4 in terms of whether the task was an appropriate one from which to generalise experiences outside of the experimental setting. For example, how representative was the task in terms of the types of task that individuals experience in real-life? This can be debated from two levels, in terms of participants' perceptions of the type of task they were completing, and secondly, in terms of the amount of feedback they received.

To elaborate on the first criticism, if the purpose of the experiments in chapter 4 was to present participants with a skill task, then it seems necessary to establish whether the task was perceived as one requiring skill. For example, it seems questionable whether participants saw the task as one that essentially involved skill or whether they thought their performance was just down to luck. If participants thought their performance was due to luck, then it is unlikely they would have been bothered when they did well or poorly. If they did poorly, then they could just say that the task was simply one where ability had little to do with their performance. This speculation is plausible. After all, participants were not asked whether they thought the task was a skill one. In addition, several participants' data had to be excluded from the final analysis because they reported that they thought that it was the computer that was controlling their responses. It may have been that of the data included for analysis, many of these participants could also have been suspicious about the outcomes. In this sense, there seem to be grounds for suspecting that participants may not have been responding to the task in terms of it inferring anything about their ability.

However, other aspects of the results strongly suggest that participants really were concerned about how they performed. In all cases where one group performed better than the other, the group who performed better rated the task more positively. If participants were really not concerned about their ability, then their experiences of the task should have been independent of how well they actually performed. Similarly, the differences in WellY and WellO ratings suggest that participants were concerned about how they performed relative to others. If they thought that the task was essentially one of chance, then when they did poorly, they should have thought that others would have done just as well as them. After all, if performance was random, then others should also perform randomly, that is, performance should have been independent of the individual. Thus, whilst there is clearly good reason to suspect that participants could have perceived the feedback in a way that did not influence their feelings about their ability, their actual ratings suggest the opposite. It seems that despite the potential for participants to excuse their performance, they did not do so.

The second criticism concerns the amount of feedback that participants received. Recall that participants completed a prediction task six times and received six pieces of feedback telling them how many predictions they had got right on the previous ten trials – can this really be equated to how individuals experience improvement outside of the experimental setting? Do individuals really experience improvement so minimally outside of the experimental setting?

I suggest that there are many examples that they do. For example, in the Psychology Department at Stirling University, first year students are given a series of multiple-choice questions four times each semester. In each test, participants are given 10 questions. This seems remarkably similar to the paradigm adopted in chapter 4, except that the feedback is received over a longer period of time. As a tutor for many of these students, I have become increasingly aware of their comments regarding how they have either done better or worse than their previous marks. Moreover, the study by Remedios et al (*in press*) on the effects of grades on course enjoyment supports the suggestion that improving, or surpassing prior expectations as it has been interpreted in chapter 4, plays an important role in students' experiences of the tasks they engage in. This study was set in a real-life classroom and variables were not manipulated as they were in chapter 4. Instead, participant's grade expectations were monitored and then matched against the actual grades they attained during the semester. Regression analyses revealed that the best predictor of course enjoyment was the degree to which participants surpassed their initial grade expectations. So when participants actually did better than they expected to outside of the experimental setting, then as in the experiments in chapter 4, they rated their experiences more positively.

5.5.3 Real effects?

It is clear that the experiments in this thesis contain methodological flaws, but these flaws do not appear to have had any significant bearing on the phenomena observed both in chapters 3 and 4. I suspect that the flaws in methodology probably contributed to the non-significant results in Experiments 1-4 (and experiments 8-10 in the first instance), but these flaws should not detract from the potentially important phenomena that were identified. For example, in Experiment 5, when the methodological issues identified from the four previous experiments were resolved, significant difference between the two groups emerged between the two groups across all the main dependent measures. Also, the fact that the effects of improvement were replicated in Experiments 8 and 9 suggests that the findings were robust. Thus whilst it is entirely reasonable that caution should be exercised over presuming too much from a limited set of empirical data, it seems just as important not to underplay the existence of real phenomena. In examining the possibility of a Type II error, one should be careful not commit a Type I error instead.

5.6 Final Conclusions

The thesis initially aimed to investigate how participants would respond, in terms of persistence and enjoyment, to differing levels of success when tasks were presented with a mastery-focus. The experiments in chapter 3 provided consistent evidence that when participants were allowed to persist with a task for as long as they wanted, they persisted longer when they did poorly compared to when they did well, regardless of whether the task was presented with a mastery or a performance-focus, even though they were not enjoying themselves. These persistence findings for the initial phase suggest that whenever participants are presented with tasks and given the opportunity to persist with them for as long as they want, they cannot help but be concerned about how their performance might be interpreted by others. It seems that this type of persistence paradigm cannot but induce concerns about performance. Indeed, it may be that such experimental paradigms are unable to examine mastery-type behaviour because of the unwitting pressure they put on participants to do well.

However, the most interesting finding in chapter 3 was observed in both Experiments 4 and 5 when participants were allowed to persist with the task for as long as they wanted, and then again in a subsequent free-choice persistence period. Here it was observed that those who had done poorly, persisted initially with the task but in the subsequent free-choice period, did not. The opposite patterns of persistence were observed for those who did well.

Achievement-goal theorists such as Dweck and Nicholls claim that when individuals are in achievement settings, they are constantly trying to resolve ability-related issues.

However, these theories do not easily account for the behaviour of participants during the free-choice period in Experiments 4 and 5, given that they already had the chance to resolve any ability issues during the initial persistence phase. It seems therefore that whilst achievement-goal models have specified when individuals will be mastery or performance-orientated, and take into account the role of factors such as perceived ability, they are still vague concerning what constitutes good performance or mastery. This vagueness leaves it unclear as to exactly why participants persisted with the tasks during the free-choice period. I have suggested the possibility that they persisted simply because of the pleasure they were deriving from performing well, and that the role of pleasure may be an underplayed determinant of behaviour in task situations.

In terms of the improvement experiments, there seems to be good evidence that improvement positively influences participants' experiences of tasks under one set of conditions but not under another. That is, improvement helps when participants do well at a task, but not when they do poorly. The underlying structure of improvement has been described using an expectations-of-success explanation, that is, improvement effects occur because individuals continually surpass prior expectations of success. This suggests that improvement may be the consequence of a much simpler process, namely, surpassing prior expectations.

Built into the analyses for the findings in chapters 3 and 4 is the suggestion that in the absence of normative information, individuals set themselves targets. In all ten experiments in this thesis, participants have been not been given any normative standard to compare themselves against. Instead they have had to gauge how well they have been doing based on their own perceptions of what constituted good or bad performance. I have suggested that these targets might be set artificially high because participants use inappropriate reference groups, or maybe their targets are actually considerably lower, but that they are reluctant to brag about their achievements. These explanations may be domain specific, that is, they may be restricted to the tasks and task situations described in this thesis. For example, there may be tasks that participants have a reasonably clear idea of what the external norm is (e.g., most students know that the average I.Q. is 100) and there maybe certain situations when it may serve individuals to overplay their performance (e.g., when trying to impress another person when they have no knowledge whether how you performed was good or bad). What remains apparent is that when asked to perform a task that requires some level of skill, even when there is minimal evidence that they are going to be evaluated, individuals, or at least the participants who volunteered for the experiments in this thesis, seem to be determined to prove competence.

How individuals set the boundaries for what constitutes competence is a more difficult question to answer, but clearly one that needs to be addressed before inferring success or failure. The term success is ubiquitous, but it seems that we need to be careful not to confuse high levels of success with *feelings* of success because clearly it is the latter that determines behaviour. Clearly there are many factors that contribute to feelings of success; the findings from the experiments in this thesis suggest that two components, namely pleasure and improvement, may also play a significant role.

References

Alexander, C. N. and Rudd, J. R. (1981). Situated identities and response variables. In J. T. Tedeschi (Ed). *Impression Management theory and Social Psychological Research*. New York; Academic Press.

Anderson, S. and Rodin, J. (1989). Is Bad news always bad?: Cue and feedback effects on intrinsic motivation. *Journal of Applied Social Psychology*, 19, 6, 449-467.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behaviour. *Psychological Review*, 64, 359-372.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Atkinson, J. W. and Litwin, G. (1960). Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *Journal of Abnormal and Social Psychology*, 60, 52-63.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood cliffs, NJ: Prentice-Hall.

Bandura, A. (1990). Self –regulation of motivation through anticipatory and self-reactive mechanisms. *Nebraska symposium on motivation, 69-164.*

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bargh, J. A. and Chartrand, T.L. (1999). The unbearable automaticity of being. *American Psychologist,* 54, 7, 462-479.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press. Cited in Eysenck, M. W. and Keane, M. T. (1993). *Cognitive Psychology: A student's handbook.*Hillsdale: Lawrence Erlbaum Associates.

Baumeister, R. F. and Tice, D. M. (1985). Self-esteem and responses to success and failure: Subsequent performance and intrinsic motivation. *Journal of Personality,* 53, 3, 450-467.

Blanck, P. D., Reis, H. T. and Jackson, L (1984). Effects of verbal reinforcement on intrinsic motivation for sex-linked tasks. *Sex roles,* 10, 369-387.

Blunt, A. (1991). The effects of anonymity and manipulated grades on student ratings of instructors. *Community College Review,* 18, 48-54.

Bogganio, A. K. and Ruble, D. N. (1979). Competence and the overjustification effect: A developmental study. *Journal of Personality and Social Psychology,* 37, 1462-1468.

Brown, J. D. and Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology,* 68, 4, 712-722.

Brown, J. D. and Gallagher, F. M. (1992). Coming to terms with failure: Private self-enhancement and public self-effacement. *Journal of Experimental Social Psychology,* 28, 3-22.

Brunstein, J. C. and Golliwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology,* 70, 2, 395-407.

Burt, D. M., and Perrett, D.I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia. 35, 5,* 685-693.

Butler, R. (1992). What young people want to know when: Effects of mastery and ability goals on interest in different kinds of social comparisons. *Journal of Personality and Social Psychology,* 62, 934-943.

Butler, R. (1989). Mastery vs. ability appraisal: A developmental study of children's observations of peers' work. *Child Development,* 60, 1350-1361.

Butler, R. (1990). The effects of mastery and competitive conditions on self-assessment at different ages. *Child Development,* 61, 201-210.

Butler, R. (1993). Effects of task and ego-achievement goals on information seeking during task engagement. *Journal of Personality and Social Psychology,* 62, 934-943.

Butler, R. A. and Harlow, H. F. (1957). Discrimination learning and learning sets to visual exploration incentives. *Journal of genetic Psychology,* 57, 257-264.

Butler, R. and Neuman, O. (1995). Effects of task achievement goals on help-seeking behaviours and attitudes. *Journal of Educational Psychology,* 87, 261-271.

Chako, T. I. (1983). Students' ratings of instruction: A function of grading standards. *Educational Research Quarterly,* 8 (2), 341-351.

Coplan, J. D., Papp, L. A., Pine, D. M., Cooper, T., Rosenblum, L. A., Klein, D. F., and Gorman, J. M. (1997). Clinical improvement with fluoxetine therapy and noradrenergic function in patients with panic disorder. *Archives-of-General-Psychiatry,* 54, 7, *643-648.*

Covington, M. V. and Omelich, C. L. (1979). Are causal attributions causal? A path analysis of the cognitive model of achievement motivation. *Journal of Personality and Social Psychology,* 37, 1487-1504.

Crespi, L. P. (1942). Quantitative variation in incentive and performance in the white rat. *American Journal of Psychology,* 55. 467-517. Cited in D. A. Lieberman (2000). *Learning: Behaviour and Cognition (3rd Edition).* London: Wadsworth.

Csikszentmihalyi M. and I. S. Csiksentmihalyi (1990). *Optimal experience: Psychological studies of flow in consciousness.* New York: Cambridge University Press.

Csikszentmihalyi, M. (1990). The flow experience and its significance for human psychology. In M. Csikszentmihalyi and I. S. Csiksentmihalyi (Eds). *Optimal experience: Psychological studies of flow in consciousness (1990).* New York: Cambridge University Press.

Csikszentmihalyi, M., and Larson, R. (1987). Validity and reliability of the Experiences-Sampling Method. *The Journal of Nervous and Mental Disease, 175, 9,* 526-36. Cited in M. Csikszentmihalyi and I. S. Csiksentmihalyi (eds*). Optimal experience: Psychological studies of flow in conciousness (1990*). New York: Cambridge University Press.

Csikszentmihalyi, M., Larson, R. and Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adoloescence,* 6, 281-94 cited in M. Csikszentmihalyi and I. S. Csiksentmihalyi (eds*). Optimal experience: Psychological studies of flow in conciousness (1990*). New York: Cambridge University Press.

Davis, R. T., Settlage, P. H. and Harlow, H. F. (1953). Performance of normal and brain-operated monkeys on mechanical puzzles with and without food incentive. *Journal of Genetic Psychology,* 77, 305-311. Cited in E. L. Deci. (1975). *Intrinsic motivation.* New York: Plenum press.

DeCharms, R. (1968). *Personal causation: The internal affective determinants of Behaviour*. New York: Academic Press.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18, 105-115.

Deci, E. L. (1972). Intrinsic Motivation, extrinsic reinforcement and inequity. *Journal of Personality and Social Psychology*, 22, 113-120.

Deci, E. L. (1975). *Intrinsic Motivation.* New York: Plenum Press

Deci, E. L. and Ryan, R. M. (1985). *Intrinsic Motivation and self-determination in human behaviour.* New York: Plenum Press.

Diener, C. I. and Dweck, C. S. (1978). An analysis of learned helplessness: Continuos changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36, 5, 451-462.

Diener, C. I. and Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology*, 39, 451-462.

Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology, 31, 4,* 674-685.

Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist,* 41, 1040-1048.

Dweck, C. S. (1996). Capturing the dynamic nature of personality. *Journal of Research in Personality,* 30, 348-362.

Dweck, C. S. (1996). Implicit theories as organisers of goals and behaviour. In P.M. Gollwitzer and J. A. Bargh (Eds). The Psychology of action: Linking cognition and motivation to behaviour. New York: Guildford.

Dweck, C. S. and Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Journal of Personality and Social Psychology,* 95, 2, 256-273.

Eisenberger, R., Kuhlman, M. and Cotterell, N. (1992). Effects of social values, effort training and goal structure on task persistence. *Journal of Research in Personality,* 26, 258-272.

Elliot, A. J. and Church (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology,* 72, 1, 218-232.

Elliot, A. J. and Harackiewicz, J. M. (1994). Goal setting, Acheivement orientation, and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology,* 66, 5, 968-980.

Elliot, A. J. and Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70, 3,* 461-475.

Elliot, E. S. and Dweck, C. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology,* 54, 1, 5-12.

Elliott, R. L. (1994). Applying quality improvement principles and techniques in public mental health systems. *Hospital-and-Community-Psychiatry,* 45, 5, 439-444.

Epstein, J. A. and Harackiewicz, J. M. (1992). Winning is not enough: The effects of competition and achievement orientation on intrinsic interest. *Personality and Social Psychology Bulletin,* 18, 128-139.

Eysenck, M. W. and Keane, M. T. (1993). *Cognitive Psychology: A student's handbook.* Lawrence Erlbaum Associates. Hillsdale.

Feather, N. T. (1959). Success probability and choice behaviour. *Journal of Experimental Psychology,* 58, 4, 257-266

Feather, N. T. (1961). The relationship of persistence at a task to expectation of success and achievement related motives. *Journal of Abnormal and Social Psychology,* 63, 3, 552-561.

Feather, N. T. (1962). The study of persistence. *Psychological Bulletin,* 59, 2, 94-115.

Feather, N. T. (1963a). The effect of differential failure on expectation of success, reported anxiety, and response uncertainty. *Journal of Personality,* 31, 289-312.

Feather, N. T. (1963b). Persistence at a difficult task with alternative task of intermediate difficulty. *Journal of Abnormal and Social Psychology,* 66, 6, 604-609.

Feather, N. T. (1963c). The relationship of expectation of success to reported probability, task structure, and achievement related motivation. *Journal of Abnormal and Social Psychology,* 63, 3, 231-238.

Feather, N. T. (1966). Effects of prior success and failure on expectations of success and subsequent performance. *Journal of Abnormal and Social Psychology,* 3, 3, 287-298.

Feather, N. T. (1967). Valence and outcome and expectation of success in relation to task difficulty and perceived locus of control. *Journal of Personality and Social Psychology, 7, 4,* 372-386.

Feather, N. T. (1968). Change in confidence following success and failure as a predictor of subsequent performance. *Journal of Personality and Social Psychology,* 9, 1, 38-46.

Feather, N. T. (1969). Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology,* 13, 2, 129-144.

Feldman, K. A. (1976). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education,* 4, 69-111.

Festinger, L. (1954). A theory of Social Comparison Processes. *Human Relations,* 7, 117-140.

Fiske, S. T. and Taylor, S. E. (1984). *Social Cognition.* Reading, Mass: Addison-Wesley.

Fowler, H., and Miller, N. E. (1963). Facilitation and inhibition of runaway performance by hind and forepaw shock of various intensities. *Journal of Comaprative Physiological Psychology,* 56, 801-805. In D. A. Lieberman (2000). *Learning: Behaviour and Cognition* (3rd Edition). London: Wadsworth.

Gollwitzer P.M. and J. A. Bargh (1996). *The Psychology of action: Linking cognition and motivation to behaviour.* New York: Guilford.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist,* 54, 7, 493-503.

Greenwald, A. G. and Gillmore, G. M. (1997a). No pain, No gain? The importance of measuring course workload in students ratings of instruction. *Journal of Educational Psychology*, 89, 743-751.

Greenwald, A. G. and Gillmore, G. M. (1997b). Grading Leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209-1217.

Grolnick, W. S. and Ryan, R. M. (1987). Autonomy in Children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, 52, 5, 890-898.

Harackiewicz, J. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37, 1024-1037.

Harackiewicz, J., Abrahams, S. and Wageman, R. (1987). Performance Evaluation and Intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology*, 53, 6, 1015-1023.

Harackiewicz, J., Manderlink, G. and Sansone, C. (1984). Rewarding pinball wizardry: The effects of evaluation on intrinsic interest. *Journal of Personality and Social Psychology*, 47, 287-300.

Harlow, H. F. (1953a). Motivation as a factor in the acquisition of new responses. Nebraska symposium on motivation, 1, 24-49. Cited in E. L. Deci (1975). *Intrinsic motivation.* New York: Plenum press.

Harlow, H. F. (1953b). Mice, monkeys, men, and motives. *Psychological Review*, 60, 23-32.

Harris, A. M., and Covington, M. V. (1992). The role of co-operative reward interdependency in success and failure. *Journal of Experimental Education,* 61, 2, 151-168.

Heider, F. (1958). *The Psycholgy of interpersonal relations.* New York: Wiley.

Heider, F. and Simmel, M. (1944). An experimental study of apparent behaviour. *American Journal of Psychology,* 57, 243-259.

Holmes, D. S. (1972). Effects of grades and disconfirmed grade expectancies on students' evaluations of their instructors. *Journal of Educational Psychology*, 63, 130-133.

Hull, C. L. (1943). *Principles of Behaviour.* New York: Appleton-Centurt-Crofts.

Kelley, H. H. (1967). Attribution theory in social psychology. Cited in D. Levine (ed.), *Nebraska Symposium on Motivation,* Lincoln: University of Nebraska Press.

Kirsch, I. (1999). Automaticity in clinical psychology. *American Psychologist,* 54, 7, 504-515.

Koestner, R., Bernieri, F. and Zuckerman, M. (1992). Self-regulation and consistency between attitudes, traits and behaviours. *Personality and Social Psychology Bulletin,* 18, 1, 52-59.

Koestner, R., Zuckerman, M. and Olsson, J. (1990). Attributional style, comparison focus of praise and intrinsic motivation. *Journal of Research in Personality,* 24, 87-100.

Koestner, R., Zuckerman, M. and Koestner, J. (1987). Praise, involvement and intrinsic motivation. *Journal of Personality and Social Psychology,* 53, 2, 383-390.

Kruglanski, A. W. (1989). *Lay epistimics and human knowledge: Cognitive and motivational biases.* New York: Plenum press.

Lepper, M. R. and Greene, D. (1975). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology,* 31, 3, 479-486.

Lorusso, S., Poli, V. and Casmiro, M. (1994). Cognitive improvement following treatment in a case of idiopathic hypoparathyroidism. *European-Neurology,* 34, 5, 292-294.

Markowitz, J. C., Friedman, R. A., Miller, N.and Spielman, L. A. (1996). A Interpersonal improvement in chronically depressed patients treated with desipramine. *Journal of Affective Disorders,* 41, 1, 59-62.

Marsh, H. W. and Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist,* 52, 1187-1197.

McClelland, D. C., Atkinson, J. W., Clark, R. A. and Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts..

McFarlane, D. A. (1930). *The role of kinesthesis in maze learning*. University of California publications in Psychology, 4, 277-305. In D. A. Lieberman (2000). *Learning: Behaviour and Cognition* (3rd Edition). London: Wadsworth.

McFarlin, D. B. (1985). Persistence in the face of failure: The impact of self-esteem and contingency information. *Personality and Social Psychology Bulletin,* 11, 2, 153-163.

McFarlin, D. B. and Blascovich, J. (1984). On the Remote Associates Test (RAT) as an alternative to illusory performance feedback: A methodological note. *Basic and Applied Social Psychology*, 5, 223-229.

Miller, A. (1985). A developmental study of the cognitive basis of performance impairment after failure. *Journal of Personality and Social Psychology,* 49, 529-538.

Miller, A. and Hom, H. L. (1990). Influence of extrinsic and ego incentive value on persistence after failure and continuing motivation. *Journal of Educational Psychology*, 82, 3, 539-545.

Miller, A. T. and Klein, J. S. (1989). Individual differences in ego-value of academic performance and persistence after failure. *Contemporary Educational Psychology*, 14, 124-132.

Miller, N. E. (1948). Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *Journal of Experimental Psychology*, 38, 89-101. Cited in B. Weiner (1992). *Human motivation: Metaphors, theories and research*. New York: Sage.

Montgomery, K. C. (1954). The role of exploratory drive in learning. *Journal of Comparative and Physiological Psychology*, 46, 129-133.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Nathanson, D. E. and De-Faria, S. (1993). Cognitive improvement of children in water with and without dolphins. *Anthrozoes*, 6, 1, 17-29.

Nicholls, J. G. (1984). Achievement Motivation: Conceptions of ability, subjective experience, task choice and Performance. *Psychological review*, 91, 3, 328-346.

Nisbett, R. E. and Bellows, N. (1977). Verbal reports about causal influences on social judgements: Private access versus public theories. *Journal of Personality and Social Psychology,* 35, 9, 613-624.

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

Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls: In R. Rosenthal and R. L. Rosnow (Eds.), *Artifacts in behavioural research.* New York: Academic press.

Orne, M.T. (1971). From the subject's point of view, when is behavior private and when is it public: Problems of inference. *Journal of Consulting and Clinical Psychology,* 35,2, 143-147.

Page, M. (1981). Demand compliance in laboratory Experiments. In J. T. Tedeschi (Ed). *Impression Management theory and Social Psychological Research.* New York; Academic Press.

Page, M. M. (1968). Modification of figure-ground perception as a function of awareness of demand characteristics. *Journal of Personality and Social Psychology,* 9, 59-66.

Plant, R. and Ryan. R. M. (1985). Intrinsic Motivation and the effects of self-consciousness, self-awareness and ego-involvement: An investigation of internally controlling styles. *Journal of Personality,* 53, 435-449.

Powell, R. W. (1977). Grades, learning and student evaluation of instruction. *Research in Higher Education*, 7, 193-205.

Redelmeier, D. A. and Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time retrospective evaluations of two minimally invasive procedures. *Pain,* 66, 3-8.

Remedios, R., Lieberman, D. A., and Benton, T. G. (2000). The Effects of Grades Course Enjoyment: Did You Get the Grade You Wanted? *British Journal of Educational Psychology, 70,* 353-368.

Rigby, C. S., Deci, E. L., Patrick, B. C. and Ryan, R. M. (1992). Beyond the Intrinsic-Extrinsic dichotomy: self-determination in Motivation and Learning. *Motivation and Emotion,* 16, 3, 165-185.

Rosenberg, M. (1965). *Society and the Adolescent self-image.* Princeton: Princeton University Press.

Rotter, J. B. (1966). Generalized expectancies for internal vs. external control of reinforcement. *Psychological Monograph,* 80, 1-15.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of Cognitive Evaluation theory. *Journal of Personality and Social Psychology*, 43, 450-461.

Ryan, R. M., Koestner, R. and Deci, E. L. (1991). Ego-involved persistence: When free-choice behaviour is not intrinsically motivated. *Motivation and Emotion*, 15, 3, 185-205.

Ryan, R. M., Mims, V. and Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using Cognitive Evaluation theory. *Journal of Personality and Social Psychology*, 45, 736-770.

Sandelands, L. E., Brockner, J. and Glynn, M. A. (1988). If at first you don't succeed, try, try again: Effects of persistence-performance contingencies, ego involvement, and self-esteem on task persistence. *Journal of Applied Psychology*, 73, 2, 208-216.

Schaffer, R. and Murphy, G. (1943). The role of autism in a visual figure-groun relationship. *Journal of Experimental Psychology*, 32, 335-343.

Schank, R. C. (1979). Interestingness: Controlling inferences. *Artificial Intelligence, 12*, 273-297. Cited in K. A. Renninger., S. Hidi. and A. Krapp (Eds). *The role of interest in learning and development* (pp. 222). Hillsdale: LEA.

Schank, R. C. and Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* In M. W. Eysenck. and M. T. Keane. *Cognitive Psychology: A student's handbook* (1993). Hillsdale:Lawrence Erlbaum Associates.

Schunk, D. H. and Gunn, T. P. (1986). Extended attributional feedback: Sequence effects during remedial reading instruction. *Journal of Early Adolescence,* 6, 55-66.

Schunk, D. H. and Gunn, T. P. (1986). Strategy training and attributional feedback with learning disabled students. *Journal of Educational Psychology*, 49, 843-851.

Scott, D., Scott, L. M. and Goldwater, B. (1997). A performance improvement program for an international-level track and field athlete. *Journal of Applied Behavior Analysis,* 30, 3, 573-575.

Stumpf, S. A. and Freedman, R. D. (1979). Expected grade covariation with student ratings of instruction: Individual vs. class effects. *Journal of Educational Psychology,* 71, 293-302.

Tedeschi, J. T. and Reiss, M. (1981). Identities, the phenomenal self, and laboratory research. In J. T. Tedeschi (Ed). *Impression Management theory and Social Psychological Research.* New York: Academic Press.

Thompson, E. P., Chaiken, S. and Hazlewood, J. D. (1993). Need for cognition and Desire for control as mediators of extrinsic reward effects: A person X situation approach to the study of intrinsic motivation. *Journal of Personality and Social Psychology*, 64, 6, 987-999.

Thorkildsen, T. A. and Nicholls, J. G. (1998). Fifth graders' achievement orientations and beliefs: Individual and classroom differences. *Journal of Educational Psychology*, 90, 2, 179-201.

Tolman, E. C. (1932). *Purposive behaviour in animals and men.* New York: Century.

Vallerand, R. J. and Reid, G. (1984). On the causal effects of perceived competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology*, 6, 94-102.

Vasta, R. and Sarmiento, R. F. (1979). Liberal grading improves evaluations but not performance. *Journal of Educational Psychology*, 71, 207-211.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71, 3-25.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 4, 548-573.

Weiner, B. (1992). *Human Motivation: Metaphors, theories and Research.* London: Sage.

Worthington, A. G. and Wong, P. T. P. (1979). Effects of earned and assigned grades on student evaluations of an instructor. *Journal of Educational Psychology,* 71, 764-775.

Ziegarnik, B. (1927). Das behalten erledigter und unerledigter handlungen. *Psychologische Forschung,* 9, 1-85. Cited in E. L. Deci (1975). *Intrinsic Motivation.* New York: Plenum Press.

List of Appendices

Chapter 3

Chapter 4

Experiment 4: Test Anagrams and Numbers

Easy	Problem	Solution
1.	LATK	TALK
2.	**614**, 2,2,2,6,8, 50	2x50=100; x6=600; 2x8-2=14
3.	HISPEESH	SHEEPISH
4.	**156**, 2,4,4,4,7,50	2x50=100;4+4x7=56
5.	OLITTE	TOILET
6.	**117**, 9,9,9,2,1,100	100+(9+9)-1
7.	YPAHP	HAPPY
8.	**875**, 1,3,5,5,7,100	7+1=8x100=800; 5x5=25
9.	SEXCUE	EXCUSE
10.	**204**, 1,2,4,8,9,50	4x50=200; 8\2=4
11.	LCDO	COLD
12.	**396**, 1,3,4,4,4,100	4x100=400; -4
13.	OLYLNE	LONELY
14.	**451**, 1,4,5,7,7,50	4+5=9x50=450+1
15.	ICDHL	CHILD
16.	**620**, 1,3,3,5,4,100	3+3=6x100=600; 5x4=20.
17.	TPEL	PELT
18.	**425**, 2,2,2,3,5,100	2+2=4x100=400; 2+3=5x5=25
19.	UROS	SOUR
20.	**901**, 1,2,2,5,8,100	8+1=9x100=900; +1
21.	YOHL	HOLY
22.	**136**, 1,2,6,6,8,50	2x50=100; 6x6=36
23.	LDISO	SOLID
24.	**195**, 2,2,2,3,9,50	2+2=4x50=200; 2+3=5
25.	NKHAT	THANK
26.	**749**, 2,3,4,7,7,100	7x100=700; 3+4=7x7=49
27.	JBCOTE	OBJECT
28.	**304**, 1,2,3,4,5,50	5+1=6x50=300; +4
29.	TTLI	TILT
30.	**407**, 1,2,3,5,6,100	1+3=4x100=400; 2+5=7

Easy	Problem	Solution
31.	ALTL	TALL
32.	**205**, 2,6,9,9,9, 100	2x50=100; 9/9=1; 6-1=5
33.	DOTHME	METHOD
34.	**325**, 2,3,5,6,9,50	6x50=300; 2+3=5x5=25
35.	FEHC	CHEF
36.	**564**, 1,1,4,7,9,100	1+4=5x100=500; 7x9=63+1=64
37.	DOPOLE	POODLE
38.	**108**, 1,4,7,7,9,100	100+7+1
39.	IRDAO	RADIO
40.	**430**, 1,2,4,6,8,50	8x50=400; 1+4=5x6=30
41.	RPOGU	GROUP
42.	**599**, 1,3,4,4,6,100	6x100=600; -1= 599
43.	LOPIT	PILOT
44.	**344**, 1,4,4,6,6,50	6x50=300; 6+4+1=11x4=44
45.	ITANCO	ACTION
46.	**610**, 2,3,3,5,4,100	3+3=6x100=600; 5x2=10.
47.	TPRI	TRIP
48.	**725**, 2,2,3,5,5,100	2+5=7x100=700; 2+3=5x5=25
49.	ELOD	DOLE
50.	**999**, 1,2,2,8,9,100	9+1=10x100=1000; 2/2=1; 1000-1=999
51.	TECSLA	CASTLE
52.	**181**, 1,2,6,7,9,50	2x50=100; 2+7=9x9=81
53.	THRAC	CHART
54.	**174**, 2,2,2,3,6,50	3x50=150; 2+2=4x6=24
55.	EOOB	OBOE
56.	**651**, 2,2,4,7,7,100	4+2=6x100=600; 7x7=49+2=51
57.	TRINP	PRINT
58.	**763**, 1,2,6,6,8,50	8x2=16 x 50=800; 6x6=36+1=37; 800-37=763 (DIFF)
59.	TOBUTN	BUTTON
60.	**982**, 2,3,4,6,7,25	4+2=6x6=36+3=39; 39x25=975; 975+7=982 (DIFF)

Easy	Problem	Solution
61.	TALNP	PLANT
62.	**135**, 1,2,2,6,6,50	2X50=100; 6X6=36-1=35
63.	DIVOA	AVOID
64.	**417,** 1,4,4,4,9,100	4x100=400; 4x4=16+1=17
65.	FETRFO	EFFORT
66.	**764,** 1,2,7,7,9,100	100*7=700; 9*7=63+1=64
67.	GRALE	LARGE
68.	**343,** 4,4,6,6,7,50	6*50=300; 7*6=42; 4/4=1
69.	SKAST	TASKS
70.	**531,** 1,2,5,5,6,50	2*5=10*50=500; 6*5=30+1=31
71.	CKLA	LACK
72.	**896,** 1,1,2,9,8,100	9*100=900; 1+1+2=4
73.	ELTSUR	RESULT
74.	**720,** 2,2,5,5,7,50	2*7=14*50=700; 5+5=10*2=20
75.	REAPP	PAPER
76.	**601,** 1,1,2,2,3,50	1+2+3*2=12*50=600+1
77.	DIRFAY	FRIDAY
78.	**244,** 2,2,4,5,6,50	2+2=4*50=200; 6+5=11*4=44
79.	OHLC	LOCH
80.	**975,** 1,3,5,5,9,100	9*100=900; 5*5=25*3=75
81.	ESONT	STONE
82.	**111,** 9,9,9,9,9,100	9/9=1;9/9=1; 9+1+1=11
83.	ODOW	WODO
84.	**423,** 2,2,3,7,8,100	8/2=4*100=400; 7*3=21+2=23
85.	DRELU	LURED/RULED
86.	**636,** 3,3,6,6,9,100	6*100=600; 3+3=6*6=36
87.	ETONCTX	CONTEXT
88.	**743,** 1,3,4,6,7,100	3+4=7*100=700; 6*7=42+1=43
89.	EFLS	SELF
90.	**212,** 1,3,3,3,4,50	1+3=4*50=200; 3*4=12

Easy	Problem	Solution
91.	LIFM	FILM
92.	**581,** 1,5,5,5,8,100	5*100=500; 5+5=10*8=80+1
93.	ADIEM	MEDIA
94.	**965,** 2,5,5,7,9,100	2*5=10*100=1000; 7*5=35
95.	LUSAU	USUAL
96.	**364,** 1,2,3,7,8,50	2*3=6*50=300; 1+7=8*8=64
97.	PESEL	SLEEP
98.	**842,** 3,3,3,4,8,100	8*100=800; 3+3=6; 3+4=7; 7*6=42
99.	ROSECU	COURSE
100.	**923,** 1,1,3,7,9,100	9*100=900; 7*3=21+1+1=23
101.	HYORET	THEORY
102.	**196,** 1,3,5,7,9,50	3*50=150; 9*5=45+1=46
103.	SEYAS	ESSAY
104.	**308,** 2,2,3,4,9,50	2*3=6*50=300; 4*2=8
105.	CIFOEF	OFFICE
106.	**480,** 2,4,4,5,5,100	4*100=400; 5+5=10; 2*4=8; 8*10=80
107.	TILTEL	LITTLE
108.	**706,** 1,1,2,3,6,100	6+1=7*100=700; 2*3=6
109.	YKENOM	MONKEY
110.	**421,** 1,2,2,4,5,100	2*2=4*100=400; 4x5=20+1=21
111.	ELETS	SLEET
112.	**501,** 1,1,2,2,3, 50	2+3=5*2=10*50=500+1=501
113.	NEIMTGE	MEETING
114.	**387,** 2,2,2,2,9, 100	2*2=4*100=400; 2+2+9=13
115.	NISESOS	SESSION
116.	**263,** 1,3,5,7,9,50	1+3=4*50=200; 7*9=63
117.	CTKSO	STOCK
118.	**917,** 1,1,3,5,9,100	9*100=900; 5*3=15+1+1=17
119.	DYELDA	DEADLY
120.	**250,** 1,1,2,9,9, 50	1+1+2=4; 9/9=1, 4+1=5x50=250

Hard	Problem	Solution
1.	TISSEYLID	STYLISED
2.	**758**, 1,2,5,6,7, 100	7+1x100=800; 5x2=7x6=42
3.	AROCBUDP	CUPBOARD
4.	**312**, 1,2,2,6,9,25	6x2=12x25=300; 9+1+2=12
5.	PAGDELPR	GRAPPLED
6.	**478**, 1,2,3,3,6,25	6x3=18+1=19x25=475+3=478
7.	SVOEINU	ENVIOUS
8.	**825**, 1,2,4,5,8,50	2x8=16x50=800; 1+4=5x5=25
9.	RATSGETY	STRATEGY
10.	**155**, 3,4,5,6,6,50	6/3=2x50= 100; 6+5=11x 5 =55
11.	TICENAR	CERTAIN
12.	**567**, 1,2,5,6,8,50	5+6=11x50=550; 2x8=16+1=17
13.	RACEVIA	AVARICE
14.	**903**, 2,6,6,6,8,25	6x6=36x25=900; 6/2=3
15.	REMPLBO	PROBLEM
16.	**276**, 1,2,3,3,8,100	3x100=300; 8x3=24; 300-24=276
17.	EEGBCRI	ICEBERG
18.	**721**, 3,3,4,5,9,50	9+5=14x50=700; 4+3=7x3=21
19.	GOATARNR	ARROGANT
20.	**392**, 2,3,3,5,7,25	3x5=15x25=375; 2x7=14+3=17; 375+17=392
21.	NRATTY	TYRANT
22.	**433**, 2,2,5,7,9,100	5x100=500; 7x9=63+(2+2)=67; 500-67=433
23.	ROESWTFA	SOFTWARE
24.	**662**, 1,2,3,6,9,25	9x3=27x25=675; 6x2=12+1=13; 675-13=672
25.	SITSTASNA	ASSISTANT
26.	**591**, 2,2,3,3,6,25	6x(2+2)=24x25=600; 3x3=9; 600-9=591
27.	RIGAMADP	PARADIGM
28.	**971**, 2,3,4,5,9,50	5x4=20x50=1000; 9x3=27+2=29; 1000-29=971
29.	CIAPLACRT	PRACTICAL
30.	**973**, 2,2,3,3,6,25	3x2=6x6=36+3=39; 39x25=975; 975 - 2=973

Hard	Problem	Solution
31.	MIMESER	IMMERSE
32.	**731**, 3,3,6,7,8,100	8x100=800; 3x3=9x7=63+6=69; 800-69=731
33.	EMRTOAVNI	NORMATIVE
34.	**942**, 1,6,6,8,9,25	6x6=36+1=37x25=925; 8+9=17
35.	NEEEPTOCCM	COMPETENCE
36.	**606**, 4,5,6,6,7,50	6+7=13x50=650; 6+5=11x4=44; 650-44=606
37.	POMEDIL	IMPLODE
38.	**349**, 1,1,2,2,4,25	1+2+4=7*2=14*25=350-1
39.	INTENSET	SENTIENT
40.	**264**, 1,2,3,5,6,100	3x100=300; 1+5=6*6=36
41.	YIDOLSTI	SOLIDITY
42.	**879**, 2,3,8,9,9,50	8+9=17x50=850; 9x3=27+2=29
43.	GREEDHOW	HEDGEROW
44.	**565**, 3,3,3,4,5, 25	3+4=7x3=21x25=550; 3x5=15
45.	ACTIONAR	RAINCOAT
46.	**471**, 2,2,3,5,9, 50	2x5=10x50=500; 9x3=27+2=29
47.	BILIMEOM	IMMOBILE
48.	**650**, 7,7,7,7,7, 100	7x100=700; 7x7=49; 7/7=1
49.	PARTLY	PALTRY
50.	**431**, 2,3,3,3,7, 50	3x3=9x50=450; 3x4=21-2=19; 450-19=431

Experiment 4: Anagrams + Numbers - Worksheet

Anagram 1	Numbers 1
Anagram 2	Numbers 2
Anagram 3	Numbers 3
Anagram 4	Numbers 4
Anagram 5	Numbers 5
Anagram 6	Numbers 6
Anagram 7	Numbers 7
Anagram 8	Numbers 8

Participants were given 4 of these sheets so they could attempt up to 32 anagrams and 32 numbers problems.

Your experiences

Appendix 3

How much 'FUN' would you say that task was?

Lots of Fun	6	5	4	3	2	1	Not Fun

How much did you Enjoy solving the problems?

A lot	6	5	4	3	2	1	Not at all

How interesting was the task as a whole?

Very interesting	6	5	4	3	2	1	Uninteresting

On the whole, how difficult was the task?

Difficult	6	5	4	3	2	1	Easy

How well did you think you did at the task?

Well	6	5	4	3	2	1	Badly

How well do you think others would do at the task?

Well	6	5	4	3	2	1	Badly

Think about how you did. On the whole, was this down to your own ability or the inherent difficulty/easiness of the problems you had to solve?

Yourself	6	5	4	3	2	1	The Problems

Think about why you stopped solving the problems and rate the following statements on how true/untrue they are.

You had enough information to be able to comment on how much fun the task was

True	6	5	4	3	2	1	Untrue

The tasks were too easy

True	6	5	4	3	2	1	Untrue

The tasks were too hard

True	6	5	4	3	2	1	Untrue

You got bored

True	6	5	4	3	2	1	Untrue

You felt you had spent enough time on the tasks

True	6	5	4	3	2	1	Untrue

Experiment 5: Anagrams and their solutions

	EASY problem	EASY solution	HARD problem	HARD solution
		.		
1	LATKS	STALK	TISSEYLD	STYLISED
2	YPAPH	HAPPY	AROCBUDP	CUPBOARD
3	LCDOS	SCOLD	PAGDELPR	GRAPPLED
4	ICDHL	CHILD	SVOEINU	ENVIOUS
5	PTLES	SLEPT	RATSGETY	STRATEGY
6	LIDSO	SOLID	TICENAR	CERTAIN
7	NKHAT	THANK	RACEVIA	AVARICE
8	TTNIA	TAINT	REMPLBO	PROBLEM
9	OLITTE	TOILET	EEGBCRI	ICEBERG
10	ALTSL	STALL	GOATARNR	ARROGANT
11	FEHC	CHEF	ROESWTFA	SOFTWARE
12	IRDAO	RADIO	SITSTASNA	ASSISTANT
13	RPOGU	GROUP	RIGAMADP	PARADIGM
14	DOPOLE	POODLE	CIAPLACRT	PRACTICAL
15	TPERI	TRIPE	MIMESER	IMMERSE
16	ELORD	DROLE	EMRTOAVNI	NORMATIVE
17	THRAC	CHART	NEEEPTOCCM	COMPETENCE
18	EOOB	OBOE	POMEDIL	IMPLODE
19	TRINP	PRINT	INTENSET	SENTIENT
20	TOBUTN	BUTTON	YIDOLSTI	SOLIDITY
21	TALNP	PLANT	GREEDHOW	HEDGEROW
22	DIVOA	AVOID	ACTIONAR	RAINCOAT
23	GRALE	LARGE	BILIMEOM	IMMOBILE
24	SKAST	TASKS	ROIIATSLE	SOLITAIRE
25	CKLAB	BLACK	LABAALVIE	AVAILABLE
26	REAPP	PAPER	DREACTPU	CAPTURED
27	DIRFAY	FRIDAY	FCAOLFII	OFFICIAL
28	OHLC	LOCH	CAMMSIEHN	MECHANISM
29	ESONT	STONE	NSIRDGSE	DRESSING
30	ODOW	WOOD	MEEPRRO	EMPEROR

	EASY problem	EASY solution	HARD problem	HARD solution
31	EFLS	SELF	FGUESRFA	SUFFRAGE
32	LIFMS	FILMS	PRAMOTEM	METAPHOR
33	MADIE	MEDIA	COOCLANIS	OCCASIONAL
34	LUSAU	USUAL	ICEDORIP	PERIODIC
35	PESEL	SLEEP	MEELNHI	HEMLINE
36	SSEYA	ESSAY	YRBEREW	BREWERY
37	YKENOM	MONKEY	EATNISCA	INSTANCE
38	CTKSO	STOCK	ENPSICEM	SPECIMEN
39	KILYS	SILKY	CESSNITMA	SEMANTICS
40	POSUT	SPOUT	DRODHASBA	DASHBOARD
41	MADER	DREAM	STWEDANAL	WASTELAND
42	MAFEL	FLAME	ALATIGEC	GLACIATE
43	LANPT	PLANT	RETREBHN	BRETHREN
44	OSNOW	SWOON	FARFIGE	GIRAFFE
45	BLILE	LIBEL	DIAYTIC	ACIDITY
46	ACSNK	SNACK	DRAGEESS	DRESSAGE
47	DIFER	FRIED	RAFTWELAL	WATERFALL
48	AYUGD	GAUDY	DORSUQAN	SQUADRON
49	SIWTT	TWIST	DIMYROBIT	MORBIDITY
50	LDOCU	CLOUD	LICELOFL	FOLLICLE
51	ELSEP	SLEEP	STEPPIRR	STRIPPER
52	BUDTE	DEBUT	ANDEJCIU	JAUNDICE
53	DEVTU	DUVET	TALEGRUF	GRATEFUL
54	OKTSC	STOCK	SLANGIPE	PLEASING
55	FICFL	CLIFF	TIGGLIHHH	HIGHLIGHT
56	IRESA	RAISE	RAILSDUE	RESIDUAL
57	LITAR	TRIAL	FEWYORL	FLOWERY
58	RTIFL	FLIRT	TOABKUTCL	BLACKOUT
59	GITHF	FIGHT	ENTTIREC	RETICENT
60	DMMAA	MADAM	GLOWSAL	GALLOWS

	EASY problem	EASY solution	HARD problem	HARD solution
61	MOUDL	MOULD	GOLFTHOIT	FOOTLIGHT
62	HLICL	CHILL	KJOSITCY	JOYSTICK
63	RAYSC	SCARY	MILESTONE	LIMESTONE
64	THLOS	SLOTH	DAHLIOYS	HOLIDAYS
65	NHCUP	PUNCH	JOBEERMA	JAMBOREE
66	POCMH	CHOMP	GLINTFIR	FLIRTING
67	OTTID	DITTO	DAUNLOCR	CAULDRON
68	IDLES	SLIDE	GUNCUMIVA	VACUUMING
69	REDEG	GREED	SLEETFAL	LEAFLETS
70	RESDS	DRESS	AXLEMPES	EXAMPLES
71	THUDC	DUTCH	ELGINFEC	FLEECING
72	ECOSH	CHOSE	BANEMICE	AMBIENCE
73	HADES	SHADE	LISTYRCT	STRICTLY
74	RAQUYR	QUARRY	ONEXVATI	VEXATION
75	NAHYD	HANDY	TELLMANY	MENTALLY
76	TIRAE	IRATE	STRANDDUNE	UNDERSTAND
77	PLEAT	PLATE	JAGINCOL	CAJOLING
78	ERRUC	RECUR	DTHGIBLE	BLIGHTED
79	EGUDL	GLUED	SIMSTRUT	MISTRUST
80	INADR	DRAIN	OGBDAVAN	VAGABOND
81	SIYON	NOISY	GITLAPHNY	PLAYTHING
82	PLOEE	ELOPE	DORANUYB	BOUNDARY
83	CHEFT	FETCH	SHIVELEC	VEHICLES
84	DORUN	ROUND	KCOWLIESC	CLOCKWISE
85	AKNLY	LANKY	STRIPEDES	PERSISTED
86	NODPU	POUND	GRAFTRAN	FRAGRANT
87	DELIG	GLIDE	DATOTETO	TATTOOED
88	YANTS	NASTY	SLIBSLUF	BLISSFUL
89	ARSLN	SNARL	RENTCONIU	CENTURION
90	WOLLEY	YELLOW	CROWKFART	CRAFTWORK

	EASY problem	EASY solution	HARD problem	HARD solution
91	NIBLD	BLIND	SHIRLOTACI	HISTORICAL
92	SHATE	HASTE	DETAPCAL	PLACATED
93	KAJLAC	JACKAL	DENTIAFUAT	INFATUATED
94	TELFC	CLEFT	EOPONDSI	POISONED
95	STRUT	TRUST	FREAKSTAB	BREAKFAST
96	LEBLY	BELLY	RANDOMATY	MANDATORY
97	KOPAL	POLKA	SLIDDENP	SPLENDID
98	GLOTA	GLOAT	FEWDILLI	WILDLIFE
99	CABON	BACON	DECORUNT	TROUNCED
100	DACRS	CARDS	FUZYLESTL	ZESTFULLY
101	REFLI	FLIER	EAENQUILTS	SEQUENTIAL
102	REDFE	FREED	TIMESDURST	MISTRUSTED
103	SOWER	SWORE	REWARDSEN	WANDERERS
104	GIRLA	GRAIL	QUEERDANDS	SQUANDERED
105	EETTN	TENET	GILLOXENT	EXTOLLING
106	VOLES	SOLVE	EXLOGOFLYER	REFLEXOLOGY
107	BURMA	RUMBA	BLIGUIDSN	BUILDINGS
108	HIDES	SHIED	SHAMDONE	HANDSOME
109	DECID	DICED	GASSGREPN	PRESSGANG
110	ALETD	DEALT	PANEDEXIL	EXPLAINED
111	RELAP	PEARL	GHASTIRT	STRAIGHT
112	WOCRN	CROWN	DRESHYLW	SHREWDLY
113	PULSR	SLURP	DIFINGDL	FIDDLING
114	DEEVA	EVADE	EQUATEBEHED	BEQUEATHED
115	TORUD	DROUT	SIEVELUTB	VESTIBULE
116	TABLE	BLEAT	STOCISLER	CLOISTERS
117	BHCUT	BUTCH	YALCACHINP	CHAPLAINCY
118	TALUF	FAULT	REMADAUR	MARAUDER
119	DIREC	CRIED	FILEDYINET	DEFINITELY
120	YRYLD	DRYLY	SUBOILI	BILIOUS

	EASY problem	EASY solution
121	SEPEH	SHEEP
122	LISKL	SKILL
123	ACHIN	CHAIN
124	SOLAS	LASSO
125	ZARCE	CRAZE
126	BASTE	BEAST
127	BLAME	AMBLE
128	FEWAR	WAFER
129	ROSTY	STORY
130	IRERC	CRIER
131	OSGHT	GHOST
132	RETED	DETER
133	NORFD	FROND
134	UTONC	COUNT
135	COSTO	SCOOT
136	MAPLE	AMPLE
137	AXLER	RELAX
138	KOPNL	PLONK
139	PULCM	CLUMP
140	HASKR	SHARK
141	REDEC	CREED
142	TUFLE	FLUTE
143	PIEDT	TEPID
144	VOLGE	GLOVE
145	PALEP	APPLE
146	DAREY	READY
147	NAGRD	GRAND
148	RODUT	DROUT
149	LEFIT	FILET
150	PLAMC	CLAMP

	EASY problem	EASY solution
151	DALEP	PLEAD
152	IXMMA	MAXIM
153	SIDLE	SLIDE
154	TOGAR	GROAT
155	IPCRM	CRIMP
156	MASEX	EXAMS
157	LUMPE	PLUME
158	TOYFL	LOFTY
159	ALOCK	CLOAK
160	FATLO	FLOAT
161	YITDR	DIRTY
162	LOFYL	FOLLY
163	COVIE	VOICE
164	DICHE	CHIDE
165	STUDY	DUSTY
166	NIMSAJ	JASMIN
167	IRLDL	DRILL
168	UHOSE	HOUSE
169	TRAVE	AVERT
170	IVDIL	LIVID
171	CHUPO	POUCH
172	CHINE	NICHE
173	KULFE	FLUKE
174	STOAT	TOAST
175	ROCKA	CROAK
176	FRASC	SCARF
177	FREWE	FEWER
178	CHEFT	FETCH
179	YILLD	IDYLL
180	DOVEL	LOVED

	EASY problem	EASY solution
181	NUAGT	GAUNT
182	NIPOT	POINT
183	SITOJ	JOIST
184	FREER	REFER
185	HICED	CHIDE
186	SINAB	BASIN
187	WESRE	SEWER
188	DOSUN	SOUND
189	CHATM	MATCH
190	VATUL	VAULT
191	DERIP	PRIDE
192	GITHH	THIGH
193	SECRT	CREST
194	STAWE	WASTE
195	DATER	TRADE
196	GLEBA	GABLE
197	STYUM	MUSTY
198	PRIGE	GRIPE
199	RAYIH	HAIRY
200	TOLUC	CLOUT
201	WEEDS	SWEDE
202	DECTA	CADET
203	REBTE	BERET
204	PEELB	BLEEP
205	TOYFS	SOFTY
206	SWORT	WORST
207	STULY	LUSTY
208	DEBRI	BRIDE
209	CEDER	CREED
210	MEATS	STEAM

Experiment 5: Anagram worksheet

Anagram Worksheet

Please show all your working out on these sheets.
Note the code (i.e. E156; H56) of every problem you have attempted, even if you were unable to solve the anagram.

Please circle as appropriate

On the whole I am satisfied with myself

strongly agree agree disagree strongly disagree

At times I think I am no good at all

strongly agree agree disagree strongly disagree

I feel I have a number of good qualities

strongly agree agree disagree strongly disagree

I am able to do things as well as most other people

strongly agree agree disagree strongly disagree

I feel I do not have much to be proud of

strongly agree agree disagree strongly disagree

I certainly feel useless at times

strongly agree agree disagree strongly disagree

I feel I am a person of worth, at least equal with others

strongly agree agree disagree strongly disagree

I wish I could have more respect for myself

strongly agree agree disagree strongly disagree

All in all, I am inclined to feel that I am a failure

strongly agree agree disagree strongly disagree

I take a positive attitude towards myself

strongly agree agree disagree strongly disagree

Thank you for completing this questionnaire.

Please be assured that your answers will be treated in the <u>strictest</u> confidence