Agents' Performance and Emotions: An Experimental Investigation

Emanuela Lezzi

Submitted for the degree of Doctor of Philosophy

University of East Anglia

School of Economics

March 2014

Abstract

This doctoral thesis is structured in three essays. In the first essay (Chapter 2) I explore the behavioural effects of anxiety on agents' performance. I hypothesize that a certain level of tension and pressure can induce agents to exert more effort, according to theories of anxiety in psychology. The negative valence associated to this emotion might propose an impairment in performance. On the contrary, the laboratory economic experiment I have run shows that when an anxious mood is induced individuals are more likely to exert more effort. Anxiety leads to performance improvements.

In the second essay (Chapter 3) I raise a methodological issue on the use of effort tasks in economic experiments. Effort tasks are usually assumed to lead to similar results. However, the choice of the effort task can significantly drive experimental results. I have conducted an economic experiment where I compare four different effort tasks which give a measure of participants' performance or investment when they compete for a prize. Results show that there is no equivalence between the types of task applied.

The last essay (Chapter 4) is a substantial part of a joint project with Professor Daniel J. Zizzo. We ran an experiment where participants are asked to enter a 2-player prize competition. Each pair consists of a High Type participant, who performs a previous real effort task better, and a Low Type participant, who performs a previous real effort task worse. Participants receive feedback on their performance rank and their opponents' performance rank. They are also informed about the allocation of an extra monetary reward. Participants are then asked to choose their level of investment. They can also sabotage their opponent. Results show that perceived unfairness of the reward allocation rule, expectations of investment and sabotage, and competitive feelings affect participants' behaviour in the contest.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my sincere gratitude to all those who made it possible for me to complete this thesis. In particular, I am deeply indebted to my supervisors Daniel John Zizzo and Piers Fleming for their invaluable support. Any professional and personal satisfaction that comes from this intense intellectual experience would have not been possible without their professional suggestions, critical and constructive discussions, and constant supervision.

I wish to thank my examiners Robert Sugden and Luigi Mittone for their valuable comments received after my viva.

I am also grateful to Subhasish Modak Chowdhury, Liliana Harding and Kei Tsutsui for their assistance and insights.

Thanks to Gina, Helen and Shiona, and all the research students and faculty members for their friendly support.

My heartfelt thanks go to my friends Maite, Maria and Andrew, Pinar, Antoni, Cigdem and Okan, Burcak and Umut, Arabella and her kids: it is an honour to consider you my friends.

And finally, a thank you to my husband and best friend Fabio whose love and encouragement have always been a source of inspiration.

This thesis is dedicated to my parents Rita e Luigi. They have shown me that sacrifice, love and unity are the secrets of happiness.

Chapter 1

Introduction

This thesis has two objectives: it explores how anxiety affects agents' performance and it discusses a methodological issue on the use of effort tasks in economic laboratory experiments.

It first tries to investigate whether emotions, fairness and expectations affect agents' performance. It examines the effects of anxiety on workers' performance. It also studies how affective states and expectations drive the level of investment, and therefore performance, when workers compete in a contest. Moreover, it analyses how perceived unfairness and hostile emotions induce workers to engage in destructive effort that reduces opponents' performance.

Differently from the current research in economics which considers emotions in different contexts with respect to risk attitude and in relation to social norms and preferences in interpersonal interactions (Hopfensitz and van Winden, 2008; Bosman and van Winden, 2010; Reuben and van Winden, 2005; Ben-Shakhar et al., 2007), this study sheds light on the effect of anxiety on workers' performance at an individual task under a piece rate incentive scheme and on the effects of hostile emotions on workers' performance in tournaments.

The validity of this research topic is supported by a large psychological literature on affective states which nowadays builds on established findings (Isen et al., 1978; Nasby and Yando, 1982; Isen, 1987). Psychologists have shown that emotions have strong and predictive power on decision-making processes and that they can affect consistently individuals' behaviour (Isen and Means, 1983; Isen et al., 1991; Murray et al., 1990; Mackie and Worth, 1991; Schwarz et al., 1991; Ellis et al., 1984, 1985; Parrott and Sabini, 1990). For example, Isen and Patrick (1983); Isen and Geva (1987); Arkes et al. (1988) show that people in whom positive affect had been induced were more risk averse than people in control conditions, when the risk situation was a real one in which they could have a meaningful loss. Otherwise, they appeared more risk-prone (Isen and Patrick, 1983).

Secondly, this thesis raises a methodological issue on the use of effort tasks in economic experiments. Effort tasks are laboratory tools that allow experimentalists to collect measurements on individuals' level of investment or performance (Bull et al., 1987; Harbring and Irlenbusch, 2005; Vesterlund and Nierderle, 2007; Abeler et al., 2011; Gill and Prowse, 2011). Effort tasks are implicitly assumed to lead to similar results as far as they can minimize personal skills involved in the task. Effort tasks can be categorized in two groups: 1) *induced value effort tasks*, where individuals select a value that would represent their chosen performance, and 2) *real effort tasks*, tasks that imply a physical or mental effort to be performed. Real and induced value effort tasks have been implicitly assumed to be equivalent.

However, the choice of the effort task may considerably affect experimental data. This is due to the substantial difference between effort tasks. This thesis provides evidence of the non-equivalence of the effort tasks. It aims to point out that the choice of an effort task may affect the laboratory results and that any interpretation of them should acknowledge the nature of the task used.

The following chapters of this dissertation apply the experimental methodology.

This methodology has been chosen because of the nature of the data I aim to analyse. In fact, it seems difficult to gather data on emotions in the field. Even if some field data are available, they suffer from noise, identification problems, and lack of control. The laboratory allows the researcher the use of a controlled environment in which variations in individual factors can be tested while keeping all others constant. In laboratory experiments it is possible to create analogous environments that mimic real-life scenarios. Although stylised, they have the important advantage of providing a great deal of control over relevant margins.

Some experiments I carried out are enriched by the use of experimental psychology techniques that facilitate the test of behavioural hypotheses. The combined use of both economic experimental methodology and psychological techniques underlines the interdisciplinary feature of my research.

This doctoral thesis is structured in three substantial chapters. In Chapter 2 I explore the behavioural effects of anxiety on workers' performance. Specifically, I investigate the motivational role of an anxious mood on individual performance. According to a large psychological literature, anxiety consistently affects individuals' performance and, therefore, it is considered a determinant drive. The negative valence associated to this emotion might propose an impairment in performance. On the contrary, I hypothesize that a certain level of tension and pressure can actually induce workers to exert more effort.

I ran two economic experiments (Experiment 1 and Experiment 2) where participants were asked to carry out a real effort task, the slider task by Gill and Prowse (2010). In this task participants were asked to position as many sliders as they could in two minutes at a certain value indicated in the instructions. Participants performance was given by the number of sliders correctly positioned in the time limit. Before carrying out the real effort task participants' mood was manipulated. An established mood inducement technique called *Personal Recollection*, widely used in experimental psychology, was applied: students were asked to recall two personal events in which they felt very anxious. The mood manipulation lasted for 20 minutes. Before and after the mood induction stage participants elicited their mood with the use of a Likert Scale.

Results shows that an increase in anxiety leads to performance improvements and induces subjects to exert more effort. Data support the Processing Efficiency Theory by Eysenck and Calvo (1992). According to their formulation the individual's initial effect to worries is negative because worries interfere with attention to the task-relevant information, thus reducing the resources potentially available for the task. Consequently, performance is impaired. However, "[...] an increase in anxiety can therefore improve performance if the worker is motivated to increase effort to such an extent that it counterbalances and exceeds the initial negative impact of worry (p. 415)". Therefore, this essay shows that anxiety has a motivational function that induces the subject to increase effort and therefore to boost performance.

This essay adds an additional contribution to the economic literature: the causal link between personality and behaviour. I measured the proneness to anxiety either in a separated experimental session (Experiment 1) or in the same experimental session (Experiment 2) using the State Trait Anxiety Inventory by Spielberger (1972). I found that personality counts: participants with proneness to anxiety (High Trait-Anxiety individuals) are more affected by threatening stimuli and, therefore, they achieve high level of performance. These findings support the claim that a sufficiently high level of anxiety may act as incentive and that anxiety arousal and anxiety trait increase workers' performance. Hence, they should be taken into account for contract and job design.

Chapter 3 discusses the methodological issue on the use of effort tasks in economic experiments. To show that the choice of the task substantially affects the quality of the experimental results I ran an economic experiment which consisted of four treatments identical except for the effort task applied. Three real effort tasks, the *Slider Task*, the *Maths Task*, the *Grid Task*, and an induced value effort task were implemented in the lab.

The "Slider Task" by Gill and Prowse (2011) required participants to use their computer mouse to position up to forty-eight sliders, having a range of integer values from 0 to 100 and appearing at position zero, in the central position (value 50). Students had two minutes to correctly position as many sliders as they could. Subjects' performance was given by the number of sliders correctly positioned in two minutes. For the "Maths Task" (Vesterlund and Nierderle, 2007) participants were asked to sum a series of four 2-digit numbers that appeared on their screen. They were provided of a list of operations and they were asked to solve as many as they could in the time constraint (two minutes). The number of correct answers gives the participant's performance. Participants dealing with the "Grid Task" (Abeler et al., 2011) saw a 5 by 5 grid on the screen and were asked to count the number of ones randomly assigned in the cells. They had two minutes to solve as many grid as they could. Performance was given by the number of grid correctly solved. Finally, in the treatment in which the "Induced-Value Effort Task" was used participants were endowed with ten experimental points and they had to decide how much to invest and how much to keep for themselves. The part of endowment not invested (if any) was converted in pounds and paid to subjects at the end of the experiment. Participants with the highest investment won a prize.

The effort tasks, thus, provided a measure of participants' level of performance (or investment). Participants were asked to compete in a two-player all-pay auction in order to win a prize. Those with the highest performance could get the prize. In this competitive setting I expect that an anxious emotional state will have a positive effect on subjects' level of performance (or investment). I expect that the more risk averse contestant will exert less effort, and that there will gender difference in performance. These effects should not depend on the effort task used.

Results show that there is no equivalence between the types of task applied. There is evidence that participants' performance is differently affected by the same group of explanatory variables across treatments: the more the anxiety level the worse the level of effort when participants deal with the mathematical task and the grid task. On the contrary, the effects of anxiety on individuals' performance take completely different directions when either the slider task or the induced value effort task are used. There is also evidence that a high degree of risk aversion positively affects subjects' performance when the grid task is applied but it does not have any significant impact on performance when the other effort tasks are used. This suggests that the nature of the real effort task strongly affects experimental results. Hence, researchers using the economic experimental methodology should be aware of the limitation of the effort tasks tool and should acknowledge it in their work.

Chapter 4 is a substantial part of a joint work with my supervisor Daniel J. Zizzo. We aim to analyse the effects of three types of allocation rules of monetary rewards on individuals' decisions of investment and sabotage in tournaments. We conducted an experiment where participants were initially asked to perform a real effort task for five minutes, the counting task (Abeler et al., 2011). At the end of the real effort task their performance was ranked. Two groups were created: the High Type participants group, who performed better in the task, and the Low Type participant group, who performed worse in the task. Participants received information on their own type, their opponent's type, and the possibility to be given a bonus of an extra £5. Moreover, they were informed about the bonus allocation rule to be applied: randomly between contestants, according to a meritocracy rule (High Type participants received the bonus), or according to an anti-meritocracy rule (Low Type participants received the bonus). The allocation rule differed across treatments: in one treatment High Type participants were rewarded (meritocracy rule), in a second treatment the Low Type subjects received the reward (anti-meritocracy rule), and finally in a control treatment, the bonus was randomly assigned to either a High Type or a Low Type subject (random allocation).

Successively, participants were asked to enter a 2-player prize competition. Each pair consisted of a High Type participant and a Low Type participant. They were asked to choose their level of investment. They could also sabotage their opponents' level of investment and, therefore, reduce their opponents' performance and probability to win the prize.

Results show that Low Type participants were more willing to invest when they were not rewarded and to sabotage their opponents' investment when they were rewarded for their comparatively poor performance in a previous task. Their decision of investment and sabotage depended on their expectations of their opponents' level of investment and sabotage. Low Type participants provided explanations about their behaviour in the contest. They had a clear understanding of the fairness of the allocation procedures and specifically, when unfairly rewarded, they felt competitive feelings that affected their behaviour in the tournament. These results suggest that the allocation procedures for monetary rewards prior the competition affect contestants' behaviour and therefore they should be taken into account in contest design with respect to the contest's objectives.

Overall, this thesis shows how some emotions affect agents' performance. Moreover, it considers the role of perceived unfairness and expectations on agents' decisions of investment and sabotage in contests. Emotions, fairness and expectations have a substantial impact on workers' behaviour to be acknowledged in the discussion of contract design and contests design. Finally, this thesis provides an important insight on the limitation of the effort tasks tool and open a debate on the appropriate use of them in the experimental methodology.

Chapter 2

Motivation and Emotions: Effects of Anxiety on Workers' Performance

2.1 Introduction

Incentives are the essence of economics. The most basic concept of demand considers how to induce a consumer to purchase. Similarly, supply relationships are descriptions of how agents respond with more output or labour for additional compensation. Incentive problems arise when a principal wants to delegate a task to an agent[1]. The principal wants to induce the agent to behave in a way that is beneficial to him but does not have perfect knowledge of the agent's actions. The agent dislikes labour and her/his interests may not be aligned with the principal's interests. A large literature has pointed to a multitude of different mechanisms that can be used

[1]Delegation can be motivated either by the possibility of benefiting from some increasing returns associated with the division of tasks or by the principal's lack of time or lack of any ability to perform the task himself or by any other form of the principal's bounded rationality to face complex problems. However, by the mere fact of this delegation the agent may get access to information that is not available to the principal (Laffont and Martimort, 2002).

to induce agents to act in the interest of their principals[2].

However, emotions have not been considered in standard agency theory as incentives (Baker et al., 1987) and very few attempts have been made in order to implement emotional factors in contract design and in personnel management (Lazear and Oyer, 2000; Lazear, 1989).

This essay looks at the motivational role of one particular emotion, anxiety, and specifically investigates the effects of anxiety on individuals' performance. According to a large psychological literature anxiety consistently affects individuals' performance and therefore it is considered a determinant driver. The negative valence associated with this emotion might cause impairment in performance. However, in this work I hypothesize that a certain level of tension and pressure can actually induce agents to exert more effort. I conducted two economic laboratory experiments where I asked subjects to carry out a real effort task after having manipulated their current mood by using an established mood inducement technique. I measured their level of anxiety arousal and their propensity to anxiety, controlling for any confound effects due to these measurements. Results show that an increase in anxiety leads to performance improvements and subjects with a high propensity to anxiety (High Trait Anxiety) exert more effort as soon as their emotional arousal increases.

2.2 Related Literature

The received view in the standard incentive theory is that workers respond to incentives. Specifically, an increase in financial incentives provided for an activity is expected to induce workers to supply more output, or similarly, to increase their performance (Laffont and Martimort, 2002; Prendergast, 1999). *The theory of*

[2]For example options, discretionary bonuses, promotions, profit sharing, efficiency wages, deferred compensation, and so on (Lazear, 1986; Fama, 1991; Baker, 1992; Lazear and Rosen, 1981; Prendergast, 1999; Baker et al., 1987).

compensation represents the core of personnel economics: performance is positively related to effort, workers dislike labour, money is the main incentive and the relationship between monetary compensation for an activity and the level of effort exerted to carry out that activity is monotonic (Lazear, 2000a; Lazear and Oyer, 2007; Lazear, 2000b, 1999; Gibbons and Waldman, 1999; Paul R. and Roberts, 1992). Many economic models have been offered to explain this relationship, for example, in the specific form of piece rate (which is defined to be payment on the basis of output)[3] or, in general, in the form of different compensation schemes (Seiler, 1984; Brown, 1992; Brown and Philips, 1986; Goldin, 1986; Drago and Hyewood, 1995)[4].

The economic literature on the provision of incentives in firms[5] has so far neglected important factors like emotional status and traits in the design of contracts and personnel managerial practices. Few attempts have been made in this direction in economics: Lazear (1989) discusses the conditions under which *personality-contingent wages* are not equivalent to segregation practices and addresses the problems connected to the observability of personalities and the pay inequality. Generally, psychological factor concerns arise in the "mix and matching" processes where "matching the right workers to the right firms as well as matching workers to the most appropriate jobs within the firms creates economic value of a magnitude that few other economic processes can" (Lazear and Oyer, 2000). To my knowledge,

[3]Lazear (1986) provides a detailed discussion of when to pay a piece rate. Fama (1991) discusses other reasons for paying on the basis of some measured time interval. Baker (1992) points out the difficulty created by pay-for-performance payment scheme when performance measurement is a problem.

[4]A recent approach assumes that employees' get intrinsic motivation from the creation of their own output (Deci, 1972; Frey and Oberholzer-Gee, 1997; Hackman and Lawler, 1971; Hackman, 1980; Murdock, 2002) and hence, takes into account extrinsic and intrinsic motivation in the formulation of organizational and personnel policies. A review of the literature of intrinsic and extrinsic motivation is out of the purpose of this work because I will specifically focus on anxiety as an emotion that can act as incentives distinguishing its effects from the effects of any intrinsic motivation.

[5]See Prendergast (1999) for a survey.

no work in economics specifically considers emotions for their potential role as incentives. Can good mood, happiness, envy, frustration or disappointment, for example, affect individuals' performance and therefore boost individuals' effort? In which direction?

This essay aims to explore how emotions relate to agents' performance and, specifically, to address the motivational force of one emotion in particular: anxiety. My findings will be a contribution to contract and job design, and worker interaction policies in firms as well. The choice of anxiety (and anxiety trait) is guided by the knowledge that it is a strong and consistent determinant of individuals' performance and hence it has a predictive power on agents' behaviour. Indeed, the behavioural effects on task performance and individual differences in trait anxiety have been the core of quite substantial literature in psychology (Montague, 1953; Spence et al., 1956, 1957; Spence and Spence, 1966).

I will explain here briefly the concept of anxiety and the main theories in psychology on the relationship between anxiety and performance before addressing some related findings in the economic literature.

Anxiety: Definition and Theories in Psychology

According to the glossary of the Diagnostic and Statistical Manual of Mental Disorder (First et al., 1987) the term anxiety denotes "apprehension, tension, or uneasiness that stems from the anticipation of danger" (page 392). It is agreed, in fact, that there are two common principles of anxiety theory: the *anticipatory* and *aversive* feature of this emotion[6]. Some theorists see anxiety as a "basic emotion" that coordinates a quasi-autonomous process in the nervous system by communicating that a self-preservation goal is violated (Oatley and Johnson-Laird, 1987; Gray,

[6]See Lazarus (1966) for a survey on the experimental evidence of the anticipatory effects of anxiety.

1985b; Corr, 2008). In his work with humans and rats, LeDoux (1996) established the role of the amygdala[7] in the preconscious evaluation of a stimulus as representing a potential threat. In this sense, the aversive nature of anxiety is connected to the appropriate "flight" or escape response[8]. Anxiety can be also evaluated as a combination of different emotional states: for example, fear and expectancy (Plutchik, 1980) or fear and apprehension (Izard, 1977).

In this context, I will consider anxiety as a multidimensional construct (or structure where the structure of an emotion is a vector of component values; this vector is called *profile*) consisting of the following dimensions: negative valence, high unexpectedness, low control, high relation with social context, and high relation with self-esteem (Frijda, 1986). However, I shall also evaluate anxiety as "(1) an emotional state, evoked in a particular context and having a limited duration, and (2) as personality traits, characterizing individuals across time and situations" (Spielberger, 1972).

Anxiety and Performance in Psychology

Considerable research interest has focused on the relationship between anxiety and performance in psychology. It is apparent from these studies that the relationship is not a unique one, especially because anxiety interacts with other variables in the response-acquisition situation. A very important variable is, for example, task difficulty. Montague (1953) found that high-anxious subjects performed better than low-anxious subjects in the serial learning of a relatively easy list of nonsense syllables, but performed more poorly than low-anxious subjects on lists of increasing difficulty. Spence et al. (1956, 1957) have reported a series of experiments in which

[7]The amygdala is located within the medial temporal lobes of the brain and it has a primary role in the processing of memory and emotional reactions.

[8]Anxiety is an ambiguous construct usually identified with fear. However, fear is said to differ from anxiety primarily in having an identifiable eliciting stimulus. In this sense anxiety is often "prestimulus" that is anticipatory to more or less real threatening stimuli. In anxiety the nature and the location of the threat might remain unknown and thus difficult to cope whereas in fear the threat is clearly located in space and time.

the relative superiority of high-anxious or low-anxious people in paired-associate learning[9] was shown to be related to the task difficulty. Thus, the evidence in the lab seemed to Spence and colleagues to point out to some sort of non-monotonic relation between anxiety and performance[10], especially for subjects in the middle range of anxiety[11]: subjects in the middle range were found to be superior to those in the extremes in carrying out the task.

The relationship between emotional intensity and performance has been formalized in the so-called *Yerkes-Dodson* Law. According to this law the relation between emotional intensity and performance can be represented by an inverted U-curve. An increase in emotional intensity from some zero point upwards is supposed to increase the quality of performance, up to an optimal point. Further increases in intensity then lead to performance deterioration (Hebb, 1970). The optimal point is reached sooner (at lower intensity) the less well learned or more complex is the performance. An increase in emotional intensity affects finer skills, complex reasoning task, and recently acquired skills more readily than routine activity. However, the predictions of the Yerkes-Dodson Law do not always find support in the laboratory evidence: Spence and Spence (1966) found that the experimental manipulations of anxiety do not always induce poorer performance; highly anxious subjects indeed sometimes perform better than do less anxious ones.

Overall, it is assumed that anxiety refers to higher vigilance and motivates careful preparation, for example for examination and social interactions. Attention and motivation are anxiety components fully addressed in the *Processing Efficiency Theory* by Eysenck and Calvo (1992). According to their formulation the initial

[9] In the *paired associates* task subjects must learn a list consisting of words, like high-low and light-dark; then subjects are given the first word and they must respond with the second.

[10] This sort of relation has also been found by Montague (1953).

[11] Here anxiety is given by the measurement of trait (Manifest) using the Taylor Scale of Manifest Anxiety (Taylor, 1953).

effect of worries is negative because worries interfere with attention to the task-relevant information, thus reducing the resources potentially available for the task. Consequently, performance is impaired. However, anxiety has a motivational function that induces the subject to increase effort and therefore to boost performance:

> *Worry about task performance has a second effect [...]. It serves a motivational function...In order to escape from the state of apprehension associated with worrisome thoughts and to avoid the likely aversive consequences of poor performance, anxious subjects try to cope with threat and worry allocating additional resources (i.e. effort) and/or initiating processing activities (i.e. strategies)* (Eysenck and Calvo, 1992, p. 415).

An increase in anxiety can therefore improve performance if the agent is motivated to increase effort to such an extent that it counterbalances and exceeds the initial negative impact of anxiety[12].

The understanding of the effects of anxiety on individuals' behaviour and performance cannot disregard the study of anxiety trait. Research on anxiety has seen the development of theories of emotions that classified anxiety either as a primary (Gray, 1985a; Oatley and Johnson-Laird, 1987; Corr et al., 1997) or secondary emotion (Frijda, 1986; Izard, 1977; Ekman, 1992; Ortony and Turner, 1990) and theories of personality that included anxiety trait (or neuroticism) as one of the major dimensions of personality (See Digman, 1990, for a review). An attempt of unifying the research areas of anxiety within a common theoretical framework was made by Eysenck (1997) with his *Unified Theory of Anxiety*[13]. Eysenck pointed

[12] *Worry* or *self-preoccupation* is characterized by concerns over evaluation and failure, and expectations of aversive consequences. It is activated in stressful situations and is more likely to occur in individuals with high proneness to anxiety. Since in the Processing Efficiency Theory worry is the component of state anxiety precisely responsible for effects of anxiety on performance effectiveness and efficiency I will consider it in this work as an equivalent word to anxiety.

[13] Specifically, three research areas were unified in his theory: emotion, personality, and clinical or abnormal psychology.

out that there are fairly consistent individual differences in the cognitive biases operating within the emotional system.

However, in 1966 Spielberger had already proposed a conceptual framework for considering trait and state anxiety. In that framework, he argued that external stimuli are initially subject to a process of cognitive appraisal, which determines whether or not any particular stimulus is regarded as threatening. Only those stimuli that are appraised as threatening increase the level of state anxiety. Trait anxiety affects the process of cognitive appraisal.

> *[A]nxiety as a personality trait would seem to imply a behavioural disposition that predisposes an individual to perceive a wide range of objectively non-dangerous circumstances as threatening* (Spielberger, 1966, p. 17).

Since the mid-sixties, the trait-state distinction has received wide recognition in the psychological literature in explaining the tendency for individuals high in trait anxiety to experience more state anxiety than those low in trait anxiety (Spielberger et al., 1980; Dreger, 1985; Spielberger, 1985; Endler, 1997).

The measurement of anxiety trait is nowadays a well-established research procedure in psychological studies that aim to explore the connection between anxiety and behaviour. Results show that individuals' arousal, behaviour, and performance strictly depend on individual differences in anxiety trait (Eysenck, 1985; Sarason and Palola, 1960; Cassady and Johnson, 2002; Sarason et al., 1990; Hembree, 1990; Sarason, 1984, 1973).

The Neuropsychology of Anxiety

The relationship between emotions and performance has attracted growing attention among researchers due to the recent findings in neuroscience on the deviant behaviour of neurological patients with specific brain lesions, such as the frontal lobe (ventromedial prefontal lobe) damage (Damasio, 2008; Camille et al., 2004)[14]. With respect to the specific psychological state of anxiety, researchers have found that two particular brain structures are involved: the hippocampus and the amygdala (Gray, 1982; Gray and McNaughton, 2000). Felix-Ortiz et al. (2013) show that the basolateral amygdala (BLA) and the ventral hippocampus (vHPC) have both been implicated in mediating anxiety-related behaviour. In particular, their findings demonstrate that activation of BLA-vHPC synapses acutely and robustly increases anxiety-related behaviour while inhibition of BLA-vHPC synapses decreases anxiety-related behaviour. Tye et al. (2011) confirm that specific parts of the basolateral amygdala (BLA) are critical circuit elements for acute anxiety control.

However, Bannerman et al. (2004) suggest that the hippocampus plays a key role for the emotion of anxiety. In particular, the ventral hippocampus makes its own distinctive contribution to the control of behaviour in certain situations where anxiety occurs as distinct from fear, which more plausibly depends on the amygdalar functions[15]. The hippocampus and the amygdala are of course likely to be highly interactive. The idea of anxiety and fear as different emotional entities with separate underlying neural substrates is at the core of Gray's behavioural inhibition theory (Gray, 1982; Gray and McNaughton, 2000). This theory is based on extensive similarities between the behavioural effects of anxiolytic drugs and hippocampal lesions. According to the behavioural inhibition theory there is a sharp (functional, behavioural and pharmacological) distinction between fear and anxiety. Fear has the function of moving the subject *away from* danger and it is insensitive to anxiolytic drugs. Anxiety has the function of moving the subject *toward* danger. It involves inhibition of aggressive behaviour, increased risk assessment and defensive quiescence. All this manifestations of anxiety are sensitive to anxiolytic drugs.

[14]Damasio (2008) finds that patients who have damages in their frontal lobes become emotionally flat and lose their ability to make decisions, while retaining their cognitive power.

[15]Both hippocampus and amygdala are involved in anxiety-related processes and behaviour, as Felix-Ortiz et al. (2013) show.

The functions of the hippocampal system are distributed across the nominal psychological functions of anxiety and memory. It rests on the evidence that anxiolytic drugs affect "hippocampal" tests of memory (Paré, 2003; Huff and Rudy, 2004). The hippocampal memory system may provide the means for encoding both the spatial and the temporal contexts (the "where" and the "when") associated with a particular event (Eichenbaum and Fortin, 2003; Morris et al., 2003). There are several recent demonstrations of hippocampal involvement in the temporal sequencing of events (Agster et al., 2002; Fortin et al., 2002; Kesner et al., 2002). The challenge is to identify a common psychological operation that might underlie both the episodic-like memory function and a role in anxiety. Gray and McNaughton (2000) suggest that the operation performed by the hippocampus is to compare different potential response alternatives and then to select the optimal response. In the case of episodic-like memory tasks, this process will involve using conditional information provided by either spatial or temporal cues, or both, to select the appropriate learned response. In the case of tests of anxiety, the process may involve selecting between conflicting approaches and avoid responses.

Roozendaal et al.'s (2009) review shows that intensely emotional events or chronic exposure to stressful experience can create traumatic memories and even result in the development of mood and anxiety disorders. Acute and chronic stress induces long-term functional and morphological alterations in the amygdala, together with associated changes in the hippocampus and the prefrontal cortex, which might underlie the cognitive changes, increases in anxiety-like behaviour and mood alterations. These changes and alterations could have long-term consequences for cognitive performance and working memory.

Neurobehavioural and psychological evidence of the effects of emotions on decision-making processes has recently attracted the interest of economists (Zizzo, 2008, 2004b; Camerer et al., 2005). Ambler and Burne (1999) have studied the impact of affect on memory of TV advertising, using behavioural measures to assess performance in tasks of image recognition and recall. The results obtained suggest that under normal conditions, recognition and recall of affective TV material (using e.g. suspense, drama, humour) is superior to cognitive material (based on plain facts). Administration of propanolol (a blocker drug used to treat anxiety) reduces slightly recognition and recall of affective material, but increases substantially the recall of cognitive material.

Anxiety in Economics

Anxiety research in economics was pioneered by Loewenstein (1987) and Caplin and Leahy (2001). Loewenstein investigated some implications of anticipatory emotions for consumption decisions and in particular looked at anxiety effects on saving behaviour: anxiety (for the future) occurs when retirement is close or because of a possible loss of wage income; "this anxiety raises the returns, in terms of anxiety reduction, of saving, and counteracts the saving-discouraging effect of the loss of income upon retirement" (Loewenstein, 1987, p. 677). The emotional effect on behaviour can explain some anomalous saving phenomena like saving for short-term concrete goals instead of for a future retirement. Caplin and Leahy (2001) consider a general two-period decision problem under uncertainty. The novel element in their *Psychological Expected Utility Theory* is an exogenous map which assigns a psychological state due to anticipatory emotions like hopefulness, anxiety, and suspense, to the first period outcome and a lottery over the second period outcome. The agent's overall utility function is therefore given by the sum of the utility of the first period defined over psychological states and the utility of the second period which depends on the expectation with respect to the lottery. Although Caplin and Leahy present a powerful framework for modeling situations and emotions involving uncertainty, anxiety effects remain not very well addressed in their model. Rauh and Seccia (2006) reconsider the psychological expected utility theory and develop a formalization of anxiety consistent with expected utility maximization. In their two-period model anxiety is defined as the difference between expected utility with zero uncertainty and expected utility evaluated at optimal effort[16]. In this model anxiety is reduced when the information increases, and it is a function of an agent's effort, her/his skill and her/his emotional arousal. In particular, they show that anxiety can induce greater effort and expected performance as the Processing Efficiency Theory claims. A different interpretation and formalization of anxiety is given by Wu (1999) who refers to anxiety as the psychological utility induced by an unresolved gamble. He includes anxiety in his model of decision making with delayed resolution of uncertainty. Anxiety is thus defined as unproductive worrying, a source of psychological dissatisfaction. Instead of considering what alternatives can be undertaken to avoid an uncertain future, anxiety affects the evaluation of specific outcomes of gambles with delayed resolution (it measures how large the discount is for delayed resolution).

In the experimental literature few researchers investigate the choice problem under uncertainty in the presence of an anxious mood and provide explanations that take into account the anticipatory effects of anxiety. Bosman and van Winden (2010) conducted an experiment on a choice problem between a safe option and a risky option when there is a global risk, that is, when there is a probability of a certain risk that is not possible to avoid. Their results show that anxiety influenced decisions and created a tendency to take less risk. Hopfensitz and van Winden (2008) and van Winden et al. (2010) considered a setting of dynamic choice and the impact of the timing of the resolution of risk on people's willingness to take risks. They focused on anticipated and experienced emotions and found that anxiety definitely reduced risk taking.

[16]In fact the agent is uncertain about her/his skill in doing the activity or task at the first period, but she/he can make inference about it observing her/his own first period performance. Consequently, the first period effort not only increases expected first period performance but also the amount of information at the beginning of the second period about the skill.

In all this research anxiety has been studied in relation to risk attitude. As far as I know, there is only one study in the economic literature that explores anxiety in relation to task performance: Apesteguia and Palacios-Huerta (2010). The authors collected their evidence from a randomized natural experiment, that is, a real life situation in which treatments and control groups are determined via explicit randomization, on professional performance in a soccer competition. Using data from penalty kicks, they claim that teams that take the first kick in the sequence win the penalty shoot-out 60% of the time. Therefore, given the characteristics of the setting they attribute this difference in performance to psychological effects (pressure, anxiety) resulting from the consequences of the kicking order[17].

Apesteguia and Palacios-Huerta's setting involves the interaction between players and therefore it examines the behavioural anxiety effects when agents play against each other. In my context I aim to define the behavioural effect of anxiety on individual task performance. This eventually will allow me to consistently draw inferences in other strategic situations.

Literature Overview

In this chapter I am going to explore the relationship between anxiety and individual's performance. I will refer to the Processing Efficiency Theory by Eysenck and Calvo (1992), assuming therefore that anxiety has a motivational function that induces the subject to increase effort to such an extent that it counterbalances and exceeds the initial negative impact of anxiety on performance. From a psychological perspective, I will consider anxiety as a multidimensional construct consisting of various dimensions: negative valence, high unexpectedness, low control, high relation with social context, and high relation with self-esteem (Frijda, 1986).

[17] See the survey by Woodman and Hardy (2001) and Zaichkowsky and Baltzell (2001) for similar findings in the sports psychology literature.

I will also take "Trait Anxiety" and the trait-state distinction into consideration, building my behavioural hypotheses on the assumption that individuals high in trait anxiety show the tendency to experience more state anxiety than those low in trait anxiety (Spielberger et al., 1980).

Therefore, I am going to evaluate anxiety as "(1) an emotional state, evoked in a particular context and having a limited duration, and (2) as personality traits, characterizing individuals across time and situations" (Spielberger, 1972).

In order to study the causal relationship between anxiety and performance I will induce an anxious state into participants of two laboratory experiments by using a mood inducement technique. This technique consists of personal recollection of past events in which subjects were exposed to stressful experience. According to the neuropsychological literature (Roozendaal et al., 2009) this exposure can create traumatic memories through morphological alterations in the amygdala and can result in the development of an anxious mood.

From an economic perspective, I will not consider anxiety in relation to risk attitude as most economists have done so far. I will rather try to explore the effects of anxiety in relation to task performance. Apesteguia and Palacios-Huerta (2010) studied the competitive setting of football players and examined the behavioural anxiety effects when agents play against each other. In my study, I will instead focus on the behavioural effect of anxiety on individual task performance.

2.3 The Experiments: Behavioural Hypotheses, Design and Procedure

To find the relation and causal effect of anxiety on individual task performance I ran two laboratory economic experiments that, as I will explain later, are enriched by psychological techniques.

Behavioural Hypotheses

I expect that an anxious state can improve the performance of anxious agents by acting as incentives as predicted by the Processing Efficiency Theory (Eysenck and Calvo, 1992). The linear principal-agent model in economic theory predicts that a higher level of effort allows participants to achieve a higher levels of performance[18].

My first hypothesis can be formulated as follows:

HYPOTHESIS 1. *Anxiety positively affects performance.*

In order to study the effects of anxiety on performance I will induce an anxious state by using a mood inducement technique[19]. I expect that participants more prone to become anxious are more likely to become anxious if threatened by a negative stimulus. I will refer to these group of participants as the High Trait anxious subjects[20]. High Trait anxious individuals will experience a higher level of anxiety and will perform better. Hence, my second behavioural hypothesis follows:

HYPOTHESIS 2. *High Trait anxious subjects perform better when their anxiety arousal increases.*

The Experiments

---

[18]Agents' performance or output is given by q = $\mu + \epsilon$ where μ is the level of investment, a measure of skill or average output, chosen by the agent prior to a realization of the random or luck component ϵ (Lazear and Rosen, 1981).

[19]Mood inducement techniques are widely used techniques in psychological laboratory experiments that aim to manipulate subjects' mood (Martin, 1990).

[20]"The stronger the anxiety trait, the more probable that the individual will experience more intense elevations in State Anxiety in a threatening situation." (Spielberger et al., 1970, p. 1). State Anxiety indicates the individual's current level of anxiety.

The experiments, Experiment 1 and Experiment 2, were conducted in February and March 2011 at the University of East Anglia. Instructions, control questionnaire and experimental tasks were computerized[21]. The experiments were programmed and conducted with the software z-Tree (Fischbacher, 2007). 151 subjects (74 men and 77 women) aged between 18 and 41 took part in the 26 sessions. All of them were university students, enrolled in different schools (only 10% were Economics students); 45% were British while the rest were represented by a wide variety of international students. Participants received a show-up fee of £3 and a payment based on their performance[22]. Subjects were randomly assigned to the seats in the lab where partitions between them assured the anonymity and the avoidance of any communication. Instructions were read aloud and the experimenter checked the participants' answers to the control questionnaire individually.

Procedure

At the beginning of each session subjects were asked to fill in the State Trait Anxiety Inventory (STAI) questionnaire by C. Spielberger (1976). This questionnaire aims to assess the current anxious state (*State Anxiety*) and the proneness to anxiety (*Trait Anxiety*) of subjects. It consists of 20 questions on State Anxiety and 20 questions on Trait Anxiety and it is framed in order to avoid any kind of demand effect. The instructions have a neutral frame as well. Participants used a four-point Likert scale for their answers. The STAI questionnaire, hence, assigns two scores, one for State Anxiety and another for Trait Anxiety; for each subject there are therefore two measurements from 0 to 80[23].

[21]Participants received paper instructions as well.

[22]The average payment per subject was 12 pounds.

[23]To the four-point Likert scale answers scheme another possible answer was added: "I did not understand the question". Subjects that were not sure about the meaning of the question could select it. This type of answer was considered as not answering the question in the score counting. Scores were adjusted according to the number of not answered questions. If there were 3 or more questions not answered for each set of questions (Trait or State) the score for that set was considered missing (Spielberger, 1972). There were only 3 participants whose score, either State or

After completing the STAI questionnaire, participants received instructions for the second part of the experiment, which consisted of two sections. In the first section a mood manipulation technique was applied and in second section participants were asked to perform a real effort task. Instructions for the mood manipulation task were very general and framed neutrally: participants were asked to recall two personal situations over the last year in which they felt very anxious. Then they were given the possibility to write about them briefly. Finally, they were induced to think about some adjectives that could better describe their feelings in those situations. Overall, the task lasted for eighteen minutes[24]. This mood inducement technique is called *Personal Recollection* and its validity and effectiveness in inducing an anxious mood is established and proved by several psychological experiments (Martin, 1990)[25].

In the last section of the experiment participants carried out a *real effort* task[26], the *slider task* by Gill and Prowse (2009)[27]. Forty-eight sliders, with a range of integer values from 0 to 100, appeared on the screen and subjects were asked to drag each slider up to the value of 50 with the help of their mouse[28]. They had two minutes to correctly position as many sliders as they could. For each value correctly positioned they gained one point. Subjects repeated the task for ten rounds; the

Trait, was considered missing. The STAI questionnaire is in Appendix A.1.

[24]Experimental instructions are in Appendix A.2.

[25]In addition to the personal recollection technique other mood types of mood manipulation and procedures were examined in pilot experiments like music, cognitive tasks and real effort tasks (Martin, 1990).

[26]Real effort tasks are widely used in many economic laboratory experiments. They are practical tools that give a measure of the level of effort exerted by participants. Quite a few tasks have been used in economic experiments. For example, mathematical questions (Vesterlund and Nierderle, 2007; Eriksson et al., 2008; Dohmen and Falk, 2011), folding letters (Konow, 2000; Falk and Ichino, 2006; Konow et al., 2009), counting numbers in a series of grids (Pokorny, 2008; Abeler et al., 2011; Falk et al., 2006), solving some anagrams and various segretarial tasks (Gneezy, 2002; Gneezy et al., 2003). A more extensive literature review on the literature of real effort tasks is presented in Chapter 3 - "Does it matter which effort task you use?".

[27]Guidance and practical advice of the use of this task have been followed.

[28]See Figure A.1 in Appendix A.3 for a screen shot of the task.

score was reset at the end of each round[29]. At the end of the experiment one round was randomly selected for the payment[30].

Experimental Treatments

I ran two experiments, Experiment 1 and Experiment 2. They were identical in the tasks and material used but slightly different in their procedures: Experiment 1 consisted of two sessions separated over time while Experiment 2 consisted of only one session. Participants in Experiment 1 were invited twice by email; the second session took place one week later[31]. Running two separate sessions in Experiment 1 allow me to have an additional control on any experimenter demand effect (Zizzo, 2010) or confound effect due to the implementation of the STAI questionnaire[32]. Each experiment consisted of two treatments: in the main treatments a mood manipulation technique for anxiety was applied while in the control treatments a non-anxious mood was induced[33]. Table 2.1 contains details of the treatments and the number of subjects in each experiment.

Emotion Elicitation

Before and after completing the real effort task, participants rated the extent

[29]In this work I chose to use the slider task because it allows to minimize the physical skills required to carry out the task and to get an unambiguous measure of the subjects' effort exerted in the time constraint. This task provides in fact only the measure of the number of sliders correctly positioned in the time constraint. I do not have any measure of how long subjects take (or how hard was) to position each slider. Participants were incentivised to try to correctly position as quickly as possible each slider.

[30]Each point was paid £0.30.

[31]The first session of Experiment 1 lasted around 15 minutes and the second session of Experiment 2 lasted around one hour. Sessions of Experiment 2 lasted around one hour and 15 minutes.

[32]In the first session of Experiment 1 subjects filled in the STAI questionnaire.

[33]In the control treatments I used the same mood inducement technique, personal recollection, as in the main treatments. However, in these sessions participants were asked to remember pleasant events.

Table 2.1: Experimental Treatments and Number of Subjects

	Number of Sessions	Treatment	Subject	Mood
Experiment 1	2 sessions	A	40	Anxious
	2 sessions	B	37	Non-Anxious
Experiment 2	1 session	C	37	Anxious
	1 session	D	37	Non-Anxious
		Total	**151**	

to which they felt each of the eight emotions presented to them between 0 (not at all) and 4 (very much) on a Likert scale . The type of emotions were drawn from Slyker and McNally (1991) and were anxiety, tiredness, despondency, frustration, apprehensiveness, tiredness, happiness, and anger[34].

In my analysis I will consider the Trait Anxiety score and the arousal anxiety values before the real effort task as my main variables of anxiety while the State Anxiety score and the arousal anxiety values after the real effort task will be used as control for subjects' current emotional state.

2.4 Results

Performance and Anxiety

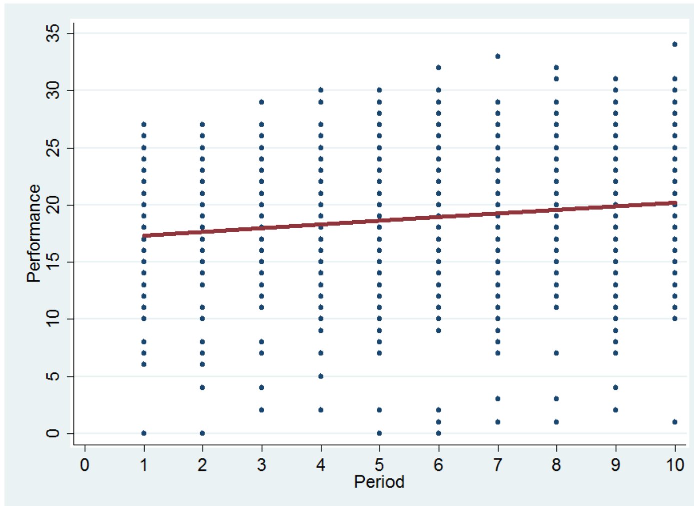
Figure 2.1 shows that performance increases over the ten periods indicating learning effects[35].

On average performance is greater in the main treatments of Experiment 1 and

[34]Slyker and McNally (1991) tested the relative effectiveness of some mood inducement procedures for anxious and depressed moods.

[35]I reproduced the performance trend of the real effort task by Gill and Prowse (2011).

Figure 2.1: Performance Trend



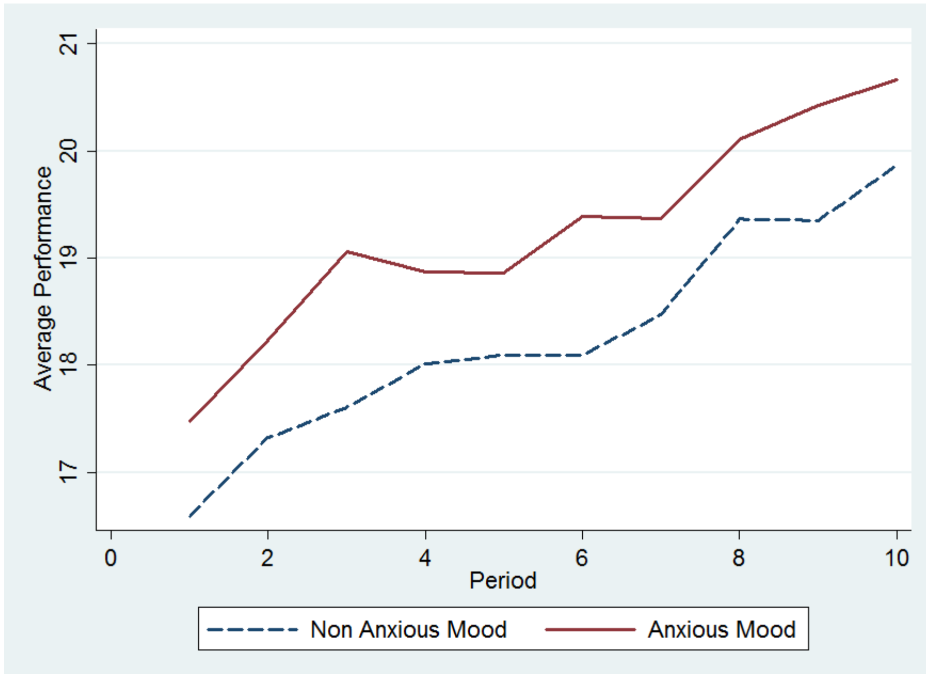
Experiment 2 where an anxious mood is induced (Treatment A and C) compared to the control treatments where a non-anxious mood is induced (Treatment B and D). The difference in treatments is confirmed by the t-test (p-value ≤ 0.001) and the non-parametric test Wilcoxon rank-sum (p-value $= 0.005$)[36] performance is significantly higher when participants' mood is negatively manipulated (Figure 2.2)[37].

While there is no significant difference in the level of performance between treatment A (Experiment 1) and treatment C (Experiment 2) when an anxious mood is induced, performance is significantly higher in treatment D compared to treatment B, treatments in which a non-anxious mood is induced (t-test, p-value $= 0.020$, Wilcoxon rank sum test, p-value $= 0.025$). The procedure of the two separated sessions over time seems to not affect the level of exerted effort when participants have an anxious mood. However, filling in the STAI questionnaire just before the real effort task seems to induce subjects to work harder when their mood

[36]The t-tests and non-parametric tests used now and onwards are two-sided unless specified otherwise.

[37]For the t-test and Wilcoxon rank-sum test I merged treatment A with C and treatment B with D. This is because I aim to compare performance by mood.

Figure 2.2: Performance by Mood



is positively manipulated.

Looking at Table 2.2, anxiety is higher in the main treatments (treatments A and C): average elicited anxiety is statistically significant (t-test, p-value = 0.006, Wilcoxon rank sum test, p-value = 0.005) compared to the average elicited anxiety in the treatments (Treatment B and D) where subjects were asked to recall pleasant events (Wilcoxon rank-sum test, p-value = 0.033). In particular, in Treatment C anxiety has the highest mean with the lowest standard deviation (Table 2.2).

Negative emotions like despondency, sadness, apprehension, tiredness, and frustration are on average higher in Treatment C compared to the other treatments (Table 2.2) while, at the aggregate level, tiredness, sadness, apprehension, anger and frustration are on average higher in treatments where an anxious mood is induced and despondency and happiness are on average higher in the treatments in which subjects are meant to be relaxed. Tiredness differs significantly between the main treatments and control treatments at a 10% level (Wilcoxon rank-sum test, p-value

Table 2.2: Average values of Performance and Emotions

	Experiment 1		Experiment 2	
	A	B	C	D
Performance	19.04 (4.24)	17.80 (5.18)	19.46 (4.56)	18.75 (5.95)
Anxiety	1.67 (1.45)	1.05 (1.31)	1.78 (1.25)	1.21 (1.25)
Happiness	2.02 (1.36)	2.46 (1.46)	2.13 (1.15)	2.62 (1.06)
Tiredness	1.75 (1.59)	1.35 (1.60)	2.48 (1.32)	2.01 (1.41)
Despondency	0.47 (0.78)	0.84 (1.21)	1.05 (0.84)	0.86 (0.97)
Sadness	0.40 (0.84)	0.62 (1.20)	0.83 (0.98)	0.27 (0.50)
Apprehension	1.05 (1.33)	1.22 (1.52)	1.73 (1.19)	1.13 (1.10)
Anger	0.45 (1.10)	1.62 (0.44)	0.35 (0.67)	0.32 (0.70)
Frustration	0.65 (1.09)	0.63 (1.09)	1.05 (1.15)	0.59 (0.89)

Note: Standard Deviation in parentheses

= 0.081) while happiness happens to be significantly higher in the control treatments (Wilcoxon rank-sum test, p-value = 0.027) (Table 2.3).

The results above support the effectiveness of the mood inducement technique used both for the main treatments and for the control ones: subjects are on average more anxious when an anxious mood is induced and happier when a non-anxious mood is induced. Moreover, it has been shown that performance is higher in the main treatments where an anxious mood is induced (Treatment A and C) compared to the control treatments where a non-anxious mood is induced.

Regressions on Performance

There is strong evidence that anxiety positively affects agents' performance as shown by the regression analysis on the performance level. I employed an Ordinary Least Square (OLS) estimation model with error clustering on all my aggregate

Table 2.3: Aggregate average values of Performance and Emotions

	Main Treatments	**Control Treatments**
Performance	17.48 (4.41)	16.59 (5.09)
Anxiety	1.72 (1.35)	1.13 (1.27)
Happiness	2.07 (1.26)	2.54 (1.27)
Tiredness	2.10 (1.50)	1.67 (1.53)
Despondency	0.75 (0.86)	0.85 (1.09)
Sadness	0.61 (0.93)	0.44 (0.93)
Apprehension	1.37 (1.30)	1.17 (1.32)
Anger	0.40 (0.92)	0.24 (0.59)
Frustration	0.85 (1.13)	0.60 (0.99)

Note: Standard Deviation in parentheses

data[38]. Table 2.4 presents the estimates of an OLS regression model with error clustering.

Anxiety has a strong significant positive effect on performance ($\beta = 0.873$): an increase in anxiety level induces subjects to work harder.

RESULT 1. *As predicted by hypothesis 1, there is strong evidence that anxiety positively affects performance.*

Subjects that experience higher levels of anxiety arousal exert more effort; this allows them to perform better the individual task. Result 1 is consistent with

[38]I also used random effects regression models and I found that random effects regression estimates converge to OLS regression estimates. For my data the error clustering model is the most appropriate considering the nature of my independent variables (time-invariant variables). The use of clusters provides a robust standard error per group (that is, per subject). Sashegyi et al. (2000) argues that where observations over time are taken for different group of subjects an econometric model must control both for intra-cluster correlation and intra-individual correlation within the same cluster. In either way, I control for the subject level non independence of observations.

Table 2.4: OLS Estimates with Error Clustering on Performance

Period	0.322***	(7.94)
Main	0.849	(1.39)
Number of Sessions	-0.852	(-1.25)
abs(Anx)	0.014	(0.01)
abs(Anx)Square	-0.059	(-0.08)
Anxiety	0.873***	(2.66)
Despondency	0.957*	(1.76)
Sadness	0.027	(0.07)
Apprehension	-0.871**	(-2.58)
Tiredness	-0.284	(-1.12)
Happiness	-0.032	(-0.11)
Anger	0.175	(0.56)
Frustration	-0.724**	(-2.19)
Age	-2.341***	(-4.74)
Gender	2.380***	(3.73)
British Nationality	1.794***	(2.81)
School of Economics	0.327	(0.27)
State Score	-0.016	(-0.37)
Trait Score	-0.005	(-0.12)
Constant	16.702***	(6.55)
Observations	1470	
R-Square	0.264	
F-test	6.300	

t statistics in parentheses;

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$;

Model 1: Error clustering with demographic variables.

the Processing Efficiency Theory: anxious subjects wish to avoid the negative consequences of a negative performance and hence they put more effort in the task. Because of its motivational role and its effects on effort and performance, anxiety might represent a type of incentive.

Results also show that overall performance increases over time (variable "Period"[39]) indicating the trend of learning effects ($\beta = 0.322$). In addition to Anxiety and Period, the regression model considers the following set of independent variables as well: treatment dummies (*Main* = 1 if an anxious mood was induced; *Number of Sessions* = 1 if a treatment consisted of two separated sessions over time), eight *emotion elicitation variables* that indicate the emotional state on a 5-point Likert scale before the real effort task, *Trait Anxiety* and *State Anxiety* scores, and two variables that detect the shape of the relationship between anxiety and performance (*abs(Anx)* is the absolute value of the difference between the anxiety values and the anxiety average value, and *abs(Anx)Square* is the square of the difference between the anxiety values and the anxiety average value). Model 1 includes some demographic variables as well (Age, School of Economics = 1 if Economics, British Nationality = 1 if British, Gender = 1 if male).

Participants in treatments where an anxious mood was induced perform better than participants in treatments where the mood manipulation aims at relaxing (Main, $\beta = 0.849$). The coefficient is not significant and this indicates that in the main treatments subjects have different sensitivities to the anxious stimulus. I find hence in the main treatments on average a higher level of anxiety and yet some groups of participants that perform better than other when an anxious mood is induced.

The coefficient of Number of Sessions is not significant as well: running a unique session for my experimental tasks does not provide any significant bias in my results.

[39]There were ten periods.

It appears that as long as anxiety arousal increases towards the mean performance increases (abs(Anx), β = 0.014) but only in a small size since the magnitude of the coefficient is decreased but the coefficient of abs(Anx)Square variable. However, none of the variables have any significant effects on performance and hence I can claim that performance in these experiments does not have a bell-shaped relationship with anxiety arousal. On the contrary, it seems that the more anxiety increases the higher the performance.

Either an apprehensive or a frustrated status negatively affects subjects' performance (Apprehension, β = -0.871; Frustration, β = -0.724) while being despondent might produce some positive effects (Despondency, β = 0.957). There is no other robust result linked to the elicitation of emotion or anxiety characteristics[40].

Moreover, there is mostly no correlation between the emotions elicited before completing the real effort task. Frustration seems to be correlated with despondency (Pearson correlation coefficient = 0.554, p-value $\leq$ 0.001, Spearman correlation coefficient = 0.525, p-value $\leq$ 0.001). Therefore, for the robustness check I dropped either the despondency or frustration variable in additional regression analysis and I found that anxiety is still significant. See Table A.3 in Appendix A.2[41].

Regressions by Trait Anxiety Groups

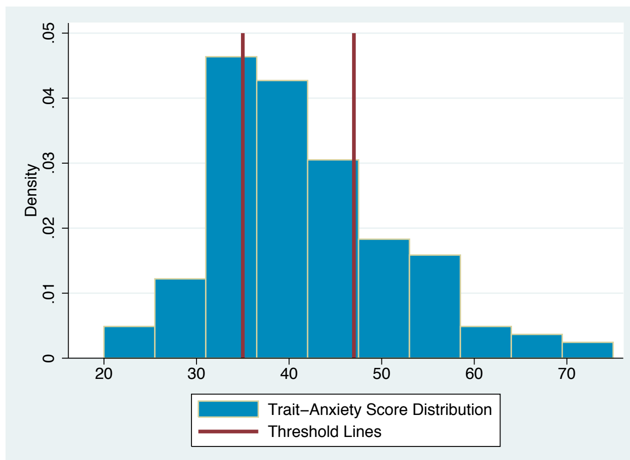
In order to more accurately identify the relation between Trait Anxiety and subjects' performance, and specifically to test hypothesis 2, the subject pool has been divided in three groups (Low Trait Anxiety, Medium Trait Anxiety, and High

[40]In additional regression analysis I introduced some interaction terms between anxiety and other variables like period, despondency, apprehension and gender. I found some significant effect between anxiety and despondency. See Table A.2 in Appendix A.3.

[41]I also found evidence of some demographic effects. See Table 2.4.

Trait Anxiety) according to subjects' Trait Anxiety score. The thresholds used to create the three trait anxiety groups are the trait score upper quartile and the trait score bottom quartile values of the Trait Anxiety score distribution, specifically 35 and 47.368 (Figure 2.3)[42].

Figure 2.3: Trait Anxiety Thresholds



The Trait Anxiety score distribution is slightly skewed to the right (mean = 41.50, median = 39, St. Dev = 10.09, skewness = 0.73) and there is a consistent group of participants whose score is within 30 and 40. However, I have overall a nearly normal distribution of the Trait Anxiety score which allowed me to create the three different groups of Trait-Anxiety subjects and finally to identify different behavioural effects[43].

Hence I ran three sub-sample OLS regressions with error clustering. Table 2.5 shows the estimation results.

Low Trait anxious (Low TA column) subjects were not affected by the mood inducement technique and indeed their level of anxiety does not have a significant impact on performance. Moreover, I noticed that Low Trait anxious participants

[42]Different groups have been considered in order to check the effect of a specific threshold values. For example I simply divided my observations in three groups with the same number of observations; the results obtained do no significantly differ from those reported here.

[43]See Figure A.2 for the Trait Anxiety scores distribution in Appendix A.2.

Table 2.5: OLS Estimates with Error Clustering on Performance with Trait Anxiety Groups

	High TA		Medium TA		Low TA	
Period	0.341***	(3.38)	0.282***	(6.18)	0.395***	(4.01)
Main	1.419	(0.86)	1.603**	(2.23)	-5.797***	(-3.68)
Number of Sessions	-1.549	(-1.52)	-1.933**	(-2.37)	-0.780	(-0.67)
abs(Anx)	-1.044	(-0.20)	-2.853	(-1.19)	2.187	(0.65)
abs(Anx)Square	-0.269	(-0.16)	1.164	(1.31)	-0.314	(-0.20)
Anxiety	1.637**	(2.38)	-0.203	(-0.44)	0.494	(1.46)
Despondency	0.738	(0.87)	0.768	(1.41)	-4.428***	(-3.38)
Sadness	-0.868	(-1.12)	0.833	(1.59)	-1.080	(-1.32)
Apprehension	-0.264	(-0.40)	-0.908**	(-2.40)	-0.385	(-0.56)
Tiredness	-0.819*	(-1.83)	-0.381	(-1.57)	-0.211	(-0.59)
Happiness	0.121	(0.16)	-0.519*	(-1.88)	-0.422	(-0.69)
Anger	-0.367	(-0.45)	0.737**	(2.00)	-3.065***	(-2.95)
Frustration	-0.178	(-0.36)	-1.196**	(-2.59)	1.382	(1.68)
Age	-3.808**	(-2.20)	-1.676***	(-3.40)	-3.806***	(-4.96)
Gender	1.612	(1.18)	1.462*	(1.84)	6.263***	(8.48)
British Nationality	2.490*	(1.92)	1.455	(1.63)	1.262	(1.42)
School of Economics	4.308	(1.41)	-3.179**	(-2.34)	4.617***	(3.64)
State Score	-0.072	(-0.76)	-0.012	(-0.21)	0.247***	(3.09)
Trait Score	0.280***	(3.31)	-0.051	(-0.45)	0.568***	(3.90)
Constant	4.056	(0.67)	23.581***	(4.94)	-5.007	(-1.08)
Observations	380		770		320	
R-Square	0.414		0.362		0.657	
F-test	5.040		6.916		26.973	

t statistics in parentheses;

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

give up when they are despondent (Despondency, β = -4.428). By contrast, for High Trait anxious individuals (High TA column) anxiety is a determining factor for performance (Anxiety, β = 1.637): an increase in anxiety allows subjects to achieve better results. The effect of anxiety on individuals' performance is related to the proneness to anxiety.

For the Medium Trait Anxiety (Medium TA column) group a significant impact on performance is given basically by apprehension, anger and frustration: the negative coefficients of apprehension and frustration indicate that Medium Trait anxious subjects work less hard as soon as they feel more apprehensive and frustrated (Apprehension, β = -0.908; Frustration, β = -1.196) but a feeling of anger seems to push Medium Trait anxious participants to perform better[44].

RESULT 2. *As predicted by hypothesis 2, High Trait anxious subjects perform better when their anxiety arousal increases.*

To summarize, anxiety significantly increases High Trait anxious subjects' performance while it does not affect Low Trait and Medium Trait anxious subjects' behaviour. High Trait anxious individuals are more likely to respond to threatening stimuli and therefore to become anxious (See Table A.5 in Appendix A). This is confirmed by the coefficient of the variable "Trait Score" for the High Trait anxious subjects (β = 0.280) which is strongly significant (p-value = 0.001).

Frustration and Apprehension decrease Medium Trait anxious subjects' performance while despondency has negative effects on Low Trait anxious subjects' performance[45].

[44] For robustness check I dropped either despondency or frustration variable in additional regression analysis and I found that anxiety is significant and robust for High Trait Anxiety Group and positively affects individuals' performance.

[45] Performance always increases over time and Medium Trait and Low Trait anxious subjects exert more effort when an anxious mood is induced. The number of sessions have a negative effect on performance for Medium Trait anxious subjects, thus filling in the STAI questionnaire in the same lab session of the real effort task does change my results for the Medium Trait Anxiety group.

2.5 Discussion

My results show that anxiety is positively related to performance: an increase in anxiety induces an increase in performance. Hence I can assert that anxiety has a motivational role and a positive effect on agents' performance. Moreover, I found that the effects of anxiety on performance are related to propensity to anxiety. High Trait anxious subjects are more sensitive to the mood manipulation and therefore their emotional state allows them to achieve better results.

My findings support the Processing Efficiency Theory that states the motivational force of anxiety and the Yerkes-Dodson law that links performance to emotional arousal and motivation: High Trait anxious individuals are threatened by some stimuli and their emotional arousal increases. Therefore, in order to avoid negative emotions associated with bad performance they exert more effort and hence they are able to perform better.

In this experiment I measure subjects' performance as the number of sliders correctly positioned in the time constraint. This output is one of the proxies of the effort variable (or input) which can be used. The latter, in fact, is not observable and measurable. Individuals' performance or output is given by q = $\mu + \epsilon$ where μ is the level of investment or effort, a measure of skill or average output, chosen by the agent prior to a realization of the random or luck component ϵ (Lazear and Rosen, 1981). In this experiment, anxiety boosts (at least some type of) individuals

abs(Anx) and abs(Anx)Square variables do not identify any pattern in the relationship between the level of anxiety and performance for any Trait Anxiety group. Finally, estimates on demographic variables indicate that students enrolled in the School of Economics perform better if they are Low Trait anxious subjects, worse if they are Medium Trait anxious subjects. Male students show a better performance than female students if they are Low Trait anxious subjects. Young students perform always better, especially if they are High Trait anxious subjects. Table A.4 in Appendix A shows the estimates of OLS regressions with error clustering by Trait Anxiety groups without demographic variables.

to exert more effort acting as an incentive.

It is worth, at this stage, to reassert how I defined anxiety: a general state of tension and apprehension triggered, in this case, by a mood inducement technique. Roughly speaking, anxiety is a complex state characterized by a subjective feeling of apprehension and heightened physiological reactivity. However, there is a clear distinction between the construct of anxiety and apprehension or tension. Apprehension is a concept closely allied to stress; it refers to a state of the organism created by stress. Broadly, apprehension (or tension) "is conceptualized as a state of disequilibrium brought about by some psychological need, leading to behaviour that tends to satisfy the need and thereby restore equilibrium" (Levitt, 1967, p. 13). This implies consciousness of the occurring emotional state whereas anxiety is not experienced consciously because of the temporary effects of defence mechanisms. The subsequent feeling is popularly known as "apprehension" or "nervous tension".

Let us consider the distinction between the emotional state of anxiety and an anxious mood. Generally speaking, moods are usually distinguished from emotions by one of three criteria: longer duration, lower intensity, and diffuseness or globality (e.g. Isen, 1984; Morris and Schnurr, 1989). Of the three criteria, "diffuseness" is the most interesting. It has been adopted, in fact, as the main criterion for mood by a number of authors (e.g. Isen, 1984; Ruckmick, 1936)[46]. "Diffuseness" indeed defines a class of affective states that is a distinct set of emotions. Emotions have an object. They are "about" something. One is happy about something, angry at someone, afraid of something, etc. Emotions are therefore "intentional" phenomena; they involve a subject-object relationship[47]. Moods lack such an object. The diffuseness of affective states, that allows classifying them as moods, can thus be characterized by the absence of orientation upon an object[48]. Moods are, therefore, affective states

[46]Ruckmick states: "[Mood] has no particular cognitive element. [..] There is also generally no cognitive impulse about it. It does not lend itself to any definite action" (1936, p.72).

[47]Frijda points out that emotional behaviour is directed toward or away from, or at least oriented upon, a particular thing (Frijda, 1993).

[48]There exists in fact "joyful" behaviour, not focused upon one particular object, and likewise

without an object or without a specific object. The latter qualification is added because some affective states have the environment as a whole as their object. In certain anxiety states, for example, the environment is felt to be an unsafe place, not offering any possibility for control.

The distinction between emotion and mood can of course be blurred. For example, the object of an affective state is not the same thing as its cause. Moods, while not having an object, may originate in a specific event involving a specific object. A mood may also be the consequence of a particular emotion. An emotion turns into a mood, or gives rise to a mood, when the focus upon the object is lost and the feeling become diffuse, having no object. The individual may know what caused his or her mood, in an unfocused state without an object.

I do not claim in this study any strict difference between the emotion of anxiety and the anxious mood induced by means of the mood inducement technique. The type of technique used, personal recollection, implies the recollection of particular events, maybe involving a specific object, in which the emotion of anxiety was experienced. This emotion may give rise to an anxious mood induced by the mood manipulation. In this study I am interested in how anxiety affects the performance in an individual task. I do not investigate specifically how object-specific anxiety affects the individual performance. The self-report, submitted by participants at the end of the mood manipulation, gives a measure of the emotional state post-manipulation[49]. The STAI questionnaire, completed at the beginning of the experiment, and hence before the mood manipulation, specifically aims to assess both the current anxious state (State Anxiety) and the proneness to anxiety (Trait Anxiety) of participants.

an "angry mood" applied to sequences of irritated or angry responses to a variety of events.

[49]More precisely, the post-manipulation self-report gives a measure of the subject's emotional state after the mood manipulation and prior the remaining laboratory tasks. Subjects may be anxious because of the uncertainty about the tasks they will be asked to do. They in fact receive specific instructions about the remaining laboratory tasks after the measurement of their emotional state.

In the mood manipulation phase subjects were asked to think about two personal situations over the last year that provoked a feeling of anxiety[50]. Thinking about two (and not more[51]) events induces the emotional state that occurred, and it is more likely that High Trait anxious subjects were more sensitive to this type of recollection. As Spielberger pointed out (1976; 1980) individuals high and low in trait anxiety differ in their susceptibility to psychological stress. Consequently, when High Trait anxious subjects were asked to perform a task, the negative valence of anxiety along with the worried thoughts initially distracted their attention from the task. Afterwards, they adopted an approaching behaviour towards the mildly threatening stimulus applied: as soon as subjects started working on the real effort task their attention was shifted from the threatening stimulus. This allowed subjects to cope with worrying thoughts, to be aware that their state of tension might compromise their performance and finally their payment, and therefore to work harder. This coping strategy aims to avoid negative consequences and disappointment of possible poor result at the task due to the emotional state. As Result 2 shows, High Trait anxious agents' performance increased monotonically when their anxiety arousal increases.

I would expect, according to the Yerkes-Dodson law, that very high anxiety arousal could have negative effects on performance. However, in my experiment the level of arousal induced is not extremely high. The type of mood inducement technique applied, Personal Recollection, is effective but does not produce an extreme level of emotional arousal. In some psychological experiments *deception* more effectively induces high arousals of anxiety (Hertwig and Andreas, 2008; Kelman, 1966; Dabbs and Helmreich, 1972). But in an economic laboratory "personal recollection" seemed to be a more appropriate manipulation and my pilot experiments confirmed its effectiveness. Though, only a moderate level of anxiety was on average elicited from participants. For this reason the shape of the relationship between

[50] I specifically used the adjective "anxious". However, anxiety and apprehension were both in the list of emotions at the phase of measurement of emotional arousal.

[51] The recollection of quite a few negative events is less likely to provoke a higher arousal.

anxiety and performance found in this experiment is not bell-shaped as stated by the Yerkes-Dodson law. Rather I found a positive relationship that indicates that my results highlight the positively increasing left-hand side of the bell-shaped curve of the relationship between performance and emotional arousal.

It might be claimed that an increase in anxiety arousal (provoked by the mood manipulation) might reflect a decrease in self-confidence. That is to say that somehow inducing under-confidence (by using the mood manipulation technique) can produce the same effects on performance attributed by the literature to an increase in anxiety arousal. The relationship between anxiety and self-confidence is formalized in the multidimensional anxiety theory which assumes a series of two-dimensional relationships between cognitive anxiety, self-confidence and performance (Martens and Smith, 1990). Specifically, cognitive anxiety is defined as "negative expectations and cognitive concerns about oneself, the situation at hand, and potential consequences" (Morris and Hutchings, 1981, p. 541), and self-confidence is conceptualized as one's belief in meeting the challenge of the task to be performed. As Martens et al. (1990) describe, cognitive anxiety and self-confidence are at opposite ends of a continuum. In their meta-analysis on sport performance Woodman and Hardy (2003) confirm the (positive) relationship between self-confidence and performance and the (negative) relationship between cognitive anxiety and performance hypothesized by the multidimensional anxiety theory. However, they do not find any consistent evidence on the relationship between cognitive anxiety and self-confidence. Similar patterns are identified by Koivula et al. (2002) and Parfitt and Pates (1999) in their studies on sport performance.

Therefore, it is not straightforward to assess that an increase in anxiety arousal may produce a decrease in self-confidence and hence the same effects on individuals' performance. In my study I did not collect any data on self-confidence. It would have been interesting to introduce in the experiments the commonly used Competitive State Anxiety Inventory-2 (Martens et al., 1990) to get a measure of cognitive anxiety and self-confidence. This is left for future research.

The effects of anxiety on subjects' performance strictly depend on the proneness to anxiety. Low and Medium Trait anxious participants have not been dramatically affected by the mood inducement technique and therefore their level of anxiety does not have a significant impact on the exerted effort. This explains why the coefficient of the variable "Main", which refers to treatment A and C together, treatments where an anxious mood was induced, is not significant in the OLS regression shown in Table 2.4. High, Low and Medium Trait Anxiety groups were present both in treatment A and C (and in treatments were a non-anxious mood was induced, B and D) but only High Trait anxious participants were affected by the threatening stimulus of the mood manipulation. Hence, only a group of participants experienced an increase of their anxiety level. It is in fact the variable "Anxiety" that affects subjects' performance, as it is shown by its coefficient in the same regression[52]. I cannot exclude that some participants in treatments B and D, where a non-anxious mood was induced, declared a high level of anxiety at the beginning of the real effort task. Their emotional state might be due to the uncertainty about the task and their ability to perform it well. Table 2.3 shows for example that the mean of anxiety is high in treatment D (treatment consisting of only one session where I induced a non-anxious mood) where the mean of effort is significantly higher compared to treatment B (treatment consisting of two sessions where I induced a non-anxious mood).

In this work I applied three measures of anxiety. Specifically, I obtained a State and Trait measure using the STAI questionnaire (Spielberger, 1972) and a 5-point Likert scale for the anxiety elicitation before and after the real-effort task[53]. These

[52]The coefficient of the variable "Apprehension" has a very similar magnitude of the anxiety variable but with opposite sign in the aggregate regression in Table 2.4. The separate regressions by Trait Anxiety groups show the different impact of these two variable on performance (Table 2.5).

[53]At the end of the real effort task I asked subjects to elicit the intensity of their emotions on a 5-point Likert scale. These groups of emotions cannot be meant to have predictive power since they refer to an emotional state elicited at the end of the experiment. Table A.1 in Appendix A.2 shows

measurements are of a self-report type. Trait Anxiety and State Anxiety scores are correlated (Spearman coefficient = 0.5439, Pearson coefficient = 0.611)[54]. The State Anxiety scores distribution is positively skewed (mean = 34.38, median = 32.31, St. Dev. = 9.37, skewness = 1.11) and the 75% of subjects have a score below 40. This means that overall participants arrived in the lab relatively relaxed and that no confound effects due to subjects' negative emotional state occurring before the experimental sessions was detected (See Figure A.3 for the State Anxiety scores distribution in Appendix A).

Although I used standard measurements of anxiety there are other techniques used in the experimental psychological laboratory that could offer a more accurate measure of anxiety arousal and trait. For example, a widely used measurement is the *dot-probe task* that assesses the "selective attention" of an individual and therefore her/his proneness to vigilance to threat through her/his reaction time: a High Trait anxious subject responds more quickly to threatening stimuli. The application of this task would definitely enforce and enrich my results[55].

The motivational force of anxiety casts away the common view of anxiety as a negative emotion and brings to the discussion of a potential involvement of this emotion as an incentive. Specifically, a moderate[56] level of anxiety can be beneficial because it boosts performance. But the question can be: how to induce anxiety in a workplace? How does the principal ensure a certain level of pressure in the working environment? Can a psychological pressure be induced by monitoring?

---

some summary statistics. The mean for sadness, tiredness, anxiety, despondency, apprehension and happiness is higher in the control treatments. Only anger seems to be significant at 10% level (Wilcoxon rank sum test, p-value = 0.060). I checked for correlation with my dependent variable as well and I found a positive correlation between anxiety, despondency, apprehension and frustration (0.55, 0.49, and 0.41 respectively) but no variable is correlated with effort.

[54]"Persons high in Trait Anxiety tend to be higher in State Anxiety, even in relatively neutral situations" (Spielberger et al., 1970, p. 15).

[55]The z-Tree software used for my experiments does not allow the time measure in fractions of seconds. Hence, I did not implement the dot-probe task in my sessions.

[56]I do not provide any evidence on the effect of excessive level of anxiety.

An assumption of standard economic theory[57] is that incentives and monitoring are substitutes: an increase in monitoring increases the probability of detecting shrinking. Monitoring is in fact a way of preventing shirking since it is assumed to be frequent and highly significant activity in all principal-agent relationship (Frey, 1993; Sappington, 1991; Frey and Oberholzer-Gee, 1997)[58]. In the standard linear principal-agent model (Milgrom and Roberts, 1992) monitoring has no direct effect on effort; however, it reduces the agents' risk premium, permitting stronger incentives and therefore greater effort. In both models monitoring is associated with greater effort and expected performance. However, monitoring can be counter-productive because it may be a signal of distrust and thus reduces intrinsic motivation (Frey, 1993). Interesting future research can investigate this potential relation between anxiety caused by monitoring and the level of effort achieved.

2.6 Conclusion

This essay explores the motivational role of anxiety and its effects on individual levels of effort. Emotions have not been studied in the standard theory of incentive and personality implications have not found a large space in economics literature.

I believe that an anxious state can induce agents to work harder and therefore it can act as an incentive. To investigate my hypothesis I ran two economic experiments enriched by a psychological technique: I manipulated the mood of the participants and then I implemented a real effort task. When a certain level of arousal occurs, anxious subjects achieved better results than non-anxious subjects. This result is consistent with the Processing Efficiency Theory and with the Yerkes-Dodson

[57]The Efficiency Wage Theory (Akerlof and Yellen, 1986).

[58]Principal-agent theory assumes that individuals exert work effort to the point where net utility is maximized. Moreover, individuals are labour-averse: whereas his earned income provides benefits, the effort of earning procures disutility; therefore individuals will have an incentive to shirk.

Law which state the relationship between a given level of anxiety and performance. Moreover, I claim that "personality counts" (Lazear, 1989): High Trait Anxiety individuals respond better to the stimuli of the mood manipulation which is the cause of behavioural effects.

My results contribute to the literature of incentive and contract design, suggesting that emotions can have a motivational function and are important elements that drive behavioural choices on effort. Future research can explore the way in which these types of incentives may be beneficial in a context of group work, monitoring, and "mix and match" practices aimed at optimizing the team level of effort and performance.

A final remark is at this point necessary. The results I have presented are given by the analysis of data collected in the laboratory in which participants were required to carry out some tasks. Specifically, they were asked to perform a real effort task, the slider task. I did not run any control treatment to check the robustness of the obtained results with respect to the effort task used. The conclusion drawn in this chapter may depend on the effort task used.

Chapter 3

Does It Matter Which Effort Task You Use? A Comparison of Four Effort Tasks When Agents Compete for a Prize

3.1 Introduction

In laboratory economic experiments researchers ask participants to carry out tasks. Some of these tasks are used to measure participants' performance or to elicit their decisions of effort or investment. These tasks are called *effort tasks*. Effort tasks can be *real effort tasks* if subjects are required to put substantial (physical or mental) effort to perform them, or *induced value effort tasks* when participants' choice of their level of effort depends on a given cost structure assigned by the experimenter.

Generally, experimentalists implicitly assume that effort tasks are equivalent and that real effort tasks and induced value effort tasks lead to similar results.

However, there is no systematic study that either shows the methodological equivalence between effort tasks or that claims that one (or some) of them is (are) more appropriate tools for measuring subjects' performance or investment. There is a surprising variety of effort tasks used in economic experiments that can be operationalized in different ways. All of them have specific limitations and yet they have been typically used as virtually interchangeable.

This chapter addresses a crucial methodological research question in experimental economics: "does it matter which effort task you use in the laboratory economic experiments?". The importance of this research question arises from the necessity of making use of tools in the laboratory that do not compromise the ability to generalize the experimental results. The use of effort tasks which significantly affects laboratory results has the potential to cause a dramatic loss of external validity of the research.

Usually, only one effort task is implemented in the laboratory setting but no test is carried out to check whether the experimenter can collect similar experimental data by using another effort task. Researchers have not considered this important limitation of their studies and may not realise that their results depend on the effort task they have chosen.

This chapter tests the equivalence between effort tasks. Thus, to test the assumption that the use of different effort tasks does not change the quality of my results, I ran an economic experiment which consisted of four treatments identical except for the effort task applied. Three real effort tasks and an induced value effort task were implemented in the lab. The effort tasks provided a measure of subjects' level of performance (or investment). Participants were asked to compete in a two-player all-pay auction in order to win a prize. Those with the highest performance could get the prize. In this competitive setting I expect that an anxious emotional state will have a positive effect on subjects' level of performance (or investment). I expect that the more risk averse contestant will exert less effort, and that there will be gender difference in performance. These effects should not depend on the effort task used.

On the contrary, results show that experimental data on performance (or investment) significantly depend on the effort task chosen. I can therefore claim that the choice of the effort tasks to be implemented in the economic laboratory is an important methodological issue. Real effort tasks may not be equivalent among each other. Only some of them can minimize the individual's abilities in order to give a clean measure of exerted effort. Moreover, the operationalization of these tasks in the laboratory can represent a limitation of experimental studies. Finally, this work argues that real effort tasks and induced value effort tasks might not lead to similar results. Hence, the interpretation of the laboratory data and the conclusions of experimental research should acknowledge the effort task chosen.

It can be argued that the level of effort exerted, and hence the participants' performance, might depend on the relevance of the task to some groups of participants (stereotypes). Steele and Aronson (1995) provided strong evidence of the so-called *stereotype threat*: salience of some form of stereotype has a detrimental effect on performance of some identified groups of participants[1]. In my experiment I can rule out the possibility that performance was affected by any stereotype threat linked to the type of the effort task. I did not state, in fact, in the instructions of the experiment any scope related to gender, race, nationality, etc.

This chapter is structured as follows. In the next section the difference between induced value effort tasks and real effort tasks will be briefly given before clarifying advantages and limitation of each type of task. Some examples of their use in the literature will be provided. Section 3.3 illustrates the behavioural hypotheses, the experimental design, and procedure. Section 3.4 shows the experimental results. Section 3.5 and 3.6 discuss and conclude respectively.

[1]A more detailed discussion of the stereotype threat is provided in the discussion section of this chapter.

3.2 Related Literature

The laboratory experimental method in economics provides *ceteris paribus* observations on how a group of factors can affect individuals' behaviour and decisions. Many economic experiments aim, for example, to study the effects of different environmental and institutional variables on individual effort and task performance. For this purpose, the investigator might ask participants to perform a task, that is any piece of work assigned to or demanded of a person. Task characteristics can vary between tasks. Some of them could require cognitive skills or physical effort whereas others could require only "procedural knowledge"[2].

It is conventional to refer to *real effort tasks* as those tasks that imply a physical or mental effort and to *induced value effort tasks* as those tasks where individuals use their procedural knowledge to select a value that would represent their chosen effort.

Though laboratory effort tasks are less comparable to the ones that would occur in a natural work setting, their use gives the possibility to understand how individuals' choices depend on different variables and incentives. They are practical tools that allow to measure the level of effort exerted by subjects under different conditions, measure which would be difficult to get in a non-experimental setting.

Real and induced value effort tasks have been implicitly assumed to lead to similar results. However, the choice of the effort task may considerably affect experimental data. This is due to the substantial difference between effort tasks. In particular, real clerical tasks may differ in task learning, required skills and intrinsic motivation. Moreover, the implementation and operationalization of a real effort task might not always been carried out with a control for the heterogeneity in the distribution of the abilities.

[2] "Procedural knowledge" is a repertoire of skills, rules and strategies for using knowledge. Experimenters are usually interested in procedural knowledge and, therefore, they are concerned in writing easy and accessible instructions that would allow all subjects to have the knowledge to understand how their decisions affect their payoffs (Camerer and Hogarth, 1999).

Measurements of effort in a controlled environment where individual values are specified by a precise utility function is offered by the use of an induced value effort task, even though the over-simplistic setting and the lack of context might compromise the validity of the experiment.

In this work I aim to point out that the choice of an effort task may affect the laboratory results and that any interpretation of them should acknowledge the nature of the task used. Therefore, I will try to answer the question "does it matter which effort task you use in the laboratory?", and more specifically, "do real effort tasks and induced value effort tasks lead to similar results?".

In the rest of this section I will organize my discussion as follows: I am going to illustrate first what induced value effort tasks are (section 3.2.1) and second what real effort tasks are (section 3.2.2), clarifying in each section their advantages and limitations, and providing some examples of their implementation in the related literature.

3.2.1 Induced Value Effort Tasks

A clear test in economic experiments of the hypotheses requires controlling for subjects' preferences. But this is a major difficulty because individuals' evaluations and utilities are not observable by the experimenter. One way to deal with this is to provide the subject an appropriate monetary incentive which depends on a specified utility function. This so-called *induced value method* assumes that given a costless choice between two alternatives, identical except that the first gives a higher payoff than the second, the first will be always chosen over the second by an individual whose utility is a monotone increasing function of the monetary reward" (Smith, 1976).

The induced value method has been widely applied in testing, for example, contest theory. In contests agents exert costly effort to compete for a prize. In equilibrium, effort is chosen in such a way that marginal effort costs equal marginal gains. A direct empirical test of this theory requires that researchers know different parameters such as, for example, the number of agents, the effort cost functions, the exact level of the prize, and the production function including the nature of the error term (if any). The knowledge of these parameters allows the derivation of a precise prediction of chosen effort. Researchers can therefore decide appropriate parameters. The test of the theory will occur by observing subjects' chosen level of effort.

The above type of test would not be possible with the same level of confidence in a non-laboratory environment. "The existence of these superior control is the most important asset behind running experiments and such control can be achieved by using a reward structure *to induced predefined monetary values on actions*" (Smith, 1976).

One of the first contest experiments was conducted by Bull et al. (1987). They asked subjects in a two-player tournament to choose an integer number, called "decision number", between zero and 100. Corresponding to each decision number there was a cost listed on a table given to subjects. Schotter and Weigelt (1992) followed the same procedure of Bull et al. in an experiment on asymmetric tournaments. Sheremeta and Wu (2012) asked subjects to pick up a number from 0 to 120, representing their level of effort in a tournament experiment where the role of the principal was endogenous. Harbring and colleagues (Harbring and Irlenbusch, 2003, 2005; Harbring and Lünser, 2008) used different parameters in their experiments in order to measure agents' effort/bidding in tournaments. Participants could choose in fact an integer number from 0 to 100.

Induced value effort tasks are also implemented in other type of contests. Breitmoser et al. (2010) used an induced value effort task in their R&D races experiment where participants' choice of effort (or investment) was binary: low or high effort. In examining bidding decisions in a lottery contest Millner and Pratt (1989) gave an initial endowment to subjects of £12 to buy lottery tickets of the value of £0.10.

In this case the level of effort (investment) could vary from 0 to 12 taking a value up to one decimal place.

Induced value effort tasks have not been only used in contest experiments. Research on fairness and social preference aims to investigate the subjects' level of effort under certain condition using this straightforward tool. Gächter and Thöni (2010), for example, were interested in the impact of wage comparisons on agents' productivity. In their experiment participants could choose a level of effort between 1 and 20 and their related costs were linear in effort. Fehr and Schmidt (2004) investigate experimentally the role of fairness in the multi-task principal-agent model where principals could either offer a piece-rate contract or a voluntary bonus to the agent. Agents could select their level of effort from a set of integer from 1 to 10, whose cost was given by an increasing and convex cost function. An increasing positive cost function was used by Fehr et al. (1997) as well. In their experiment on the impact of reciprocity on contract enforcement effort was restricted to the interval [0,1]. Participants could choose their level of effort from a list of 14 options. Effort was a decimal number up to 3 decimal places. Similarly, Fehr et al. (1993) provided participants a table showing a list of ten options for the choice of effort. In their experiment designed to test the impact of fairness on market prices, participants were asked to choose a level of effort from 0.1 and 1 up to one decimal place. In the table provided, associated costs calculated according to an increasing cost function appeared as well. Generally, a chosen effort task is used in research on fairness and reciprocity in labour markets. Experimentalists decide the parameters of their experiments, such as the type of cost functions and initial endowment, and ask participants to select their desirable level of effort from a given list. Levels of effort are always within a specified range (Charness, 2004; Brandts and Charness, 2004; Charness et al., 2004; Gächter and Falk, 2002; Fehr and Gächter, 2000).

The above mentioned papers in contest theory and fairness in the labour market are just a few examples of the use of induced value effort tasks in the laboratory. This chapter is not intended to provide a full review of the use of chosen effort tasks in economic experiments. It just aims in this section to illustrate some possible ways of implementing this technique. To summarize, in experiments where an induced value effort task is implemented subjects generally receive an endowment and they are asked to choose a number subject to their endowment. The chosen number would represent their level of effort (or investment). Each possible level of effort is associated with a certain level of cost according to a specified cost function imposed by the experimenter.

While the imposition of a monetary cost function in induced value effort tasks assures experimenters full control over the cost of effort[3], whether the conditions implemented in the laboratory are also present in reality is an open debate. Is the induced value effort task a good replication of the agents' decision of real effort? One way of adding realism to the laboratory experiments is to use so called *real effort tasks*.

3.2.2 Real Effort Tasks

Real effort tasks require substantial effort, they imply that subjects have to actually work on an experimental task. For example, some researchers ask participants to solve some mathematical questions (Vesterlund and Nierderle, 2007; Eriksson et al., 2008; Dohmen and Falk, 2011). In Konow (2000), Falk and Ichino (2006), and Konow et al. (2009) subjects have to fold some letters, and in Fahr and Irlenbusch (2010) participants are asked to crack walnuts.

Following a classification adopted by Bortolotti (2004), real effort tasks can be categorized as either *cognitive tasks*, where task performance mainly depends on cognitive abilities, or *physical tasks* which require some physical strength to

[3] In particular, the experimenter can control the extent of any convexity in the cost of the activity and can also determine how the cost varies over individuals and over any repetition of the game.

be performed[4]. Tasks involving mathematical skills are an example of cognitive tasks. For instance, Sutter and Weck-Hannemann (2003), Dohmen and Falk (2011), Brüggen and Strobel (2007), and Kuhnen and Tymula (2008) all ask subjects to solve mathematical equations and multiplications. Vesterlund and Nierderle (2007) use additions of two-digit numbers. A two-variable optimization task is introduced by van Dijk et al. (2001) where subjects have to search by trial and error the maximum of a function[5].

Lévy-Garboua et al. (2009) ask instead to decode a number from a grid of letters. Linguistic skills are needed for the task operationalized by Charness and Villeval (2009) where students are asked to solve some anagrams. Memory and logic are necessary to solve mazes in Gneezy (2002) and in Gneezy et al. (2003), and to succeed in memory games (Ivanova-Stenzel and Kübler, 2005) or word games (Burrows and Loomes, 1994). In experiments run by Hoffman et al. (1994) and by Hoffman and Spitzer (1985) participants face some "current events" quiz.

Physical tasks may be, for example, secretary tasks such as typing abstracts (Hennig-Schmidt et al., 2005), or entering information about library books into a database (Gneezy and List, 2006; Kube et al., 2006), or copying some information into a database (Ottone and Ponzano, 2008) or entering strings of characters on the screen (Dickinson, 2001). Konow (2000), Konow et al. (2009), Falk and Ichino (2006), and Carpenter et al. (forthcoming) ask students to fold some letters and to stuff them into envelopes. Azar (2009) asks participants to find a letter in a non-justified grid of letters[6], and Heyman and Ariely (2004) design an experiment in which a

[4]Real effort task can be also divided between *real effort task without real outcomes*, such as solving arbitrary mathematical problems, and *real effort task with real outcomes*, such as folding letters or fund-raising. It might be claimed that real effort tasks with real outcomes may incentivize subjects to put more effort than real effort tasks without real outcomes do. In my experiment, all the three real effort tasks I use are real effort tasks without real outcomes.

[5]See also Bosman and Van Winden (2002), and Bosman et al. (2005) and for a modified version of the game see Dickinson and Villeval (2004), Montmarquette et al. (2004), and Sloof and van Praag (2008).

[6]The distinction between cognitive tasks and physical tasks is not sharp. In this case, for

computerized ball had to be dragged from one place to another. Manual tasks such as cracking walnuts (Fahr and Irlenbusch, 2010) or a door-to-door fund-raising (Gneezy and List, 2006) have also been used. Finally, in Pokorny (2008), Abeler et al. (2011), Falk et al. (2006), and Gill and Prowse (2011) subjects have to count the number of zeros in a series of tables or to drag some sliders to a certain position on the screen[7].

Overall, real effort tasks share common characteristics: first, tasks should be easy enough to be possibly performed by everyone without an extensive training, and second they should be uninteresting so that they minimize subjects' intrinsic motivation. Real effort tasks try to achieve a greater external validity of the experiment, because the experimental design replicates the real exertion of effort outside of the laboratory. Compared to induced value effort tasks, however, they imply a loss of control over the monetary cost function. In real effort experiments disutility from work (or effort) cannot be modelled as a monetary cost as it is usually done in experiments where the induced value method is applied. When moving from an induced value effort task to a real effort task, how to measure disutility is a crucial issue.

One possible way to address the above problem is to use some proxies for exerted effort (and experienced disutility). I can identify at least four proxies for effort in the experimental literature: quantity, quality, level of difficulty of the task, and time spent completing the task.

For example, experimenters consider as performance the number of solved equations, the number of sliders correctly positioned, the number of entered records, etc.

example, cognitive skills may be useful for the search of letters. However, the cognitive skills involved would be very minimal and hence the task might be better considered as a physical task.

[7]This work is not intended to be a complete review of the effort tasks. For example, I do not consider studies that implemented a real effort task to legitimize subjects' initial entitlement (Hoffman and Spitzer, 1985).

In other experiments, where the output could have different levels of quality and thus the output quantity is not a good proxy of effort, experimenters consider the number of wrong answers or mistakes.

However, the number of wrong answers or mistakes are not always indicators of lack of effort. There is usually a skill component required in each task that can be minimized but never ruled out completely. Therefore, some subjects are more likely to make more mistakes than more skilful subjects no matter how much effort they exert.

For this reason the level of difficulty of a task should be taken into account when choosing the task. Sometimes, the level of difficulty is used to measure effort: in Gneezy et al. (2003) subjects can choose the desired difficulty level of the maze themselves.

Time can also be used as a proxy for effort under the assumption that the longer time is spent on a specific task the higher effort is exerted (and hence the disutility experienced). However, the time spent on a task can be also a measure of ability. A more skillful subject might spend less time in completing the task.

Some examples in the literature can be found about the combined use of these four proxies. For instance, Hennig-Schmidt et al. (2005) and Kube et al. (2006) consider the number of both entries and mistakes in their library task. Carpenter et al. (forthcoming) adjust the output for quality. Azar (2009) takes into account time spent working on the task in addition to quantity of the output and difficulty of the task.

Yet, all the proxies discussed are all measures of performance rather than effort and disutility of work. I shall focus on performance which is what we usually observe in the lab. Performance is a function of effort (and ability). Effort and ability cannot be directly measured in the lab. Moreover, no measure discussed so far can actually disentangle completely effort and individual differences in innate abilities and skills.

Thus, the presence of heterogeneity points out the difficulty of the identification of subjective costs of decision discussed by Smith (1979).

It follows from the discussion above that the choice of an effort task represents a crucial issue in designing an experiment. On the one hand, researchers face a trade-off between control of the investigated variables and more realism in the experimental environment, and on the other hand, if they opt for a real effort task they need to take into account all the implications of the use of such a task in terms of operationalization.

Surprisingly, in the experimental literature there is no clear account of the limitations and the problematic use of effort tasks in the lab. Very often, laboratory experiments in the same area of investigation apply either induced value effort tasks or real effort tasks making implicitly the assumption that both techniques lead to similar results.

In this work I aim to investigate whether the choice of an induced value effort task or a real effort task actually leads to similar results and whether any difference in my results is due to the type of effort task used. Hence, my experiment will answer the following questions: "does it matter which effort task you use in the lab?".

3.3 The Experiment: Research Hypotheses, Design and Procedure

3.3.1 Research Hypotheses

In order to test for any equivalence between four effort tasks chosen within the most popular tasks in experimental settings I ran an economic experiment which consists of four identical treatments except for the effort task used. I have chosen three real effort tasks and an induced value effort task for this experiment. The real effort tasks are: the slider task (Gill and Prowse, 2011), the mathematical task (Vesterlund and Nierderle, 2007), and (a variation of) the counting task (Abeler et al., 2011).

In each treatment subjects are asked to expend costly effort to compete in a contest for a prize. Subjects' decision of the level of exerted effort may be driven by some emotional factors. For example, it depends on the arousal of anxiety that occurs in a competitive setting. My hypothesis is that anxiety improves subjects' performance. This hypothesis is supported by the *Processing Efficiency Theory* (Eysenck and Calvo, 1992) according to which the individual's initial effect to anxiety is negative because worries interfere with attention to the task-relevant information, thus reducing the resources potentially available for the task. However, anxiety has a motivational function that induces the subject to increase effort and therefore to boost performance in order to:

> *[...] avoid the likely aversive consequences of poor performance. [A]nxious subjects try to cope with threat and worry allocating additional resources (i.e. effort) and/or initiating processing activities (i.e. strategies)* (Eysenck and Calvo, 1992, p. 415).

I expect therefore that an increase in anxiety can improve performance because the individual is motivated to increase effort to such an extent that it counterbalances and exceeds the initial negative impact of worry[8].

The level of exerted effort (and the likelihood of winning of each participant) in contests also depends on the individual's risk attitude since the outcome of a contest is typically uncertain. A number of authors have investigated the effects of risk aversion in contests. Many of them compare equilibria under risk aversion with the corresponding contest in which players are risk neutral and investigate whether risk aversion reduces total effort on rent-seeking. Results have usually

[8]The first essay of this thesis discusses more in details how anxiety improves performance.

shown that risk aversion decreases rent-seeking efforts. However, the results have been obtained under restricted specifications. For instance, Hillman and Katz (1984) assume that the rent is 'small'. Skaperdas and Gan (1995) consider rent-seekers' utility functions with constant absolute risk aversion and assume that the contest success function is logistic with a power or exponential form. Konrad and Schlesinger (1997) using general utility and contest success functions indicate that "it is possible for the contest with risk averse players to dissipate more of the rents than the same contest with risk neutral players" (p. 1677). This suggests that risk aversion has an ambiguous effect on rent-seeking efforts.

Millner and Pratt (1991) carried out an experiment that investigates the possibility that risk attitude may explain rent overdissipation. They grouped together more risk averse participants and they compared effort in aggregate between groups. Results show that the more risk averse group exerted lower effort in aggregate than the less risk averse group[9].

Following this experimental evidence I expect that the more risk averse individuals will exert less effort than less risk averse individuals.

Individual behaviour in contests depends on gender as well. A number of experimental studies do find gender differences in behaviour and performance (Croson and Gneezy, 2009; Niederle and Vesterlund, 2011). Specifically, women are less likely to enter tournaments than men are (Vesterlund and Nierderle, 2007; Balafoutas and Sutter, 2012) and they do not perform as well as men under tournament incentives (Gneezy et al., 2003; Gneezy and Rustichini, 2004). Differences in performance between men and women depend on the type of task as well. Gneezy and Rustichini (2004) used a field setting wherein pairs of schoolboys and schoolgirls compete in a footrace. In this type of task, boys achieved better performance relative to girls.

Researchers explain gender differences with individuals' degree of confidence and

[9] The effects of risk aversion in contests have been largely discussed (van Long and Vousden, 1987; Treich, 2010; Cornes and Hartley, 2012, 2003; Yamazaki, 2008).

attitude toward competition[10]. Since there is no possibility for participants in my experiment to choose whether to enter the competition or not, I can observe their decision of investment only. Following the experimental evidence, I should expect that men will exert more effort than women in contests and hence they will perform better.

Therefore, I formalize my hypotheses as follows:

Hypothesis 1: *Anxiety does not affect performance in contests.*

against

Hypothesis 1a: *Anxiety positively affects performance in contests.*

Hypothesis 1 follows the neoclassical economic theory: emotions do not affect agents' performance in contests. The alternative behavioural hypothesis (hypothesis 1a) refers to the Processing Efficiency Theory, an established theory in psychology that states the motivational function of anxiety on task performance.

Assuming that there is equivalence between effort tasks, any (positive) effect of an anxious mood would not change across treatments.

Based on the evidence in the experimental literature on the effects of risk attitude and gender on performance in contests, I will formalize my null and alternative hypotheses as follows:

Hypothesis 2: *Risk aversion does not affect individuals' performance in contests.*

[10]Gender differences can be explained, though not completely, by differences in attitude toward risk: women tend to be more risk averse than men, they bid more aggressively and hence they receive significantly lower earnings from the contests than men (Morgan et al., 2012; Ong and Chen, 2012; Price and Sheremeta, 2012).

Hypothesis 2a: *The higher the risk aversion, the lower the individuals' performance in contests.*

Hypothesis 3: *There is no gender difference in performance when men and women compete in contests.*

Hypothesis 3a: *Men perform better than women in contests.*

Assuming that there is equivalence between effort tasks, any effect of risk aversion would not change across treatments. Any gender effects would be consistent across treatments as well. Effects of anxiety, risk aversion and gender might have different magnitude but they should have the same direction on individuals' performance in each treatment.

3.3.2 Design and Procedure

The experiment was conducted in Spring 2012 at the Zicer laboratory of the Centre for Behavioural and Experimental Social Science (CBESS) at the University of East Anglia. Instructions, control questionnaire and experimental tasks were computerized[11]. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). 206 subjects (46% male and 54% female) aged between 18 and 43 took part in the 24 sessions. All of them were university students, enrolled in different schools (only 12% were Economics students); 54% had British nationality.

Each session lasted approximately one hour. Subjects were randomly assigned to the seats in the lab where partitions between them assured the anonymity and the avoidance of any communication. Instructions were read aloud and the experimenter checked individually the participants' answers to the control questionnaire[12].

Sessions consisted of three stages: 1) in the first stage participants were asked to fill in a questionnaire; 2) in the second stage participants were asked to participate

[11] Participants received paper instructions as well.

[12] See Appendix B.1 for experimental instructions.

in a 2-player contest; and 3) in the last stage subjects completed a risk aversion questionnaire.

The Initial Questionnaire

At the beginning of each session, before receiving the instructions for the successive tasks of the experiment, participants were asked to fill in the State-Trait Anxiety Inventory (STAI) questionnaire by C. Spielberger (1976). This questionnaire aims to assess the current anxious state (State-Anxiety) and the proneness to anxiety (Trait-Anxiety) of subjects. It consists of 20 questions on the State-Anxiety and 20 questions on the Trait-Anxiety and it is framed in order to avoid any kind of demand effect[13]. The STAI questionnaire assigns two scores, one for State-Anxiety and another for Trait-Anxiety; for each subject there are therefore two measurements from 0 to 80[14].

The Contest and the Effort Tasks

In the second stage of the session participants received the instructions for a 2-player contest. They were asked in fact to enter a competition against another participant in the room in order to win a prize of 10 pounds.

The competition is designed as an all-pay auction, a type of contest in which the contestant who expends the highest effort wins the prize with probability 1. Both players have a positive valuation of the prize, they both know their own valuation and the valuation of their opponent. Contestants choose their effort $e_i \in [0, \infty)$ simultaneously, and the cost of effort is linear $C(e_i) = e_i$.

Therefore, in the case of two contestants $i = 1, 2$ contestant 1 wins with probability

[13] In Chapter 2 of the thesis I have described more in details the procedure of the STAI questionnaire.

[14] Scores are adjusted according to the number of not answered questions. If there are 3 or more questions not answered for each set of questions (Trait or State), the score for that set is considered missing (Spielberger, 1972). See Appendix A.1 for the instructions and statements of the STAI questionnaire.

$$p_1(e_1, e_2) = \begin{cases} 1 & \text{if } e_1 > e_2 \\ \frac{1}{2} & \text{if } e_1 = e_2 \\ 0 & \text{if } e_1 < e_2 \end{cases} \tag{3.1}$$

The probability with which contestant 2 wins is $p_2 = 1 - p_1$.
Contestant 1's payoff is

$$\pi_1 = \begin{cases} P_w - e_1 & \text{if } e_1 > e_2 \\ 5 - e_1 & \text{if } e_1 = e_2 \\ P_l - e_1 & \text{if } e_1 < e_2 \end{cases} \tag{3.2}$$

where P_w and P_l are the prizes given to the winner and to the loser of the tournament respectively[15].

Contestants' effort in this experiment is given by the performance in the effort task. Four different effort tasks, the "Slider Task", the "Maths Task", the "Grid Task" and the "Induced Value Effort Task" were applied separately in four different treatments, Treatment A, B, C and D respectively. Table 3.1 shows some experimental details for each treatment[16].

Participants dealing with the *Slider Task* saw on their computer screen forty-eight sliders, with a range of integer values from 0 to 100, appeared at position zero[17] and they were asked to position with the help of their mouse each slider in the central

[15]In this experiment P_w is equal to £10 and P_l is zero pounds.

[16]Treatments were run on a between-subject basis. The number of subjects per treatment is slightly different because some students did register for the session but they did not show up. The average payment per subject was 10 pounds and the payment was based on subjects' performance. Participants received a show-up fee of £3 in the sessions where a real effort task was applied (Treatment A, B and C) in order to align the average payment between treatments. In treatment D, in fact, average earnings from the effort task was already high because the endowment not invested was converted in pounds and paid to subjects. In the other treatments there was no initial endowment.

[17]Sliders appeared at the very left position. See Appendix B.2 for a screen shot (Figure B.1).

Table 3.1: Experimental Treatments, Effort Tasks and Number of Subjects

Treatment A	**Treatment B**	**Treatment C**	**Treatment D**
Slider Task	Maths Task	Grid Task	Induced Value Effort Task
50 participants	54 participants	52 participants	50 participants
Total: 206 participants			

position (value 50). Once the slider was dropped, the value of its position was shown to the right. If not 50, participants could adjust the position again. There was no limit on the number of attempts per each slider. Students had two minutes to correctly position as many sliders as they could. Subjects' performance is given by the number of sliders correctly positioned in two minutes[18].

For the *Maths Task* participants were asked to sum a series of four 2-digit numbers that appeared on their screen. They were provided of a list of operations and they were asked to solve as many as they could in the time constraint (2 minutes). The number of correct answers gives the subject's performance. Table 3.2 shows an example of this task[19].

For the *Grid Task* a 5 by 5 grid appeared on the screen and subjects were asked to count the number of ones randomly assigned in the cells. They had two minutes to solve as many grid as they could[20]. An additional difficulty characterized this task: some numbers in the grid disappear after few second. However, those numbers needed to be taken into account in the counting. In this task, participants' effort is

[18]Guidance and practical advice of the use of this task have been followed.

[19]Participants could see on the screen only one question per time. After submitting their answer, they could proceed to the next question.

[20]This task is similar to the task by Abeler et al. (2011). Their task consisted of a series of 10 by 15 grids filled in randomly by numbers zeros and ones. Subjects were asked to count the number of zeros in each grid. No time limit was imposed.

Table 3.2: Mathematics Tasks

Please add up the following numbers:	
Question 1:	21 + 23 + 45 + 67 + 88 =
Question 2:	34 + 14 + 44 + 67 + 19 =
Question 3:	51 + 65 + 21 + 13 + 33 =
Question 4:	82 - 13 + 49 + 72 - 15 =
...	....

given by the number of grid correctly solved[21].

In Treatment D I have used an *Induced-Value Effort Task*: participants were endowed with ten experimental points and they had to decide how much to invest and how much to keep for themselves[22]. The part of endowment not invested (if any) was converted in pounds and paid to subjects at the end of the experiment. Participants with the highest investment won the prize (one prize per pair).

In each treatment subjects repeated the effort task for ten rounds. They did not receive any feedback between rounds. At the end of the session one round was randomly selected for the payment. Contestants with the highest performance (or investment) in each pair won the prize. If contestants had the same level of performance (or investment) the prize was randomly assigned by the computer to one of the two.

Risk-measurement Questionnaire

In the third stage participants completed the Holt-Laury Questionnaire[23], an incentive compatible lottery widely used to measure subjects' risk aversion. The risk

[21]See an example of a screen shot in Appendix B.2 (Figure B.2).

[22]Invested amounts could be up to two decimal points.

[23]Holt and Laury (2002).

aversion measure can take values from 0 to 9. The higher the number the more risk averse the participants.

At the end of the experiment, a computer screen showed to participants some information on their payment: the outcome of the lottery, the selected round of the effort task, the outcome of the contest and hence the final earnings. In each session pairs were fixed and there was no feedback within rounds.

Emotions Elicitation

To get a measure of the anxiety arousal during the contest subjects were asked to elicit their emotional state using the Visual Analogue Scale (VAS): a bar appeared on the screen whose value 0 indicated the mood evaluation "not anxious at all" and value 100 the mood evaluation "very much anxious". The bar could take any integer number between zero and 100 and participants could scroll the bar up to the desired position. The VAS measurement has been applied three times in the session: at the beginning of the real effort task, after 5 rounds, and at the end of the last round.

The same type of measurement was applied to collect data on emotional arousals of other 7 emotions. The used list is by Slyker and McNally (1991): tiredness, despondency, frustration, apprehension, tiredness, happiness, anger[24].

In this analysis the self-elicitation arousal values collected before the first round and before the 6th round of the effort task will be considered as the main explanatory variables of the subjects' mood while the arousal anxiety values at the end of the tenth round of the effort task will be used as control for subjects' current emotional state. Table 3.3 shows the procedure of the experiment.

3.4 Results

Performance

[24] Slyker and McNally (1991) tested the relative effectiveness of some mood inducement procedures for anxious and depressed moods.

Table 3.3: Procedure of the Experiment

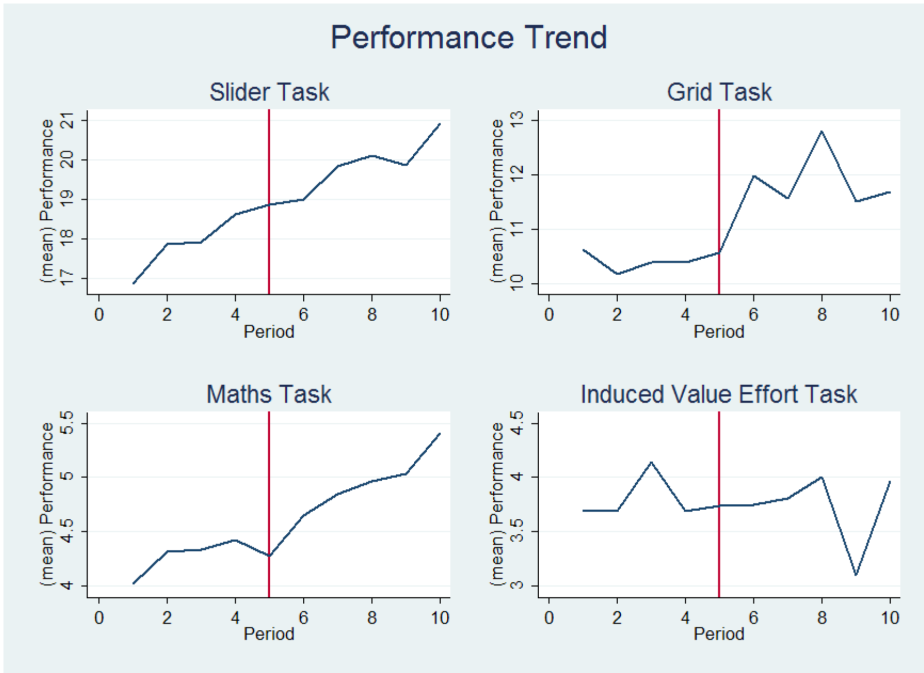
STAGE 1	STAGE 2	STAGE 3
STAI Questionnaire		
Instructions - Practice Round		
	Emotion Elicitation (VAS)	
	5 rounds Effort Task	
	Emotion Elicitation (VAS)	
	5 rounds Effort Task	
	Emotion Elicitation (VAS)	
		Risk-attitude measurement
		Results and Payment

Figure 3.1 shows that performance increases over time especially in treatments where a real effort task was used, indicating some learning effects. The red vertical line in the figure indicates the fifth round: at the end of this round subjects' emotions arousal was elicited. On average, after the fifth round subjects' performance increases monotonically in the slider task and in the maths task. The Spearman correlation coefficients are positive and strongly significant (slider task: 0.222; maths task: 0170). Also in the grid task there is a positive trend after the fifth round, as confirmed by a significant Spearman coefficient (0.103). Learning effects do not appear in the treatment where an induced value effort task is applied: on average performance does not increase monotonically. The Spearman correlation coefficient is in fact not significant.

Treatment C (the grid task) allows us to observe the highest heterogeneity between subjects (highest variance) in both sets of rounds (Table 3.4).

In the slider task (Treatment A) the maximum number of sliders that participants can correctly position is 48 per round; they can decide to not do the task at all (so that the minimum is 0). However, at least 6 and at most 32 sliders were correctly positioned in 2 minutes. The maths task and the grid task seem to be more difficult

Figure 3.1: Performance Trend



tasks: at least one subject could not solve any mathematical question or grid. There is no upper limit in questions/grid that could be solved for the maths task and for the grid task: the computer could generate an infinite number of grids or mathematical questions in 2 minutes. The best that participants could do is 12 for the maths task and 25 for the grid task. Performance (or investment) in the induced value effort task has a much defined range value because subjects receive an endowment of 10 experimental points and they cannot neither invest more than that nor invest a negative amount. The modal range value for the investment decision is between 3.5 and 4. However, some participants decided to invest everything or not at all (see Table 3.4)[25].

[25]Figure B.3 and Figure B.4 in Appendix B.2 show the distribution of average performance in each treatment. It is straightforward to notice how individuals' heterogeneity is more evident in treatments where a real effort task has been used compared to treatment D where participants carried out the induced value effort task. It seems that the heterogeneity due to individuals' difference in skills might affect the average performance level.

Table 3.4: Aggregate Descriptive Statistics of Performance - by Sets of Rounds

	Obs	**Mean**	**Std.Dev**	**Min**	**Max**
Rounds 1 - 5					
Slider Task	250	18.04	4.329	6	28
Maths Task	270	4.27	1.953	0	9
Grid Task	260	10.43	6.059	0	25
Ind. Value Effort Task	250	3.79	3.595	0	10
Rounds 6 - 10					
Slider Task	250	19.95	4.532	9	34
Maths Task	270	4.98	2.303	0	12
Grid Task	260	11.91	6.031	0	25
Ind. Value Effort Task	250	3.73	3.719	0	10

A Wilcoxon Mann-Whitney test and the t-test have been run to check for significance between sets of rounds. The Wilcoxon-Mann-Whitney test confirms that performance is significantly different between the two sets of rounds in treatments in which a real effort task is used (slider task, maths task, and the grid task). The t-test indicates that performance is higher in the second half of the rounds in those treatments. No difference among rounds can be found in Treatment D where an induced value effort task is applied[26].

Although comparing the performance values and distributions among treatments could bring some interesting discussion points, it is very important to clarify that the present work does not wish to make any claim on the difference in performance between treatments. Rather my focus will be on how some independent variables, e.g anxiety, risk attitude, and gender can affect a dependent variable (individuals' performance) in a significantly different way according to the effort task used.

[26]Table B.1 in Appendix B.2 shows the z-values for the Wilcoxon Mann-Whitney test and the t-values for the t-test with the correspondent p-values.

Table 3.5: Anxiety, Performance by Gender, and Risk Aversion - Average Values

	Slider Task Treat. A	Maths Task Treat. B	Grid Task Treat. C	Ind. Value Effort Task Treat. D
Rounds 1 - 5				
Anxiety	45.12	56.07	49.67	26.42
Performance - Male	19.508	4.6	10.928	3.246
Performance - Female	16.685	4.067	9.963	4.341
Rounds 6 - 10				
Anxiety	40.20	59.41	43.21	24.88
Performance - Male	21.7	5.257	12.168	3.189
Performance - Female	18.323	4.806	11.681	4.264
Risk Aversion	5.38	5.91	5.5	6.12

Anxiety

On average the level of anxiety is higher for participants that carried out the mathematical task (Treatment B) compared to participants that faced the other tasks in both sets of rounds (Table 3.5). Differently from the other treatments, the level of anxiety in Treatment B increases in the second set of five rounds. Overall the real effort tasks used allow us to see a similar range of variance in the elicitation of the average level of anxiety (see Table B.3 in Appendix B.2 for summary statistics).

The difference in average level of elicited anxiety between the two sets of rounds in the treatment where the grid task is applied (Treatment C) is statistically significant at the 10% level as the Wilcoxon signed-rank test shows (z-value = 1.662, p-value = 0.096)[27]. The differences in the average elicited anxiety in the other treatments are not significant[28].

[27]A t-test indicates that subjects dealing with the grid task elicited a significantly higher average level of anxiety in the first five rounds compared with the second five rounds (t-value = 1.925, p-value = 0.029).

[28]The Wilcoxon signed-rank test shows in Treatment A (slider task) that subjects have elicited a higher average level of anxiety at the beginning of the first round compared the average level of

Risk attitude

Table 3.5 shows that on average participants were more risk averse when the induced value effort task was applied (Treatment D). A high degree of risk aversion compared to Treatment A and C is captured in Treatment B as well where the maths task was applied[29]. Subjects that in Treatment D faced the induced value effort task were significantly more risk averse than subjects performing the slider task in Treatment A and significantly more risk averse than subjects performing the grid task in Treatment C[30].

Performance by Gender

In both sets of rounds men perform significantly better than women in the slider task and in the maths task (Treatment A and B)[31]. There is instead no significant difference between the average performance in Treatment C (grid task) between men and women. In both set of rounds all participants perform roughly at the same level. On the contrary, women decided on average to invest significantly more than men in the induced value effort task (Treatment C) in both sets of rounds[32].

Regressions on Subjects' Performance

Table 3.6 shows the estimates of Ordinary Least Squares (OLS) regression models with error clustering on the aggregate level of performance[33] of the first five rounds of the effort task. The independent variables include the VAS score on anxiety, the risk

anxiety elicited at the beginning of the sixth round (10% level of significance). See Table B.2 in Appendix B.2.

[29]See Table B.4 in Appendix B.2 for descriptive statistics on risk aversion.

[30]Two-sample t tests and Wilcoxon ranksum tests were calculated.

[31]See Table B.5 in Appendix B.2 for descriptive statistics on performance by gender.

[32]Two-sample t tests and Wilcoxon ranksum tests were calculated.

[33]The use of clusters provides a robust standard error per group (that is, per subject). Sashegyi et al. (2000) argue that where observations over time are taken for different group of subjects an econometric model must control both for intra-cluster correlation and intra-individual correlation within the same cluster.

aversion score given by the Holt-Laury lottery, the State-Anxiety and Trait-Anxiety score and some demographic information on participants: gender (Male = 1 if the subject is male), school (School of ECO = 1 if the subject is enrolled in the School of Economics at the University of East Anglia), and nationality (British Nationality = 1 if the subject is British). The age of each participant is included as well.

Table 3.6: Regressions with Error Clustering on Performance of the First 5 Rounds

	Slider Task	Maths Task	Grid Task	Ind. Value Task
Anxiety	0.027	-0.017	-0.061**	0.007
Risk Aversion	0.147	-0.135	0.840**	-0.133
State-Anxiety	-0.006	0.028	0.083	0.006
Trait-Anxiety	-0.146**	0.013	-0.096	0.043
Male	2.54**	0.298	-0.112	-0.777
Age	-0.087	0.029	-0.246	0.113
British Nationality	-2.164**	0.422	4.618**	-0.511
School of ECO	-3.457	-0.548	3.552	0.302
Constant	24.029***	3.510	11.855	0.308
Clusters	47	53	52	49
Observations	235	265	260	245
R-Square	0.2381	0.0891	0.2321	0.0773
F - test	3.88	0.92	2.33	1.23

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

There is strong evidence that a higher level of elicited anxiety negatively affects subjects' performance in the treatment in which the grid task was applied (Treatment C): an increase in anxiety level brings to deterioration in performance. Anxiety seems not to be a determinant of subjects' performance when the other effort tasks are applied. Rejecting hypothesis 1, anxiety does affect performance in contests when the grid task is applied. However, the same direction of effects is not observed when the other effort tasks are applied.

In the same treatment (Treatment C - grid task) risk aversion is strongly significant as well. It positively affects participants' performance. In the other treatments, however, this variable does not show a significant coefficient. Rejecting hypothesis 2, risk aversion does affect performance in contests when the grid task is applied. However, the same direction of effects is not observed when the other effort tasks are applied.

Male participants seem to perform better than female participants in the slider task and in the maths task compared to the other tasks. As predicted by the hypothesis 3a, male participants perform significantly better than female participants. However, this happens only in Treatment A (the slider task).

The variable that represents the proneness to anxiety, measured by the Trait-Anxiety score, has a negative coefficient in the treatments where either the slider task or the grid task is used. However, only in the former one the coefficient is significant at 5% level[34].

The same group of independent variables (anxiety, risk aversion, State-Anxiety score and Trait-Anxiety score, and four demographic information) are used to estimate the coefficients of OLS regressions with error clustering on the level of performance in the second five rounds. However, in these new models I consider the level of elicited anxiety (VAS score) submitted by subjects at the beginning of sixth round. Table 3.7 shows the coefficients of the regression models.

From the two sets of regressions described above (Table 3.6 and Table 3.7) I can highlight the following results:

[34] The Spearman correlation coefficient between performance and the Trait-Anxiety score has a negative and significant sign. See Table B.6 in Appendix B.2. Graph B.9 and Table B.7 in Appendix B.2 show the distribution of the scores of Trait-Anxiety and some summery statistics of the scores of Trait-Anxiety respectively.

Table 3.7: Error Clustering on Performance of the Second Set of 5 Rounds

	Slider Task	Maths Task	Grid Task	Ind. Value Task
Anxiety	-0.000	-0.026**	-0.046*	0.023
Risk-Aversion	-0.219	-0.092	0.459	-0.171
State-Anxiety	0.005	0.033	-0.011	-0.004
Trait-Anxiety	-0.090	0.015	-0.051	0.044
Male	3.268**	0.202	-0.914	-0.691
Age	0.020	0.144	-0.137	0.134
British Nationality	-2.630**	0.948	4.447**	-0.633
School of ECO	-0.741	-0.828	5.263*	-0.719
Constant	24.031***	1.686	13.794*	0.101
Clusters	47	53	52	49
Observations	235	265	260	245
R-Square	0.2438	0.1264	0.1704	0.1180
F - test	3.87	1.92	1.32	1.66

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

RESULT 1. *Rejecting hypothesis 1, anxiety negatively affects performance in contests.*

However, the same direction of effects is not observed across treatments.

RESULT 2. *Rejecting hypothesis 2, risk aversion positively affects performance (in the first set of rounds) when the grid task is applied.*

However, the same direction of effects is not observed across treatments.

RESULT 3. *As predicted by hypothesis 3a, male participants achieve higher level of performance than female participants in contests.*

However, male participants achieve higher level of performance than female participants when the slider task is applied. The same direction of effects is not observed across treatments.

Assuming equivalence between effort tasks, I should notice that any kind of direction on performance caused by an anxious mood, risk aversion, gender is replicated in every treatment, regardless the effort task used. Results show that participants' performance is differently affected by the same group of explanatory variables according to the effort task used.

In addition to the variables used so far, I have considered the other seven emotions elicited together with anxiety at the beginning of the round one and at the beginning of round six[35].

Also considering other emotional states and not only anxiety I find that results vary across treatments where the only factor that changes is the effort task applied[36].

The same type of results is provided by the estimates of the Random Effects Model (RE) and Generalized Linear Model (GLM). We find that the RE regression estimates and the GLM estimates converge to OLS regression estimates for all the treatments (see Tables B.11, B.12, B.13, and B.14 for the GLM estimates and Tables B.15, B.16, B.17, and B.18 for the RE estimates in Appendix B.2). For my data an error clustering model seems the most appropriate considering the nature of my independent variables (time-invariant variables). All the models allow us to control for the subjects' level non independence of observations.

[35]Graphs B.7 and B.8 in Appendix B.2 show the percentage of participants' emotional arousal at the beginning of the first round and at the beginning of the sixth round respectively. Graph B.6 in Appendix B.2 shows instead the percentage of participants' emotional arousal at the end of the tenth round.

[36]The sign and the significance of the other variables in the regressions also change among treatments. For example, risk aversion is positive and significant in Treatment C (grid task) while it can be negative in other treatments. The same for Trait-Anxiety whose coefficient in Treatment A (slider task) is negative as well as in Treatment C (grid task) but positive in the other two treatments. However, it is significant only in Treatment A. British nationality has a positive and significant effect only on performance in Treatment C (grid task).

3.5 Discussion

In this chapter I have tried to answer the following question: "does it matter which effort task you use in the economic experimental laboratory?". This research question arises from the methodological discussion about the use of effort tasks in the lab. Effort tasks are widely applied in economic experiments in different research areas: gift-exchange games, contests, sequential labour markets, etc. In my experiments subjects are asked to perform a task where they choose or exert real effort. Hence, the task can be an induced value effort task or a real effort task.

Although real and chosen effort have been implicitly assumed to lead to similar results, several factors might be responsible for differences between the two methods. The introduction of a real effort task can trigger heuristics developed in everyday experience which might not be present when subjects are asked to make a choice between a list of numbers. As pointed out by van Dijk et al. (2001), "[work] involves effort, fatigue, boredom, excitement and other affections not present in the abstract experiments" (p. 189), where a (almost effortless) choice of costs on a convex or linear function is supposed to be a reliable measure for real effort.

In addition, intrinsic utility might be derived from work. There is some evidence that people are more willing to provide voluntary work, than contributing money, for a given purpose (van Dijk et al., 2001). Moreover, results of a real effort task should be seen by subjects under their control and not due to innate abilities or other. Hypothetically real effort tasks minimize the individual differences in abilities and show the real subjects' decision on the level of effort to exert. Real clerical tasks in general allow us to observe how task learning, individual skills, and motivation interact with each other.

On the other hand, the introduction of more realism in lab experiments by implementing real effort tasks comes at the considerable cost of losing some control over important variables. This control is assured by the induced value method.

With an induced value effort task the experimenter can impose the cost of disutility associated with the task and control for subjective costs of decisions. Subjects' outcome is in fact a function of effort, abilities and previous experience. The observation of only outcomes in real effort experiments adds complexity to the analysis of the data and to the possibility of disentangling between those factors in order to actually evaluate the real level of effort exerted.

This work does not aim to promote one technique over another and to claim the advantages of one laboratory method over another. In this work I point out that the choice of an effort task in laboratory experiments should be considered carefully, and most of all, experimenters should acknowledge that their results strongly depend on the type of effort task used, no matter if participants are asked to choose a level of costly effort or to exert real effort. As previously discussed, real effort tasks differ between each other substantially. They can be cognitive or physical tasks (or a combination of both) and their operationalization can be carried out in a completely different way. First of all, it is not clear which proxy can offer a reliable measure of effort (quantity of the output, quality of the output, level of task difficulty chosen, time spent on the task) and second, it is a priori difficult to establish whether subjects will derive any utility from performing the task. While some experimenters believe that the the task should be very easy, mechanical and

> mind-numbing [task] devoid of any intrinsic motivation [..] that participants view as being utterly uninteresting and without any redeeming value (Heyman and Ariely, 2004, p. 790)

others do not exclude that some of the participants might enjoy performing it. Or alternatively, applying a task very easy and without any potential interest could be dangerous as well, in the sense that participants could get bored very quickly and decide to give up the task. Therefore, a task does not have to be boring to minimize the differences in latent ability or intrinsic interest.

In this experiment I used three real effort tasks and an induced value effort task.

The three real effort tasks can be considered relatively easy and mechanical, and somehow all of them can trigger some positive utility in performing them. The allotted time is the same for all the tasks[37] and the same measure of the outcome - the quantity - is applied.

The behavioural hypotheses in such a competitive economic setting refer to emotional arousal, and in particular anxiety, to risk aversion, and to gender effects. These factors can affect subjects' performance. Specifically, I expect that an anxious mood has a positive impact on individuals' performance consistently with the Processing Efficiency Theory which states the motivational function of anxiety. The initial emotional status of worries that might bring to negative consequences on performance and payoffs acts as an incentive to exert more effort. Moreover, performance can be affected by the individual's degree of risk aversion. According to the experimental literature, risk averse agents would exert lower effort. Finally, I expect gender difference in performance: men perform better than women in a competitive setting, according to a well-established experimental economic literature (Croson and Gneezy, 2009; Niederle and Vesterlund, 2011; Gneezy et al., 2003; Gneezy and Rustichini, 2004).

Assuming equivalence between effort tasks, I should notice that any kind of direction on performance caused by an anxious mood, risk aversion, gender (or by other variables considered) is replicated in every treatment, regardless the effort task used.

Results show that participants' performance is differently affected by the same group of explanatory variables according to the effort task used. There is evidence that the mathematical task and the grid task are anxiety inducing: the more the anxiety the worse the performance. On the contrary, the effects of anxiety on individuals' performance take completely different direction when either the slider

[37]It is the same for the induced value effort task as well.

task or the induced value effort task is used. This suggests that the nature of the real effort task strongly characterizes the exertion of effort, in this case inducing anxiety. It is well known in the literature on mood inducement techniques that mathematical tasks are anxiety inducing. Psychologists use this task in the laboratory in order to provoke an anxious mood (Keogh and French, 2001; Shostak and Peterson, 1990; Ashcraft and Kirk, 2001). This work shows that the grid task might bring participants to an anxious status as well as the maths task. The mathematical task, very often used in economic experiments, and the grid task do not appear to be equivalent either to the slider task or to the chosen effort task. Their features might substantially alter subjects' performance.

Contrary to Miller and Pratt (1991) established experimental findings, there is evidence that a high degree of risk aversion positively affects subjects' performance when the grid task is applied. However, the same direction of effects on performance is not observed when the other effort tasks are used. Experiments in the literature have used induced value effort tasks. Therefore, my result might be due once more to the nature of the effort task. For example, when subjects deal with the grid task, the higher the risk aversion the higher their performance. The same conclusion cannot be drawn if I consider the results of the other treatments.

Finally, there is evidence of gender effects in contests, where men do perform better than women, but only in Treatment A where subjects carried out the slider task, confirming the literature findings on the effect of the nature of the task on performance differences between men and women.

Other explanatory variables, as for example alternative emotional status, attempt to explain subjects' performance in contests. The same observation holds: there is no consistent effect of these variables on individuals' performance among treatments.

The effect of anxiety on performance does not depend on the specific measure of anxiety. Other measures of (state and trait) anxiety are in the regressions[38]. The proneness to anxiety (Trait-Anxiety) for example matters only for participants carrying out the slider task. Other emotions have been considered as well. Anger, sadness, despondency and frustration are very significant only in some treatments.

It is important to point out that my discussion on the variables of this experiment does not refer to the magnitude of their impact on participants' level of performance. I am not interested in either measuring the relative effects of each variable through treatments or proposing the advantage of one type of task over another. My claim in this work is that the significance of the independent variables considered on performance strongly depends on the effort task used and that experimental results may be largely driven by the experimental method. Many experimenters present their results and stress the importance of the impact of the factors studied on their main variable without considering the strong role that their chosen effort task has on their results. Their results might have a completely different interpretation if a different effort task is happened to be used.

Surprisingly, no regular check of the robustness of the laboratory results is usually carried out. A single task has typically been implemented in each study and little has been said about the implications of different abilities and skills required in each task. An exception is Brüggen and Strobel (2007): they compared the mathematical task with an induced value effort task in a gift-exchange game. They examine the differences between the effect of levels of wages on individuals' effort in the setting of a chosen effort (induced effort task) and a real effort (real effort task). They find individuals reciprocate to higher level of wages exerting more effort and therefore they conclude that the two laboratory methods bring similar results. However, they carefully highlight the presence of a higher variance under real effort and that the reciprocal behaviour of individuals is understated under chosen effort. Subjects in fact achieve on average a level of performance four times higher when they solve the

[38]Graph B.5 in Appendix B.2 shows the distribution of the scores of State Anxiety. Summary statistics of the scores of State Anxiety are shown in Table B.19 in Appendix B.2 as well.

mathematical task[39].

The claim that the level of effort exerted, and hence the participants' performance, might depend on the relevance of the task to some groups of participants (stereotypes) can arise. This claim look at the extensive research that follows the pioneering study by Steele and Aronson (1995). The authors carried out various investigations on the performance of African Americans on standardized academic tests. They aimed to find out whether the stereotype about Black persons' intellectual ability was relevant to their performance. Results show that Black participants greatly underperformed White participants when Black participants were asked to record their race on a demographic questionnaire filled in before the tests. Steele and Aronson provided strong evidence of the so-called *stereotype threat*: salience of the racial stereotype alone was enough to depress the performance of identified Black students.

Another series of experiments (Steele, 1997) on the stereotype about women's lower maths ability produced the same pattern. Women performed worse than men when they were told that the test aimed to check any gender difference. Steele explained that experimental conditions in which the possibility of gender differences was left to inference (rather than stated directly) did not impair women's performance by triggering doubts about their maths ability. Rather, any performance frustration and impairment came from the awareness of the possible gender-based ability limitation alleged in the stereotype.

In my experiment I can rule out the possibility that performance was affected by any stereotype threat linked to the type of the effort task. I did not make salient, in fact, any group identity and I did not state in the instructions of the experiment any scope related to gender, race, nationality, etc. Moreover, participants were randomly matched in pairs.

[39]The same subjects carry out the induced value effort task. The experiment has a within-subject design.

3.6 Conclusion

Effort tasks are widely used by experimentalists in laboratory economic experiments in order to collect measurements on participants' level of performance or investments. There are many effort tasks developed in the economic literature and they can be categorized in two groups: real effort tasks and induced value effort tasks. The former group refers to those tasks that require a substantial effort to be performed. Induced value effort tasks instead highlight the subject's choice of level of investment (or effort).

There is a wide consensus between experimental economists that while real effort tasks offer more realistic measurements of subjects' exerted effort, induced value effort tasks allow to have full control on the experimental environment and thus on the explanatory variables of the economic setting.

What, on the contrary, does not receive full attention from researchers is the possibility that effort tasks used in the laboratory can substantially affect the experimental results. Generally, real effort tasks are considered to be similar as far as they minimize as much as possible the personal skills involved in the task. Moreover, it is implicitly assumed that real effort tasks and induced value effort tasks lead to similar results.

From the importance that a correct use of the effort tasks has in the experimental methodology my research questions follow: does it matter which effort task you use in the experimental economic laboratory? Are experimental results driven by the type of effort task chosen? Do experimental economics have to account for the significant effect of the effort task on their results?

In order to answer my research questions I ran an economic experiment consisting of four identical treatments expect for the effort task used. In my experiment participants compete in a contest for a prize. Their investments' measure was given by the performance in an effort task that they were asked to carry out for ten rounds. In three treatments I applied three different real effort tasks: the slider task, the maths tasks and the grid task. In the fourth treatment I used an induced value effort tasks.

Theoretically, all the effort tasks are equivalent and they should bring similar results. On the contrary, results differ quite significantly between treatments where everything is identical except the effort task. My main variables, e.g. anxiety, risk aversion, and gender, might or might not significantly affect the provision of effort. The type of effort task used significantly affect subjects' performance.

Since the treatments of this experiment are identical in the procedure and material used except for the effort task applied, then I can conclude that there is no equivalence between effort tasks. The choice of a specific task to be implemented in the economic laboratory raises definitely a crucial methodological issue: the use of these tools affects significantly experimental results and experimentalists should be aware of this issue when they draw their conclusions.

A further experimental investigation can be carried out to explore the potential impact of the use of effort tasks on the experimental results. For example, it would be useful to get a deeper understanding of the implication of the nature of each task. This can be important for the interpretation of the experimental data. Moreover, it might help in finding solutions that overcome the limitations of each effort task.

Chapter 4

Fairness, Expectations and Sabotage: An Experiment on Tournaments

4.1 Introduction

In this chapter we aim to understand (whether and) how procedural fairness in firms and sabotage issues can together cause difficulty in achieving the principal's objectives of productivity and cooperation among workers. Motivating effort and increasing the productivity of the firm has traditionally been the main objective of the principal. Economists have always considered the principal's concern as a matter of providing adequate individual incentives. However, agents' effort is not observable, disutility of work is difficult to estimate for management, and finally agents have an interest to preserve this natural informational asymmetry.

To motivate workers' effort many organizations use relative reward scheme, for example tournaments. In rank-order tournaments agent's payment depends only on the rank of his/her performance and not on either the absolute level of performance or the size of the differences in performance across agents. Tournaments are often used for job promotions and career progression.

However, recent studies in economics have highlighted the issue that in tournaments agents are likely to engage in counter-productive activities, caused by rivalry and hostile behaviour (Bose et al., 2010; Lazear, 1989). The possibility to sabotage a co-worker, even weakly by not being helpful, can negatively affect efficient cooperation and mutual assistance in team work (Auriol et al., 2002; Kräkel, 2005; Drago and Turnbull, 1991).

A cooperative work environment is also affected by procedural fairness, namely the fairness of the set of procedures used to determine outcome relative to workers in the firm. For example, the allocation of bonuses among workers. If the allocation rule applied is perceived as unfair than there can be negative consequences on workers' productivity (Greenberg, 1990; Lind and Tyler, 1988; Greenberg, 1988, 1989). Fairness of the allocation rule can also be linked to contestants' expectations of their opponents' level of effort and sabotage. For example, agents rewarded according to an unfair rule might expect to be victim of sabotage. Because of their high expectations they might bid and/or sabotage more aggressively.

In this chapter we explore the intuition that any procedure applied by the management to define outcomes as the allocation of monetary bonuses might affect agents' behaviour when they compete in tournaments, even though the monetary bonuses are completely unrelated to the prize structure of the tournaments. Agents' behaviour in tournaments and their decisions of effort investment and sabotage should be completely unrelated to the allocation rule used to reward workers. The prize structure of the tournament is in fact fixed and arranged independently of the bonus allocation. Assuming profit maximizing agents, in equilibrium they choose the optimal level of investment and sabotage.

To investigate whether the procedure used to allocate a monetary reward between contestants affects their behaviour in tournament we ran two experiments, Experiment 1 and Experiment 2. Experiment 1 consists of four treatments: Random Allocation 1, Meritocracy, Anti-Meritocracy, and Random Allocation 2. In these treatments participants were ranked according to their individual performance in a initial (real) effort task. This allowed us to divide subjects into two groups: High Type participants, who performed better in the task, and Low Type participant, who performed worse in the task. A High Type participant was thus randomly matched with a Low Type participant. Pairs were then asked to enter a tournament where they could win a prize. Before entering the competition participants received information on their own type, their opponent's type, and the possibility to be given a bonus of an extra £5. Moreover, they were informed about the bonus allocation rule to be applied: randomly between contestants, according to a meritocracy rule (High Type participants received the bonus), or according to an anti-meritocracy rule (Low Type participants received the bonus).

In Random Allocation 1, Meritocracy, Anti-Meritocracy treatments subjects were asked to choose their level of investment[1] and sabotage, whereas in the Random Allocation 2 treatment subjects choose their level of investment only.

In Experiment 2 we incentivised the agents' elicitation of beliefs of opponents' investment and sabotage. In this experiment subjects had the possibility to sabotage.

Results show that Low Type participants were more willing to invest when they were not rewarded and to sabotage their opponents' investment when they were rewarded for their comparatively poor performance. Moreover, their decision of investment and sabotage depended on their expectations of their opponents' level of investment and sabotage.

[1]With investment we mean effort investment.

After reviewing the literature on tournaments and procedural fairness (Section 2) we are going to explain in details our research hypotheses, the experimental design and procedure (Section 3). We will then discuss in two separate sections the results of Experiment 1 and Experiment 2 (Sections 4-5). Section 6 discusses and concludes.

4.2 Related Literature

Rank-order tournaments are contests where agents have the opportunity to expend costly effort in order to affect the probability of winning prizes. Specifically, agents' effort translate into an observable output which depends on the realization of a random variable as well. Prizes are awarded to the contestant with the highest output. Examples of tournaments are examinations, college admission, athletic competitions, and promotions in organizations. Job promotions for example, rank-order tournaments are often applied as incentive schemes based on relative performance. This means that workers are rewarded according to the ranking of their outputs.

The seminal paper in this area is by Lazear and Rosen (1981) who started the formal study of rank-order tournaments in the labour market. In this paper the theoretical properties of internal labour market tournaments were studied. It was shown that wages based upon rank induce the same efficient allocation of resources as an incentive reward scheme based on individual output levels if agents are risk neutral. However, under some circumstances, risk averse workers actually prefer to be paid on the basis of rank. Dynamic aspects of multiple stages, elimination tournaments, number of contestants, distribution of the random variable, and issues of incomplete information have been investigated in later research (Green and Stokey, 1982; Nalebuff and Stiglitz, 1983; Rosen, 1986).

Rank-order tournaments have been extensively investigated in the lab. Bull et al. (1987) did the first laboratory experiment. In this experiment pairs of subjects, whose output is the sum of effort and a uniformly distributed productivity shock, compete for a set of prizes. The aim of this experiment is to test the theoretical prediction of tournament theory and to compare agents' effort in a rank-order tournament with a piece-rate incentive scheme. Experimental results showed that the mean effort levels chosen by subjects converged to their theoretical equilibrium levels. However, a large and robust variance was observed for all rank-order tournament experiments while the variance in the piece rate experiment was quite small. Later experimental findings also largely supported theoretical predictions (Weigelt et al., 1989; Schotter and Weigelt, 1992; Nalbantian and Schotter, 1997; Orrison et al., 2004; Harbring and Irlenbusch, 2003, 2005; Sheremeta and Wu, 2012).

There is evidence in the experimental literature of over-expenditure in contests[2]. Chen (2011) for example observed over-bidding in asymmetric constests, and Krakel and Nieken (2012) found twice as high effort levels than predicted in a tournament with minimum productivity requirements.

A very well established finding in the literature on tournaments is that agents increase their effort in response to an increase in the winner's prize. Sheremeta and Wu (2011) tested the comparative statics predictions of the canonical Lazear and Rosen (1981) tournament theory and consistently with the theory and with other experimental studies (Harbring and Irlenbusch, 2005; Harbring and Lünser, 2008; Falk et al., 2008) found that the incentive effects of prize spreads is significant. The authors also found several empirical puzzles that appear to contradict the canonical theory but that can be explained by a different assumption of the agent's utility function[3].

[2]See Sheremeta and Wu (2012) for a survey.

[3]Specifically, the authors suggest to replace the assumption of a separable agent utility by a non-separable utility function.

4.2.1 Sabotage in Tournaments

Lazear (1989) was the first to recognize the negative implications of sabotaging activities in tournaments for the efficiency properties of the tournament as an incentive device. The option of sabotage may cause poor performance of tournaments particularly if the cost of sabotage[4] is low compared to the cost of the effort that enhances an agent's own performance. In his paper Lazear suggested a counter-measure that can be taken, that is the "pay compression", a reduction in the spread between the payment for the winner and the loser, which typically reduces the incentive to expend effort that aims to sabotage other competitors. However, this also implies a reduction in effort that increases the agent's own performance[5].

A way of reducing sabotage among employees within a firm is to introduce competition from outside: if unsatisfactory performance of a unit or by the firm overall can cause the firm to go out of business than sabotage would represent an additional cost. Therefore, in case of competition for promotion the organization would be better off by appointing someone from outside (Chan, 1996; Chen, 2005).

Shubik (1954) studied the problem of sabotage when contestants are heterogeneous. He considered three shooters of different qualities that can shoot at each other in a randomly determined sequence. Shubik showed that in the case of sequential shooting the most able shooter will not survive with the highest probability. In fact, whoever gets an early chance to eliminate an adversary uses her/his shot to try to eliminate the stronger of the adversaries.

The advantage of eliminating a strong future opponent was explored further in the literature (Chen, 2003; Münster, 2007). If a player is very good at turning effort

[4] Sabotage is a costly activity.

[5] Bose et al. (2010) supported Lazear's 'pay equality' policy by claiming that the principal often gains more from the reduced sabotage than she/he loses from the inability to tailor the reward structure to the individual capabilities of team members. The agents also can gain from the reduced sabotage that an egalitarian reward structure secures.

into performance, the player will prefer to expend a lot of effort on this instead of sabotaging her/his opponents. On the contrary less productive players will sabotage more heavily, especially the most productive player ('the favourite') since she/he is the most dangerous competitor and sabotaging her/him is more effective than sabotaging a weaker player. In this case the opportunity to sabotage has a strongly equalizing property (Münster, 2007). However, the number of contestants has to be sufficiently large so that the total attack is sufficient to make the abler player perform worse than the less able ones (Chen, 2003).

Gürtler and Münster (2010) discussed the sabotage problem among heterogeneous contestants in dynamic tournaments: players have the opportunity to sabotage after each round of the contest. The authors showed that 'favourites' (or abler players) are sabotaged more strongly than 'underdogs' in the final round compared to static tournaments whereas in the first round[6] underdogs are sabotaged more heavily.

Similar sabotaging issues occur in elimination tournaments when players can try to influence the outcome of contest rounds that determine their future competitors. Suppose for example that a contest consists of two rounds, each contestant can decide whether to help the weaker player in the other semi-final round or not. The motivation comes from the opportunity to reduce the chance that the stronger player advances to the final round by helping a weaker player (Amegashie and Runkel, 2006)[7].

Relatively little experimental work has been carried out on sabotage in tournaments[8]. Harbring and Irlenbusch and their co-authors (2005, 2007, 2008, 2011) have

[6]The authors assume two rounds.

[7]Deciding to not help a colleague and potential opponent, and in general being reluctant to behave cooperatively is a weaker form of sabotage. Tournaments are competitive promotion systems that prevent helping efforts and cooperative behaviour in the organization (Auriol et al., 2002; Kräkel, 2005; Drago and Turnbull, 1991).

[8]See Chowduri and Gurlter (2013) for a survey.

done most of the experiments in this area. Experimental results by Harbring and Irlenbusch (2005) showed that sabotage activity increases as the spread between winner and loser prizes widens. Therefore, in the presence of sabotage a principal who seeks to maximize the total effort expended should optimally reduce the spread between winner and loser prizes (as compared to the optimal spread when sabotage is not possible). In their experiment a group of four agents compete in a rank-order tournament and a fifth agent acts as a principal who has to decide the amount of the winner and loser prizes. The principal in this experiment can mitigate the impact of sabotage by chosing an appropriate incentive contract. Similar results were achieved by Falk et al. (2008).

Harbring and Irlenbusch (2008) considered additional aspects of contest designs that could discourage agents from sabotaging each other, such as the number of contestants, and the number of prizes to be awarded. They ran experiments with up to eight participants and they found that increasing the number of winner prizes does have a significant effect on effort: effort is higher when there are as many winner prizes as there are the loser prizes. However, participants did not also increase their sabotage activity. Harbring and Irlenbusch (2011) tested the effects of communication between both contestants and principal on effort and sabotage, which can actually reduce destructive activities if agents agree on the prize structure. They examined the effect of framing as well and they concluded that sabotage is significantly lower when the laboratory situation was presented as an employment context.

Harbring et al. (2007) ran an experiment on asymmetric tournaments where heterogeneous players can mutually sabotage each other. Sabotaging behaviour varies with the composition of the group of contestants: favourites tend to sabotage each other's effort more than they sabotage underdogs but underdogs engage in sabotage with other underdogs as frequently as they do with favourites. Moreover, the authors found that if the identity of the subject that engages in sabotage can be revealed, sabotaging decreases.

Concealing information to contestants as a way to reduce sabotage is considered by Gürtler and Münster (2010) as well. They showed theoretically and experimentally that in a two stage tournament agents that exert more effort in the first stage are sabotaged more than agents that expend less effort in the first stage. Therefore there is less incentive to exert a high level of effort[9]. The problem can be solved by hiding information on contestants' performance.

None of the studies we have discussed so far considers the impact of agent's expectations on opponents' level of effort and sabotage. To our knowledge only Uske (2009) has focused on this aspect. He analysed agents' behaviour with respect to the equilibrium effort levels in tournaments and he observed that agents tend to systematically over-bid (to choose a high level of effort). He explained the observed behaviour through agents' beliefs (whose elicitation was incentivized): agents significantly overestimate their opponents and overreacted to their own beliefs by chosing a high level of investment. In this setting there is no possibility to sabotage and hence there is no discussion about the belief of the opponent's level of sabotage and how this can affect decisions of effort and sabotage as well.

In our study we aim to analyse the impact of different types of procedures through which a monetary bonus is allocated to agents on their decision of investment of effort and sabotage. The bonus is allocated to contestants before entering the competition and it is absolutely independent of the competition itself. We did not find any work in the literature that shows if and how 'procedural fairness' can affect outcomes in tournaments.

4.2.2 Procedural Fairness

The study of procedural justice (or procedural fairness) has a long tradition in psychology and management. The term is used to describe the fairness of the procedures

[9]Choices of effort and sabotage were binary.

used to determine outcome[10] for employees. Together with 'distributive justice', which describes the fairness of the outcome an employee receives, it represents the source of 'organizational justice' which refers to the study of the role of fairness as it directly relates to the workplace (Folger and Greenberg, 1985)[11].

Studies in psychology and management have provided evidence on the relationship between justice perception and work behaviour and have explained how employees' job performance may increase or decrease depending on perceptions of inequitable rewards (Greenberg, 1990; Lind and Tyler, 1988; Greenberg, 1988, 1989). Procedural fairness may be the explanatory variable of employees' behaviour in relation to a variety of work aspects such as participation (Moorman, 1991), budgetary participation, job tension, and interpersonal trust (Lau and Tan, 2006), etc. Lau and Tan (2006) found that procedural fairness has a full intervening effect on the relationship between budgetary participation and job tension and that it is a crucial variable for workers participation and behaviour. The authors claimed that from a practical perspective it is necessary to pay attention to the fairness of the procedures designed and applied in organizations. A work environment that ignores 'procedural fairness concerns run the risk of creating negative organizational attitudes, dissatisfaction with principal's decisions, non-compliance with rules and procedures' (Lind and Tyler, 1988, p. 179).

There is a growing literature in economics on procedural fairness (Hoffman and Spitzer, 1985; Ruffle, 1998; Hoffman et al., 1994; Bolton et al., 2005). These studies demonstrate that different mechanisms that are used to determine the roles in an experiment can have strong impacts on how fair individuals perceive the situation

[10]For example rewards.

[11]One element of organizational justice is the informational justice. The latter is manifested by providing knowledge about procedures that demonstrate regards for people's concerns. Workers are given adequate accounts and explanations of the procedures used to determine desired outcomes (Muchinsky, 2002).

to be and that this can have a strong impact on behaviour.

A recent paper by Ku and Salmon (2012) has considered that perceived procedural unfairness acts as an institutional factor that might make individuals less tolerant of income inequality. In their setting individuals know *ex ante* that a certain growth program proposed by the policy-maker will create income inequality and they are asked about their willingness to support the program. They ran a laboratory experiment in which they varied the mechanism which assigned people to their role (advantaged vs. disadvantaged) and they manipulated the information displayed regarding the characteristics of each player in a pair. After receiving this information disadvantaged players were asked to decide how much of their endowment to pass to the advantaged co-player. The money transferred jointly with the co-player endowment will generate an investment return which will be then share between the two players. The mechanism rules they applied were: random assignment, meritocracy (a mechanism based on performance), arbitrary rule (an assignment rule based on participants' answers of an arbitrary question), and corruption (a mechanism that rewards uncooperative and unethical behaviour). Their experimental results showed that even when the initial positions or roles were randomly assigned, disadvantaged individuals do not choose a transfer amount which will maximize their own income (as well as social efficiency). This confirms the presence of inequity averse preferences, and, particularly compared to the baseline mechanism (random assignment), disadvantaged individuals make choices that are associated with lower efficiency (and lowered inequality) when the other three mechanisms ('meritocracy', 'arbitrary rule', and 'corruption') are applied.

The work by Ku and Salmon (2012) sheds light on the significant impact of procedural fairness on individual (cooperative and non-cooperative) behaviour. However, it investigates a situation in which players are asked to decide how much of their endowment to contribute. It does not consider a competitive setting in which endowment (or final reward) depends on a competition against other individuals. In their experiment there was no opportunity for destructive or hostile activity which can easily occur in real-life situations.

The latter possibility is explored by Zizzo (2004a). In his experiment participants received money by betting and possibly by an arbitrary allocation procedure that aimed to create different perceptions of procedural fairness. Subjects were then asked to choose both to eliminate and to redistribute some money by taking money from others. He found that over 80% of the subjects cared about reducing the endowment of rich subjects at least as much as or more than reducing the endowment of the poorer ones. The way in which subjects perceived the fairness of the procedure significantly affected their decision of taking money from others.

4.3 The Experiments: Research Hypotheses, Experimental Design and Procedure

This research differs from the studies discussed so far since it considers mechanisms of procedural fairness to allocate an outcome (a monetary reward) to participants of our experiments before they enter a competition. In the competition they are asked to choose their level of investment in order to win a prize. Moreover, they are given the opportunity to sabotage their opponent.

The prize of the competition is independent of the allocation rule of the monetary reward and therefore we expect that individual behaviour in competition would not be affected by the allocation rule. However, we do believe that the fairness of the procedure used to assign the monetary reward would significantly drive participants' decision of the level of investment and sabotage in competition. Hence, we ran Experiment 1 in order to explore the effects of the allocation rule of a reward on participants' level of investment and sabotage in tournaments.

4.4 Experiment 1

Experiment 1 was conducted in January 2013 at the Centre for Behavioural and Experimental Social Science (CBESS) at the University of East Anglia. Instructions, control questionnaires and experimental tasks were computerized[12]. The experiments were programmed and conducted with the software z-Tree (Fischbacher, 2007). Overall, 96 subjects took part in 12 sessions. Participants were all undergraduate and postgraduate students of the University of East Anglia. 45% of them were enrolled either in the School of Economics or in the Norwich Business School; 29% of them have British nationality and 38% of participants were male. The average age was 22.

Experiment 1 consists of four treatments: Random Allocation 1, Random Allocation 2, Meritocracy, and Anti-Meritocracy. In Random Allocation 1, Meritocracy, and Anti-Meritocracy subjects take part in a tournament where they are asked to choose simultaneously their level of investment and sabotage (Lazear, 1989). Players in Random Allocation 2, instead, participate in a tournament without sabotage (Lazear and Rosen, 1981) and therefore they choose only their level of investment.

Before entering the contest participants carry out a real effort task. Then they are ranked according to their performance and divided in two types: 1) *High Type* participants that performed better in the real effort task, and 2) *Low Type* participants that performed worse in the real effort task. Subjects receive feedback on their rank and on their opponent's rank. A High Type participant is then randomly matched with a Low Type participant. Information about this kind of matching is displayed on the computer screen: each subject know about her/his own type, and about her/his opponent type. Successively, a bonus of five pounds is assigned according to three different rules:

1. **Random Allocation**: the bonus is assigned randomly either to the Low

[12] Participants had a paper copy of the instructions on their desks as well.

Type contestant or to the High Type contestant;

2. **Meritocracy**: the bonus is assigned to the High Type contestant;

3. **Anti-Meritocracy**: the bonus is assigned to the Low Type contestant.

In the Random Allocation 1 treatment and the Random Allocation 2 treatment a random allocation rule is adopted. In the Meritocracy treatment the bonus is assigned according to a meritocracy rule whereas in the Anti-Meritocracy treatment the bonus is given according to the anti-meritocracy rule.
Table 4.1 shows the experimental design of Experiment 1.

Table 4.1: Experimental Design - Experiment 1

	Bonus randomly assigned	**Bonus assigned to High Type contestants**	**Bonus assigned to Low Type contestants**
Investment and Sabotage	Random Allocation 1	Meritocracy	Anti-Meritocracy
No Sabotage	Random Allocation 2		

The Tournament Game

In the Random Allocation 2 treatment the competition is a two-player rank-order tournament (Lazear and Rosen, 1981). In this type of tournament players are identical, that is they have the same utility function and same cost of investment. In order to win the winner prize a player's outcome must be higher of her/his opponent's outcome. The outcome is defined by the following production function:

$$q_j = e_j + \epsilon_j$$

where e_j is the level of investment or effort, a measure of skill or average output of player j, and ϵ_j is a random component or luck drawn out of a known distribution with zero mean and variance σ^2. Player i has a similar technology to player j and simultaneously makes a decision on her/his level of investment or effort.

In the Random Allocation 1, Meritocracy, and Anti-Meritocracy treatments the competition is a two-player rank-order tournament with sabotage (Lazear, 1989). In this type of tournament players have the opportunity to expend costly effort to reduce the opponent's outcome. Players' production functions are as follows:

$$q_i = f\left(e_i, s_j\right) + \epsilon_i$$

$$q_j = f\left(e_j, s_i\right) + \epsilon_j$$

The term s_j is the j's '**sabotage**' inflicted on the other player, e_j is the level of investment or effort of player j, and ϵ_j is a random component or luck drawn out of a known distribution with zero mean and variance σ^2. Player i has a similar technology to player j and simultaneously makes a decision on her/his level of investment or effort and sabotage.

In each treatment there were two fixed prizes available per each pair of players: W_1 and W_2 where W_1 and $W_2 > 0$ and $W_1 > W_2$. W_1 is the prize that goes to the winner, the 'winner prize', and W_2 is the prize that goes to the loser of the contest, the 'loser prize'.

Investment and sabotage are costly. Each level of investment and sabotage is associated with a cost according to two different predefined cost functions.

In the Random Allocation 1, Meritocracy, and Anti-Meritocracy treatments the costs of the investment and sabotage are deducted from the prize a player receives whereas in the Random Allocation 2 the prize a player gets is reduced by the cost of investment only.

In our experiments participants receive neutrally framed instructions on this task. In the Random Allocation 2 treatment they are asked to choose an integer number between 0 to 100 as their *Number A*, which would represent their investment decision. In the Random Allocation 1, Meritocracy, and Anti-Meritocracy treatments they are also asked to choose a number, this time between 0 and 50, called *Number B*. This latter number represents their level of sabotage. Tables of associated costs are provided. Subjects could practice the task before making their decisions. The task is one-shot game. In Appendix C.1 we summarize the model of tournament without sabotage (Lazear and Rosen, 1981) and the model of tournament with sabotage Lazear (1989), we give the specification of each model and the parametrization used for our experiments. In Appendix C.2 we provide the experimental instructions of the task.

Research Hypotheses

As already mentioned, the prize of the tournament is independent of the allocation rule of the monetary reward. Therefore, we assume that in the tournament individuals would choose the level of investment and sabotage such that their pay-off is maximized. According to the parameterization of our experiment, the optimal level of investment and sabotage in equilibrium is 32 and 9 respectively.

However, receiving (or not) the bonus can induce individuals to change their willingness to invest and sabotage in the competition. For example, individuals who do not receive the bonus would be either more willing to invest in order to not miss the opportunity to get the prize, or less willing to invest because they do not wish to engage in costly activities. Moreover, if the bonus is assigned according to a perceived unfair procedure individuals could have hostile feelings and could be for example more willing to engage in sabotage. The procedural fairness of the assignment of the monetary reward might significantly affect contestants' behaviour in tournament.

We do not state any direction of the effects of the allocation rule of a monetary reward on individuals' behaviour in tournaments. Hence, we formalize our research hypotheses as follows:

HYPOTHESIS 1: *The allocation rule of a bonus does not affect individuals' behaviour in tournaments.*

against

HYPOTHESIS 1a: *The allocation rule of a bonus does affect individuals' behaviour in tournaments.*

Our null hypotheses implies that, assuming profit maximizer individuals, the chosen level of investment should be 32, and the chosen level of sabotage should be 9. These optimal choices would be consistent among treatments and participant's type in our experimental sessions.

Experimental Procedure

Subjects undertake two separate experimental sessions: Session 1 and Session 2. Session 1 is an online session: subjects are asked to complete an online questionnaire before registering themselves to the laboratory sessions which would take place at least one week later. Table 4.2 shows the experimental sessions and tasks. We are going now to illustrate each task.

Session 1

The online questionnaire is divided into three tasks, each visualized in a computer screen. In the first task (Task 1) participants are administered eight items of the *Dispositional Envy Scale* by Smith et al. (1999). This scale gives a measure of the subjects' tendency to feel envy (*Trait Envy*). Agreement or disagreement on each statement is indicated on a 1 to 5 Likert-type scale, in which 1 indicates "*strongly disagree*" and 5 indicates "*strongly agree*". The score range is between 1 and 40. The higher the score the more kin the subject is to feel envy.

In the second task (Task 2) participants are asked to provide some of their demographic information. They answer ten questions about their gender, age, school, nationality, etc. Some of the questions, e.g. "What is your favourite sport?" or

Table 4.2: Experimental Sessions and Tasks - Experiment 1

Session 1 - online session	
Task 1	Dispositional Envy Scale (Trait Envy)
Task 2	Demographic Information (10 items)
Task 3	Social Desirability Scale
Session 2 - laboratory session	
	General instructions
Task 4	Real Effort Task
Task 5	Tournament Instructions
	Practice session - 5 rounds
	Control Questionnaire
	Feedback on performance and matching
	Bonus Allocation
	One-shot contest game
Task 6	Questionnaire
Task 7	"Sabotaging the experimenter" task
Task 8	Risk aversion measure
	Expectations elicitation
	Results and Payment

"How many friend have you got on Facebook?" serve to fill in few minutes and hence to engage the working memory before been asked to complete another psychological questionnaire. The second psychological questionnaire (Task 3) is the *Social Desirability Scale - 17*. This provides a measure of subjects' social desirability[13]. Participants are asked to state whether each item described them with a "True" or "False" answer. The score range is between 0 and 16: the higher the score the higher the subject's social desirability.

The online questionnaire takes only a few minutes and entitles subjects to participate in Session 2, the laboratory session.

[13]Item 4 " *I have tried illegal drugs* " has been dropped from the list of 17 items because of ethical issues.

Session 2

Session 2 is a laboratory session consisting of five successive tasks[14]. The first task is a real effort task by Abeler et al. (2011). We slightly modified their task which consisted of a series of 10 by 15 grids filled in randomly by numbers zero and one. The authors asked subjects to count the number of zero in each grid. No time limit was imposed. In our modified task a 5 by 5 grid appears on the screen and subjects are asked to count the number of ones randomly assigned in the cells[15]. They have five minutes to count correctly as many grids as they can. They are encouraged to do so but they are no given other information about the aim of this task[16]. Per each grid correctly solved participants get one point score. Participants' performance is therefore given by their score at the end of the time allowed.

After completing the real effort task (Task 4) subjects receive the instructions for the second laboratory task (Task 5). This task is a 2-player tournament with two prizes. In the Random Allocation 1, Meritocracy, and Anti-Meritocracy treatments the tournament allows for sabotage while in the Random Allocation 2 treatment there is no sabotage opportunity (we previously discussed these two types of tournaments).

Instructions are followed by five rounds of practice. Each participant is matched with the computer of her/his lab partition. The computer randomly selects its level of investment (and sabotage) and this was known[17]. Feedback on the computer's

[14]Session 2 lasts for approximately 75 minutes.

[15]The task is used in the experiment of the third chapter of this thesis as well.

[16]The task is not incentivised and subjects do not know at this stage how this task would be linked to the successive allocation of the bonus.

[17]Adopting the method of practice round against the computer solves the issue of dealing with the subjects' beliefs on potential opponent's choice of investment and sabotage. Subjects' decisions of investment and sabotage during the real round might be affected by their experience during the practice round. The fact that subjects knew they were playing against a computer, and that the computer would randomly select the level of investment and sabotage, plausibly reduces the likelihood of the experience (in terms of outcome of the tournament game) affecting subjects' later

selections was given at the end of each round. The control questionnaire is another opportunity to practice the task as well.

Successively, participants are ranked according to their performance in Task 4 (the real effort task). If there are two or more subjects with the same performance score the number of attempted grids is considered: per each attempted grid subjects receive a point score. Hence, the group of participants in the experimental session is divided in two subgroups: 1) *High Type* subgroup, consisting of participants whose performance is in the top half of the group, that is participants that performed better the (real effort) task, and 2) *Low Type* subgroup, consisting of participants whose performance is in the bottom half of the group, that is participants that performed worse the real effort task.

Hence, one participant of the High Type subgroup is randomly matched with another participant of the Low Type subgroup. Participants receive information on their own type and their opponent's type on their computer screen.

Finally, subjects read on the screen the possibility to get a reward of five pounds (bonus) and the allocation rule according to which the bonus is assigned to one of the two contestants: 1) randomly - Random Allocation rule; 2) to the High Type subgroup - Meritocracy rule; or 3) to the Low Type subgroup - Anti-meritocracy rule.

For example, in the Meritocracy treatment, participants read on their computer screen:

> The experimenter has ranked the performance of the participants in this room in the task in which you were asked to count the number of ones in the grids.

real round. At any rate, insofar as it did affect later behaviour, it would just be in the form of introducing 'white noise' in the data. The practice round should help their understanding in the calculation of their final payoff.

> Your performance is in the top half of the group and your co-participant's performance is in the bottom half of the group. You performed better than your co-participant.
>
> The experimenter has decided TO ASSIGN A BONUS OF 5 POUNDS to participants that performed BETTER in the task.
> You have been assigned the bonus. Your co-participant has not been assigned the bonus.

At this stage participants are asked to make their decisions of investment (and sabotage if they are participating in Random Allocation 1, Meritocracy, and Anti-Meritocracy treatments) in a one-shot tournament game.

Task 6, the third task of the laboratory session, is a questionnaire which aims at identifying the current emotional state of participants, their perception of fairness and overall their feelings in Task 5 (the tournament). In particular, items were framed in such a way to catch the potential feeling of envy and unfairness of High Type participants when an unfair anti-meritocracy rule is applied. The questionnaire joins together various questionnaires (or part of them) on envy, fairness, hostile feelings, etc. Specifically, the first nine items measure *episodic envy* by Cohen-Charash (2009). We created items 10-14 to measure whether envy is caused by low warmth towards the opponent and high competence of the opponent according to the *Warmth and Competence Model* by Fiske et al. (2002). Items 15-17 measure subjects' *competitive feelings*: these three items of the original questionnaire Cohen-Charash (2009) have been adapted. From Smith et al. (1999) items have been taken and adapted to measure *subjective injustice belief* (items 18-22), *inferior beliefs* (items 23-25), *hostile and depressive feelings* (items 26-32), and *objective injustice* (items 33-35). We created item 36 to measure the joy of competition while items 37-48 are adapted statements used to measure the current arousal of twelve emotions (*anxiety, anger, envy, irritation, jealousy, surprise, happiness, contempt, sadness, fear, joy,* *and shame*), a list adopted by economists (Bosman and Van Winden, 2002; Bolle et al., 2010).

In total the questionnaire consisted of the 48 items just described. Participants are asked to scroll a bar whose range is between 0 (not at all) and 100 (very much) to indicate whether each item of the questionnaire describes them or not. See Appendix C.3 for the list of the items of the questionnaire. Since the questionnaire is administered after the tournament task, that is after decisions of investment and sabotage are made, and since the frame of the items is the same for both High Type and Low Type participants, results of the questionnaire might not be strongly informative. We discuss the results of the questionnaire in Appendix C.4.

We named Task 7, the fourth task of the laboratory session, "*Sabotaging the experimenter task*". Participants had the possibility either to take from or to contribute to the experimenter's pot of extra money up to two pounds. They expressed their choice dragging a slider towards right if they wished to contribute or towards left if they wished to take. Amounts of money were expressed in 2 decimal places. If the subject decided to contribute to the experimenter's pot, the amount of money indicated by the subject was deducted by her/his final earnings. If the subject decided to take from the experimenter's pot, the amount of money indicated by the subject was added to her/his final earnings.

We introduced this task in our experiment because we understand the possibility that participants might have hostile feelings towards the experimenter who decided the allocation rule to be applied in each experimental session. This additional task has an explorative nature. We discuss it and the results in detail in Appendix C.5.

In the final task, Task 8, participants completed the Holt-Laury Questionnaire[18], an incentive compatible lottery widely used to measure subjects' risk aversion.

[18] Holt and Laury (2002).

Finally, at the end of the session (after Task 8 and before the payment screen) we asked subjects to elicit their beliefs on their opponent's level of investment and sabotage. This question was not incentivised. Information on the payment screen included the outcome of the tournament, the outcome of the lottery, and hence the final earnings. There was no participation fee. Table 4.3 indicates our sample size per treatment.

Table 4.3: Sample Size - Experiment 1

Random Allocation 1	Meritocracy	Anti-Meritocracy	Random Allocation 2	**Total**
26	22	24	24	**96**

4.4.1 Experiment 1: Results

Investment and Sabotage

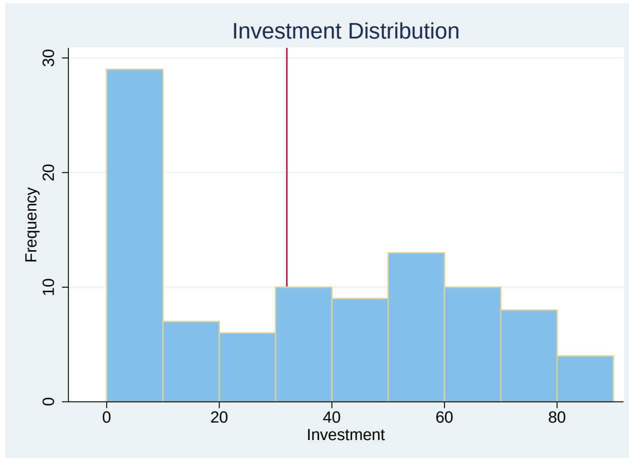
Overall, the average level of investment is 33, very close to the Nash equilibrium[19]. However, about 23% of participants chose 0 as level of investment and 40% of participants chose a level of investment greater than 45. 25% of subjects decided an investment level greater than 57. The maximum investment is 90 (Figure 4.1).

The average sabotage level is 12.7 and the median is 10. About 35% of participants chose to not sabotage (sabotage = 0). The top 25% of subjects chose a level of sabotage greater than 24. The maximum level of sabotage is 50 (Figure 4.2).

Table 4.4 shows the average level of investment and sabotage by type. High Type participants invest less than Low Type participants only in the Meritocracy treatment compared to the other treatments: the average level of investment of High

[19]We remind that according to our parametrization of the tournament game the optimal investment level is 32 and the optimal level of sabotage is 9.

Figure 4.1: Investment Distribution



Type subjects is 24.818 while the average level of investment of Low Type subjects is 49.364[20]. The difference is statistically significant (Wilcoxon rank-sum test, p = 0.073)[21]. Low Type participants invest on average more in the Meritocracy treatment (49.364) compared to the Random Allocation 1 treatment (22.154) and this difference is statistically significant (Wilcoxon rank-sum test, p = 0.042). Also, on average, Low Type participants invest more in the Meritocracy treatment compared to the Random Allocation 2 treatment where the bonus is randomly allocated and sabotage is not allowed (Wilcoxon rank-sum test, p = 0.102). There is no significant difference between treatments for sabotage decisions[22]. See Table C.6 in Appendix C.7 for descriptive statistics of sabotage by Type.

Regression Analysis on Investment and Sabotage

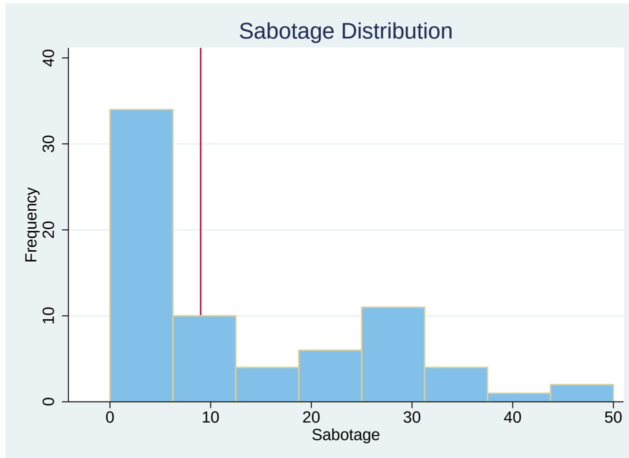
Decisions of investment and sabotage are made simultaneously (in particular,

[20]Table C.5 in Appendix C.7 shows descriptive statistics of investment by type.

[21]This test and onwards are 2-sample tests.

[22]However, if we consider sabotage by type, Low Type participants sabotage more in the Anti-Meritocracy treatment than Low Type participants in the Meritocracy treatment as a t test shows (p = 0.097).

Figure 4.2: Sabotage Distribution



subjects' enter their desired level of investment and sabotage in boxes appearing in the same computer screen). Moreover, we have seen that investment and sabotage are censored variables[23]. Therefore, we used bivariate Tobit regressions in order to analyse the impact of various independent variables on subjects' decisions on investment and sabotage.

Table 4.5 shows the estimates of four Bivariate Tobit regressions on investment and sabotage (Model 1, 2, 3 and 4). The explanatory variables include participants' type (Low Type = 1 if the subject belongs to the Low Type subgroup), a bonus dummy (Bonus = 1 if the subject has received the bonus), the score of risk aversion, participants' expectations of investment (Exp. Inv.) and sabotage (Exp. Sab), and some interaction terms between participants' type and bonus, and between participants' type and expectations of investment and sabotage. Treatments dummies are included as well: Meritocracy and Anti-Meritocracy. We will discuss the coefficients of these treatments dummies in comparison with the Random Allocation 1 treatment, which is considered our baseline treatment. We do not include data

[23]Many participants choose to not invest or to not sabotage at all. Therefore, there is a left censoring.

Table 4.4: Average Investment and Sabotage by Type - Experiment 1

	Investment		Sabotage	
	High Type	Low Type	High Type	Low Type
Overall	32.791	33.625	10.972	14.416
Random Allocation 1	28	22.153	8.461	13.846
Meritocracy	24.818	49.363	12.909	10.636
Anti-Meritocracy	37.916	33.333	11.916	18.5
Random Allocation 2	40.166	28.916	.	.

of the Random Allocation 2 treatment in which there is no opportunity to sabotage[24].

There is strong evidence that participants invest significantly more in the Meritocracy treatment and in the Anti-Meritocracy treatment than in the baseline treatment, and that they are less likely to invest if they receive the bonus. Moreover, their investment is lower the higher their risk aversion.

There is evidence that participants who have high expectation of their opponents' investment decide to invest more. However, if we introduce in the regression the interaction term 'Low Type x Exp. Sab.' expectation of investment that does not affect significantly the investment decisions anymore (Model 3). It seems that the participant's type has no effect on investment decisions.

Participants' type does not affect sabotage decisions as well, unless we consider the interaction terms between type and expectations: Low Type are less likely to sabotage than High Type (Model 3 and 4). However, if they receive the bonus they are more willing to sabotage their opponent (Model 2, 3 and 4). Finally, there is

[24] We will discuss separately participants' decision of investment in the Random Allocation 2 treatment.

Table 4.5: Bivariate Tobit Estimates on Investment and Sabotage with Expectations

	Model 1	Model 2	Model 3	Model 4
Investment				
Low Type	6.277	-11.738	-21.237	-10.150
Meritocracy	27.498***	38.790***	40.643***	38.705***
Anti-Meritocracy	29.748***	22.256**	22.680**	22.209**
Bonus	-17.623**	-36.717**	-38.097**	-36.781**
Risk Aversion	-6.457***	-5.898***	-6.013***	-5.824***
Exp. Inv.	0.368***	0.348***	0.292	0.363**
Exp. Sab.	-0.001	0.027	-0.006	0.030
Low Type x Bonus		37.125	38.845	37.097
Low Type x Exp. Inv.			0.124	-0.030
Low Type x Exp. Sab.			0.104	
Constant	34.094**	38.289***	42.123***	37.070**
Sabotage				
Low Type	2.286	-14.518	-34.859***	-25.123**
Meritocracy	8.899	19.954**	22.632***	20.934**
Anti-Meritocracy	13.772**	7.386	8.081	7.921
Bonus	0.108	-17.523*	-19.768**	-18.066*
Risk Aversion	-3.481***	-3.017***	-3.330***	-3.102***
Exp. Inv.	-0.101	-0.118	-0.234*	-0.069
Exp. Sab.	0.304*	0.327**	0.241	0.022
Low Type x Bonus		34.376**	36.718**	33.908**
Low Type x Exp. Inv.			0.270	
Low Type x Exp. Sab.			0.230	0.461*
Constant	17.093**	20.808**	30.020***	25.292***
sigma1				
Constant	29.163***	28.847***	28.770***	28.961***
sigma2				
Constant	17.621***	17.272***	16.265***	16.634***
rho				
Constant	1.363***	1.301***	1.272***	1.274***
Observations	72	72	72	72
Wald test	34.69***	35.08***	35.74***	34.82***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Exp. Inv.' = Expectation of Investment;

'Exp. Sab.' = Expectation of Sabotage.

strong evidence that the higher the risk aversion the lower the sabotage[25].

We can summarize our results discussed so far as follows:

RESULT 1: *As predicted by hypothesis 1a, the allocation rule of a bonus does affect individuals' behaviour in tournaments: subjects that receive the bonus are less likely to invest; investment is overall lower when a monetary bonus is randomly allocated. The allocation of the bonus affects the sabotage decision as well: Low Type participants sabotage more if they receive the bonus.*

Subjects' Behaviour and Nash Equilibrium

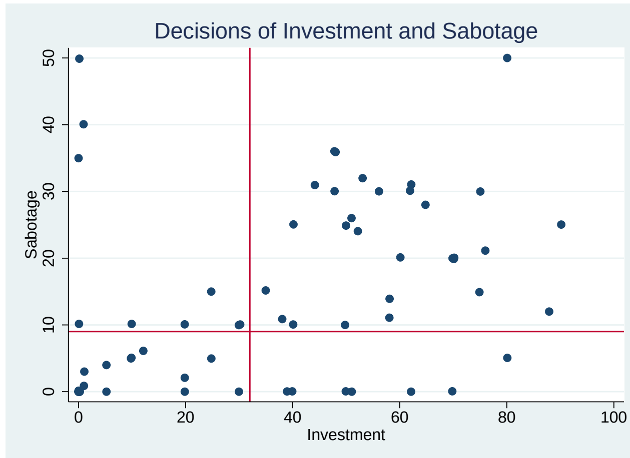
In Graph 4.3 we plot the decisions of investment and sabotage which are made simultaneously. It seems there is a positive correlation between the two decisions (Spearman correlation coefficient = 0.542, $p < 0.001$)[26]. We can observe as well that many subjects decided to not sabotage at all but to invest a certain amount, and some participants did not invest at all but they sabotage up to the maximum level[27]. Overall, 22% of participants chose to not invest and simultaneously to not sabotage, 26% of them chose to not invest but to sabotage up to the maximum level (3% chose sabotage = 50). 35% of participants chose to not sabotage but to invest up to 90.

[25]We introduce demographic variables in the models 13-15 (see Table C.7 in Appendix C.7). Students enrolled in the School of Economics are less likely to invest compared to students enrolled in other schools. The other demographic variables do not have any significant effects on investment and sabotage decisions. Finally, we include in our regression analysis the social desirability score. While this variable has no effect on investment's decisions, it is strongly significant for participants' choice of sabotage (coefficients have negative sign). See models 16-18 for sabotage in Table C.8 in Appendix C.7.

[26]The correlation coefficient between these two variables is indeed positive. See Table C.9 in Appendix C.7.

[27]Graphs C.8, C.9, and C.10 in Appendix C.7 plot the investment and sabotage decisions by participants' type, in the Meritocracy treatment, and in the Anti-Meritocracy treatment respectively.

Figure 4.3: Decisions of Investment and Sabotage - Overall



In the graph the lines indicate the optimal level of investment (32) and the optimal level of sabotage (9), according to the parameterization of the tournament model chosen for this experiment. High Type subjects invest significantly less than the optimal value when they receive the bonus (t test at 10% level of significance, p = 0.083) and significantly more than the optimal value if they do not receive the bonus (t test at 10% level of significance, p = 0.085).

The difference between the observed average sabotage and the optimal sabotage value is significant in the Anti-Meritocracy treatment (t test, p = 0.018)[28]. In particular, Low Type subjects sabotage significantly more than the optimal value in the Anti-Meritocracy treatment (t test, p = 0.028)[29]. High Type subjects' sabotage is not significantly different from the (Nash) equilibrium level.

Regression Estimates on Investment - Random Allocation 2 treatment

[28] Sign test, p = 0.032.

[29] Sign test, p = 0.073.

In the Random Allocation 2 treatment subjects do not have the opportunity to sabotage. They are asked to decide their level of investment only. To compare participants' decisions of the level of investment between treatments we run a separate regression analysis with data on investment only. We apply a Tobit regression model because data are left-censored. Estimates on investment are shown in Table 4.6. In models 5-8 we compared subjects' investment decisions of the Random Allocation 1 treatment, the Meritocracy treatment and the Anti-Meritocracy treatment with the Random Allocation 2 treatment. We introduce a dummy for participants' type (Low Type = 1 if the subject belongs to the Low Type subgroup), a bonus dummy (Bonus = 1 if the subject has received the bonus), the score of risk aversion, participants' expectations of investment (Exp. Inv.), the social desirability score, and an interaction term between participants' type and bonus. In model 8 we consider also some demographic variables: gender (Male = 1 if the participant is a man), age, nationality (British Nationality = 1 if the participant is British), and school (School of ECO = 1 is the participant is enrolled in the School of Economics).

While we do not have a clear direction of investment decisions in the Meritocracy and Anti-Meritocracy treatments compared to the Random Allocation 2 treatment, there is strong evidence that in the Random Allocation 1 treatment subjects are significantly less likely to invest compared to the Random Allocation 2 treatment. Since the two treatments differ between each other only for the possibility to sabotage, we can conclude that investment is overall lower when there are opportunities to sabotage and the bonus is randomly allocated.

As in previous regressions, bonus and risk aversion variables have negative and significant coefficients: participants that receive the bonus are less likely to invest and the more risk averse the subject is the lower the investment. Expectations of investment positively and significantly affect the investment decision.

Table 4.6: Tobit Estimates on Investment Data Only

	Model 5	Model 6	Model 7	Model 8
Low Type	-9.252	-10.310	-5.116	-4.869
Random Allocation 1	-20.136**	-20.141**	-20.485**	-19.992**
Meritocracy	10.468	7.284	7.324	3.221
Anti-Meritocracy	-0.956	-1.147	0.819	3.468
Bonus	-25.652**	-24.076**	-21.890*	-20.115*
Risk Aversion	-4.628***	-4.846***	-4.211**	-4.614**
Low Type x Bonus	26.589	27.131	17.875	15.156
Social Desirability Score		-0.766		
Exp. Inv.			0.341***	0.340***
Male				0.645
Age				0.121
British Nationality				0.454
School of ECO				-9.702
Constant	68.790***	76.976***	48.837***	51.788**
sigma				
Constant	31.117***	30.752***	29.147***	28.614***
Observations	96	95	96	95
Left-censored observations	22	22	22	22
Right-censored observations	0	0	0	0
Likelihood ratio test (LR)	17.56**	17.04**	27.93***	29.62**

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Exp. Inv.' = Expectations of Investment.

4.4.2 Experiment 1: Discussion

In this experiment we explored the effects of three allocation procedure of monetary rewards on individual decisions of investment and sabotage in tournaments. The monetary reward that we called 'bonus' has been allocated according to a meritocracy rule in the Meritocracy treatment, according to a non-meritocracy rule in the Anti-Meritocracy treatment, and finally randomly between contestants in Random Allocation 1 and Random Allocation 2 treatments. Participants' performance score refers to how they carry out a task prior entering the contest. The contest is a two-player tournament with 2 fixed prizes. The outcome of this contest is designed to be independent of the bonus. This implies that individual behaviour in the tournament would be not affected by the allocation of the bonus and that individuals would choose an optimal level of effort and sabotage assuming they are profit-maximizers.

As predicted by our alternative hypothesis, the allocation of the bonus before the tournament does actually affect contestants' behaviour. Individual behaviour varies across treatments and type. For example, Low Type participants invest more when the bonus is allocated according to a meritocracy rule than a random rule. Overall, we find that investment is lower when the random allocation rule is applied. When on the contrary a meritocracy (or anti-meritocracy) rule is followed the average level of investment is higher. Moreover, we find that subjects that receive the bonus are less likely to invest effort in the tournament. For example, we have seen in Table 4.4 that High Type participants invest less in the Meritocracy treatment, that is when they are rewarded according to a fair rule, than in other treatments. On the other hand, Low Type participants that do not receive the bonus (when a meritocracy rule is applied) are more willing to invest.

Regarding the participants' choice of the level of sabotage, we observe a very interesting effect of the bonus assignment: there is consistent and strong evidence that Low Type participants sabotage more when they receive the bonus. In particular, we have seen that they sabotage significantly more than the optimal level when the anti-meritocracy rule is applied. Expectations of opponents' behaviour and decisions might explain Low Type participants choice of destructive effort. We notice in fact in Table 4.5 (Model 4) that the coefficient of the interaction term 'Low Type x Exp. Sab.' is significant: Low Type participants with high expectations of sabotage they are more likely to sabotage. Interestingly, beliefs of investments do not affect Low Type participants sabotage decisions but beliefs on sabotage induce them to sabotage more when they expect to be sabotaged.

Although the coefficient of the interaction term 'Low Type x Exp. Sab.' in Model 4 is significant, the coefficient of the same interaction term in Model 3 is not significant. This puzzling result might be due either to noise in the data or to the elicitation procedure of participants' expectations. To better understand whether the expectations of Low Type participants play a key role in their decisions to sabotage, we decided to run another experiment, Experiment 2. In this experiment the elicitation of the subjects' expectations occurs before entering the competition instead of at the very end of the experimental session. Moreover, we introduced a monetary incentivation of the elicitation that would induce participants to make their best guess on their opponent's level of investment and sabotage. In this way we test any effects of participants' expectation on their decisions of investment and sabotage. Specifically, we will investigate whether Low Type participants are affected by their expectation of sabotage when they decide their level of destructive effort.

4.5 Experiment 2

In this experiment we test any effects of participants' expectation on their decisions of investment and sabotage and we will investigate whether Low Type participants are affected by their expectation of sabotage when they decide their level of destructive effort. Specifically, our hypothesis is as follows:

HYPOTHESIS 2: *Participants' expectations of their opponent's level of investment and sabotage affect participants' decisions of investment and sabotage.*

Experiment 2 was run in February 2013 at the Centre for Behavioural and Experimental Social Science (CBESS) at the University of East Anglia. Instructions, control questionnaires and experimental tasks were computerized[30]. The experiments were programmed and conducted with the software z-Tree (Fischbacher, 2007). 72 subjects took part in 9 sessions. Participants were all undergraduate and postgraduate students of the University of East Anglia. 41% of them were enrolled either in the School of Economics or in the Norwich Business School; 27% of them have British nationality and 34% of participants were male. The average age was 22 as for Experiment 1.

Experiment 2 consists of two treatments: Meritocracy and Anti-Meritocracy. In both treatments participants in are asked to take part in a tournament with sabotage (Lazear, 1989). In the Meritocracy treatment the meritocracy rule is applied, whereas in the Anti-Meritocracy treatment the bonus is allocated according to the anti-meritocracy rule. Table 4.7 shows the experimental design.

Table 4.7: Experimental Design - Experiment 2

	Bonus assigned to High Type contestants	**Bonus assigned to Low Type contestants**
Investment and Sabotage	Meritocracy	Anti-Meritocracy

The experimental sessions follow a similar procedure as in Experiment 1 (see Table 4.8). There are two separate experimental sessions over time: Session 1, the online session, and Session 2, the laboratory session. Session 1 of this experiment is identical to Session 1 of Experiment 1 in the tasks and procedure applied.

[30]Participants had a paper copy of the instructions on their desks as well.

Session 2 consists of the same five successive tasks of Session 2 of Experiment 1. However, in this experiment the elicitation of the participants' expectations of investment and sabotage occurs before the tournament game. Before deciding their level of investment and sabotage, subjects are asked to write down their best guess on their opponent's level of investment and sabotage. We call this stage "Guessing stage". Differently from Experiment 1, we incentivise the subjects' elicitation of beliefs. We applied the following linear incentive scheme for both investment and sabotage:

$$Earnings = \begin{cases} \text{P} & \text{if } g = t \\ \text{P - a*abs(g-t)} & \text{if } g \neq t \end{cases} \tag{4.1}$$

where g is the subject's guess and t is the true value. P is the prize subjects can get in this task. P is equal to £4. The parameter a is equal to 40 pence for guesses on investments and equal to 80 pence for guesses on sabotage. This means that if a participant can guess exactly the level of investment (or sabotage) of her/his opponent she/he gets £4. Per each point that the guess is 'out' a deduction of 40 pence (or 80 pence if the guess is on sabotage) is made from the prize, with a minimum of £0 earned. Only one randomly chosen guess (either on investment or on sabotage) was paid[31].

Task 6 is slightly different from Task 6 in Experiment 1: we add 14 items to the questionnaire. We create six items (49-54) to measure participant's expectation of her/his opponent's hostile feelings. Items 55 and 56 are added for an additional measure of objective injustice beliefs, and items 57 and 58 in order to supplement the measure of envy according to the Warmth and Competence Model. Items 59 and 60 are added to items for subjective injustice beliefs. Expectations of investment and sabotage are measured by item 61 and 62 respectively. Finally, items 63 and 64 provide a participants' response in terms of investment and sabotage after having

[31] See Appendix C.6 for experimental instructions on the guessing stage.

Table 4.8: Experimental Sessions and Tasks - Experiment 2

Session 1 - online session	
Task 1	Dispositional Envy Scale (Trait Envy)
Task 2	Demographic Information (10 items)
Task 3	Social Desirability Scale
Session 2 - laboratory session	
	General instructions
Task 4	Real Effort Task
Task 5	Tournament Instructions
	Practice session - 5 rounds
	Control Questionnaire
	Feedback on performance and matching
	Bonus Allocation
	Expectations elicitation
	One-shot contest game
Task 6	Questionnaire
Task 7	"Sabotaging the experimenter" task
Task 8	Risk aversion measure
	Two final questions
	Results and Payment

received information on the allocation of the bonus[32]. See Appendix C.3 for the items of the questionnaire and Appendix C.4 for the discussion of the participants' answers.

After Task 8, the Holt-Laury lottery for the risk aversion measure, and before the payment screen, participants are asked two final questions: they are asked to motivate their choice of the level of investment and the level of sabotage. Subjects could read on two successive screens the following questions 'You chose Number *[...]* as Number A. Why did you choose this number?' (first screen), and 'You chose Number *[...]* as Number B. Why did you choose this number?' (second screen)[33]. Participants provide their answers in a chat box[34].

[32]Items 59-64 are created ad hoc as well.

[33]Number A indicates the level of investment and Number B indicates the level of sabotage.

[34]A chat box is a tool of z-Tree program that allows to insert text.

Session 2 lasted for approximately 90 minutes; 36 participants took part to each treatment. Therefore, a total of 72 subjects participated in Experiment 2 (Table 4.9).

Table 4.9: Sample Size - Experiment 2

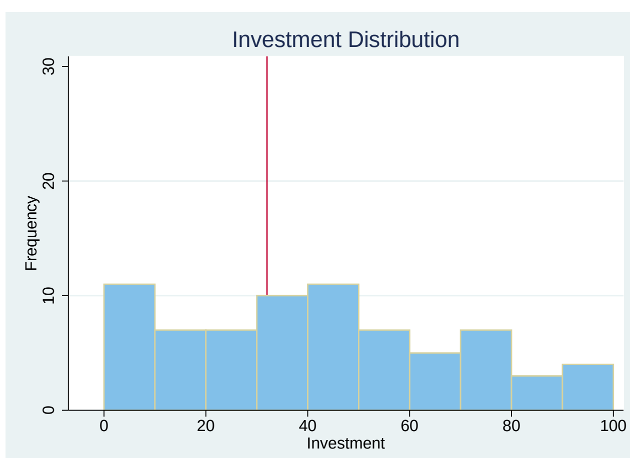
Meritocracy	Anti-Meritocracy	**Total**
36	36	**72**

4.5.1 Experiment 2: Results

Investment and Sabotage

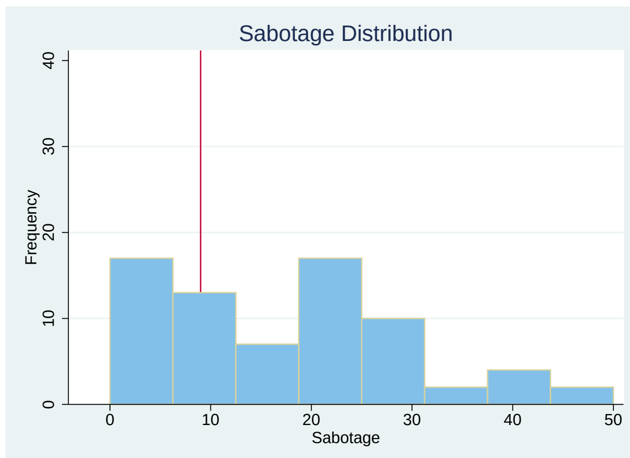
About 11% of participants in Study II chose 0 as level of investment. The maximum investment level was 100 (4%). The average level of investment was 40.29, higher than the equilibrium level. 10% chose to invest more than 78 (see Figure 4.4).

Figure 4.4: Investment Distribution



About 17% of participants chose to not sabotage at all (sabotage = 0), 25% chose a level of sabotage greater than 24 and the maximum level of sabotage was 50 (nearly 3% of participants chose the maximum level). Mostly subjects sabotage an amount between 20 and 24 (see Figure 4.5). The average level of sabotage was 17.04 and the median was 15.5.

Figure 4.5: Sabotage Distribution



There is no significant difference between the average investments in the Meritocracy treatment (39.5) and in the Anti-Meritocracy treatment (41.08)[35]. However, if we consider the average investment by subjects' type we can see (Table 4.10) that Low Type participants invest significantly more than High Type participants overall (Wilcoxon rank-sum test, p = 0.004) and in the Anti-Meritocracy treatment (Wilcoxon rank-sum test, p = 0.014). The amount invested by either High Type or Low Type participants is not significantly different across treatments[36].

The average sabotage in the Meritocracy treatment is 18.139, very close to the average sabotage in the Anti-Meritocracy treatment where the average sabotage is 15.944[37]. However, the average sabotage of Low Type participants is significantly

[35] See Table C.10 in for descriptive statistics on investment by treatments.

[36] See Table C.11 in Appendix C.7 for descriptive statistics of investment by type.

[37] See Table C.12 for average sabotage between treatments and Table C.13 for descriptive statistics of sabotage by participants' type.

Table 4.10: Average Investment and Sabotage by Type - Experiment 2

	Investment		Sabotage	
	High Type	Low Type	High Type	Low Type
Overall	30.666	49.916	14.111	19.972
Meritocracy	31.388	47.611	15.5	20.777
Anti-Meritocracy	29.944	52.222	12.722	19.166

higher than the average sabotage of High Type participants (Table 4.10). Low Type participants sabotage overall significantly more than High Type participants (Wilcoxon rank-sum test, $p = 0.103$)[38].

Expectations of Investment and Sabotage

On average participants expect that their opponent would invest 48.15, 4% of them expect that their opponent would not invest at all and nearly 21% of them expect that their opponent would invest 50 (the median value)[39]. Low Type participants that receive the bonus expect that on average their opponent would invest 48.22, while Low Type participants that did not receive the bonus expect that on average their opponent would invest 55.22. This difference is not significant. High Type participants that receive the bonus expect that their opponent would invest 43 and High Type participants that do not receive the bonus expect that their opponent would invest 46.17. There is no significant difference between expectations[40].

The average expectation of sabotage is 21.53. Nearly 10% of participants expect that their opponent would not sabotage at all (sabotage = 0), and about 14% of

[38] t test, $p = 0.068$.

[39] Only one participant out of 72 expects that their opponent would invest the maximum (100).

[40] Figures C.11, C.12, and C.13 in Appendix C.7 show the distribution of expectations of investment, overall, for Low Type participants, and for High Type participants respectively.

subjects expect that level of sabotage to be equal to 25 (the median value)[41].

The expectation of sabotage of Low Type participants, whether or not they receive the bonus, is approximately the same (if they receive the bonus expectation = 22.44, if not expectation = 23.11). However, there is some evidence of a significant difference between the expectation of sabotage of High Type participants that receive the bonus (24.11) and the expectation of sabotage of High Type participants who do not receive the bonus (16.44). The p-value of the Wilcoxon-Mann-Whitney rank-sum test is 0.068[42].

Regression Analysis on Investment and Sabotage

Table 4.11 shows the estimates of bivariate Tobit regressions on investment and sabotage[43]. We compared subjects' investment decisions of the Meritocracy treatment with the Anti-Meritocracy treatment. We introduce a dummy for participants' type (Low Type = 1 if the subject belongs to the Low Type subgroup), a bonus dummy (Bonus = 1 if the subject has received the bonus), the score of risk aversion, participants' expectations of investment (Exp. Inv.) and sabotage (Exp. Sab.), and interaction terms between participants' type and expectations.

There is strong evidence that the higher the expectation of opponents' level of investment the higher the subject's investment. The expectations of sabotage do not matter on subjects' investment decisions.

RESULT 2. *As predicted by hypothesis 2, participants' expectations of their opponent's level of investment affect participants' decisions of investment. The higher participants' expectation of their opponent's investment the higher their investment.*

The interaction term "Low Type x Expectation of Sabotage" is strongly significant

[41] Three participants out 72 expect that their opponent would sabotage the maximum (50).

[42] Figures C.5, C.6, and C.7 in Appendix C.7 show the distribution of expectations of sabotage, overall, for Low Type participants, and for High Type participants respectively.

[43] Decisions of investment and sabotage are made simultaneously. Many participants choose to not invest or to not sabotage. Therefore there is a left censoring.

Table 4.11: Bivariate Tobit Estimates on Investment and Sabotage with Expectations - Experiment 2

	Model 9	Model 10	Model 11	Model 12
Investment				
Low Type	19.676***	15.945***	-23.153	-15.594
Anti-Meritocracy	-1.746	-0.082	-0.248	1.965
Bonus	4.454	8.274	10.141*	8.119
Risk Aversion	-2.121	-0.894	-0.347	-0.744
Exp. Inv.		0.739***	0.536**	0.369*
Exp. Sab.		-0.092	-0.546	-0.006
Low Type x Exp. Inv.			0.414	0.647***
Low Type x Exp. Sab.			0.884*	
Constant	38.776***	-2.461	12.230	11.077
Sabotage				
Low Type	6.700*	4.127	-11.952*	-11.268**
Anti-Meritocracy	-3.860	-0.989	-2.758	-1.980
Bonus	1.756	-0.857	1.199	0.553
Risk Aversion	-0.157	-0.737	-0.255	-0.367
Exp. Inv.		-0.028	0.072	0.001
Exp. Sab.		0.688***	0.132	0.296*
Low Type x Exp. Inv.			-0.105	
Low Type x Exp. Sab.			0.974***	0.706***
Constant	14.180**	5.218	9.333	9.748*
sigma1				
Constant	28.820***	24.526***	23.016***	23.389***
sigma2				
Constant	14.527***	11.801***	10.454***	10.572***
rho				
Constant	1.112***	1.655***	1.598***	1.634***
Observations	72	72	72	72
Wald test	11.86**	41.18***	53.97***	54.84***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Exp. Inv.' = Expectation of Investment;

'Exp. Sab.' = Expectation of Sabotage.

both in model 11 and 12 for sabotage decisions: the higher the expectation of sabotage of Low Type participants the higher the sabotage level[44].

RESULT 3. *As predicted by hypothesis 2, participants' expectations of their opponent's level of sabotage affect participants' decisions of sabotage.*

In particular, we found that Low Type participants with high expectation of sabotage are more likely to sabotage.

Subjects' Behaviour and Nash Equilibrium

Graph 4.6 shows the plot of investment and sabotage. Decisions seem to be positively correlated[45]. Many observations show the decision to not sabotage at all. Nearly 10% of participants do not invest (investment = 0) and do not sabotage (sabotage = 0), 11% of them chose to not invest but sabotage up to the maximum level (3% chose sabotage = 50). 16% of participants do not sabotage but invest up to 51[46].

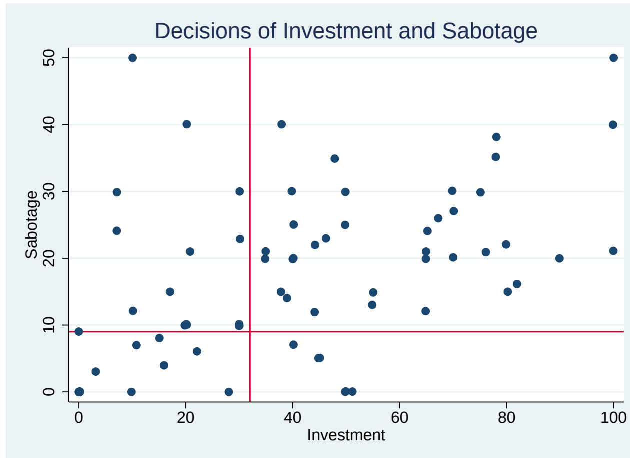
Overall the average level of investment is higher than the optimal value (Wilcoxon signed-ranks test, p = 0.025). Low Type subjects invest on average significantly more than the optimal value (Wilcoxon signed-ranks test, p-value = 0.002). In particular, they invest significantly more than 32 both when they do not receive the bonus (Wilcoxon signed-ranks test, p = 0.077) and when they do receive the

[44] As for Experiment 1 we discuss the impact of some demographic variables in the models 19-21 (see Table C.14 in Appendix C.7). Younger students are more likely to invest while for sabotage decisions nationality has a role: British students are less likely to sabotage their opponents. We ran also a separate regression to check the impact of social desirability on subjects' decisions. We did not find any significant effects.

[45] Spearman correlation coefficient = 0.479, $p < 0.001$. See Table C.15 for correlation coefficients in Appendix C.7.

[46] Graphs C.17, C.18, and C.19 in Appendix C.7 plot the investment and sabotage decisions by participants' type, in the Meritocracy treatment, and in the Anti-Meritocracy treatment respectively.

Figure 4.6: Decisions of Investment and Sabotage - Overall



bonus (Wilcoxon signed-ranks test, $p = 0.011$)[47]. High Type participants' decisions of investment are never significantly far from the optimal level.

Sabotage is on average significantly higher than the equilibrium level (Wilcoxon signed-ranks test, $p = 0.000$) and the difference between the observed average value and the optimal value is significant in both Meritocracy treatment (Wilcoxon signed-ranks test, $p = 0.000$) and Anti-meritocracy treatment (Wilcoxon signed-ranks test, $p = 0.003$). Specifically, in the Meritocracy treatment both types of participants decided to sabotage an amount higher than the optimal level (High Type, Wilcoxon signed-ranks test, $p = 0.005$; Low Type, Wilcoxon signed-ranks test, $p = 0.009$). In the Anti-Meritocracy treatment, instead, only Low Type subjects sabotage significantly more than the optimal value (Wilcoxon signed-ranks test, $p = 0.008$)[48].

[47]Sign tests give similar results.

[48]Sign tests confirm that the difference from the optimal behaviour is statistically significant.

4.5.2 Final Questions on Investment and Sabotage Choices

Methodology

A better understanding of the reasons at the basis of participants' choices of investment and sabotage may be obtained from the analysis of the participants' answers to two final open questions. They provided short answers explaining why they choose their specific level of investment and sabotage.

To analyse this set of qualitative data we applied the *Grounded Theory Method*[49] (GT), a methodology widely used in qualitative research. The GT method enables the researcher to investigate a large amount of textual information and systematically identify its properties, e.g. the frequencies of the most used keywords[50].

We followed the four steps of the GT method[51] for the two subsets of qualitative we have collected: 72 participants' answers of the first question (question about the decision of investment), and 72 participants' answers of the second question (question about the decision of sabotage)[52] .

The eleven categories we created are:

[49]Glaser and Strauss (1967).

[50]This aspect is similar to the content analysis method. We choose however the GT method because the content analysis should depart from a hypothesis. We indeed prefer to keep our qualitative analysis at an explorative level and to not state any behavioural haypothesis a priori.

[51]We first gave an overall reading of the answers in order to identify anchors that allow the key points of the data to be gathered (*step 1 - codes*). We then highlighted the codes and we grouped those of similar content (*step 2 - concepts*). We thus created eleven *categories*, which consist of either only one concept or a group of concepts of similar content (*step 3 - categories*). We finally aim to find out which categories are more likely to explain (or are correlated) to individuals' choices of investment and sabotage (*final step*).

[52]We actually have 65 out of 72 answers about investment and 66 out of 72 about sabotage. This because there were students that did not understand the question, or that skipped the question, etc.

1. Cost of the Number (A or B);
2. Competitive Feelings;
3. Random Choice;
4. Expectation of Investment;
5. Expectation of Sabotage;
6. Risk aversion;
7. Cost-benefit choice based;
8. Costs + Competitive Feelings;
9. Risk aversion + Expectation of Investment;
10. Non-competitive preferences;
11. Envy.

We used the same categories for both subsets of answers on investment and sabotage. We assigned a number from 0 to 10 to each category (categorical value), as it is illustrated in the bullet points above. We therefore used those numbers to indicate which category each participants' answer belongs to. We are now going to discuss categories' frequencies.

Decision of Investment

26% of participants wrote that their choice of number A (level of investment) was based on the cost of the number:

> 'This number does not cost many points'; 'I wanted it to be a low number so that only a small amount was deducted from my final earnings'; 'I chose 20 because the cost is not too high'.

13% of participants made their decisions based on their expectation of their opponent's investment:

> 'It is likely that the other participant would choose a high number'; 'I know Player X will choose higher number than me'; '[..] since I have chosen 50 as the guess of participant X's choice of Number A'.

The same proportion (13%) of participants seems to justify their choices by their risk attitude:

> '[...] playing it safe'; "Because the higher I choose the higher risks I will get'; 'Because I think 46 has a lower risk for me'.

A share of 10% of answers is assigned to category 7, represented together by motivation of costs and competitive feelings. For example a participant wrote: '30 was a good value for money! Somewhere the value-for-money aspect stops: you also want to beat Participant X.'.

In the Meritocracy treatment 35% of High Type participants considered costs as a determinant in their investment choice and 23% of them had concerns about their opponent's level of investment (expectation of investment). The same proportion of answers is assigned both to category 6 (cost-benefit choice based) and category 7 (cost + competitive feelings). 35% of Low Type participants chose their level of investment thinking about the level of cost (as well as High Type participant in the same treatment). However, 17% of them declared themselves risk averse to such an investment.

In the Anti-Meritocracy treatment the majority of High Type participants (28%) use their risk aversion as a motivation of their choice. Also the cost of the investment had a substantial role (21%). Interestingly, 14% of subjects were concerned about their opponent's level of sabotage. The same number of participants chose their Number A at random. Low Type participants (precisely 28% of them), who in this treatment have received the bonus, seems to think more carefully about their opponent's level of investment (category 3). They also consider carefully (17%) the cost-benefit trade-off in their choice: they want to invest but they are refrained by the high cost. 17% of them express their wish to win the 'winner prize' of the competition: 'I wanted to get the highest number so no matter what they took away I would still win'. 11% considered the cost of investment and 17% thought about their opponent's level of sabotage when they made their investment's decision.

Table 4.12 shows the percentage per category of participants' answers about their motivation for their level of investment.

Table 4.12: Percentage of answers per category - Investment Decisions

	Meritocracy		Anti-Meritocracy	
	High Type	Low Type	High Type	Low Type
1. Cost of the Number A	35.29	35.29	21.43	11.76
2. Competitive Feelings	.	11.76	7.14	17.65
3. Random Choice	5.88	11.76	14.29	5.88
4. Expectation of Investment	23.53	.	7.14	23.53
5. Expectation of Sabotage	5.88	.	14.29	11.76
6. Risk aversion	5.88	17.65	28.57	5.88
7. Cost-benefit choice based	11.76	11.76	.	17.65
8. Costs + Competitive Feelings	11.76	11.76	.	5.88
9. Risk aversion + Expectation of Investment	.	.	7.14	.
10. Non-competitive preferences	.	.	.	.
11. Envy	.	.	.	.

The value . refers to 0%;

Decision of Sabotage

Sabotage effort was more expensive than investment and subjects realized this[53]. Many of them (21%) considered this aspect in their choice of the level of sabotage. However, 18% of subjects expressed their wish to win: 'I don't want a too low number otherwise participant X would have bigger chance to win.'. Category 1, Competitive Feelings, occurred therefore with the second highest frequency in participants' answers. Category 4, Expectation of Sabotage, occurred with a high frequency as well in subjects' answers (12%). The same proportion of participants wrote either about their risk attitude in choosing their level of sabotage or about their willingness to not reduce their opponent's outcome. We called this category 'Non-competitive preferences'. For example they wrote: 'I did not want my partner to lose money"; "No point in taking money away from my opponent'.

If we consider only the answers given by High Type subjects in the Meritocracy treatment, we notice that category 0 has the highest proportion (33%), while category 1, Competitive Feelings, and category 4, Expectation of Sabotage, have the same highest percentage (16%). It seems that High Type participants considered costs in their decision of sabotage, but they thought as well about their opponent's level of sabotage. They felt also more willing to sabotage in order to win the highest prize of the competition. Low Type participants in this treatment wish also to compete aggressively (by sabotaging) in order to win the 'winner prize' (category 1, Competitive Feelings, 23%) but they also have concerns about the cost of this type of effort (category 0, Costs, 23%). However, a significant proportion of them did not want to reduce their opponents' outcome (category 9, Non-competitive preferences, 17%). Many of them selected a number at random (17%).

Risk aversion becomes the main concern for the majority of the High Type participants in the Anti-Meritocracy treatment where they are not given the bonus

[53] They could see from the instructions but it was not explicitly said.

(30%). 15% of them think about cost before making a choice and 15% of them interestingly expects that their opponent will choose a 'certain' level of sabotage. It is true, in fact, that 22% of Low Type participants explained that their choice was guided by their wish to win the competition (Competitive Feelings) and 16% of them chose a level of sabotage thinking about the balance cost-benefit (category 6).

Table 4.13 shows the percentage per category of participants' answers about their motivation for their level of sabotage.

Table 4.13: Percentage of answers per category - Sabotage Decisions

	Meritocracy		Anti-Meritocracy	
	High Type	Low Type	High Type	Low Type
1. Cost of the Number B	33.33	23.53	15.38	11.11
2. Competitive Feelings	16.67	23.53	7.69	22.22
3. Random Choice	.	17.65	7.69	11.11
4. Expectation of Investment	11.11	.	.	.
5. Expectation of Sabotage	16.67	5.88	15.38	11.11
6. Risk aversion	.	5.88	30.77	11.11
7. Cost-benefit choice based	5.56	.	.	16.67
8. Costs + Competitive Feelings	11.11	5.88	7.69	5.56
9. Risk aversion + Expectation of Investment	.	.	.	.
10. Non-competitive preferences	5.56	17.65	7.69	11.11
11. Envy	.	.	7.69	.

The value . refers to 0%;

4.5.3 Experiment 2: Discussion

The results of Experiment 1 have shown that the allocation of a bonus affects individual's behaviour in tournaments. In particular, we became very interested in Low Type contestants' decisions of investment and sabotage and we noticed that they are more willing to sabotage when they receive the bonus. This result suggested us the possible role of expectations in explaining Low Type participants' behaviour. To focus more in detail on this role we ran a second experiment, Experiment 2, where subjects are asked to elicit their expectations before entering the contest. The elicitation of their expectations is incentivised according to a linear incentive scheme.

As for Experiment 1, we find that individual behaviour does differ across treatments and participants' type. Overall, we find that Low Type participants invest more than High Type participants, especially when they receive (unfairly) the bonus. They invest significantly more than the optimal level. Moreover, they sabotage significantly more than the optimal level both when they receive and when they do not receive they bonus. We notice that participants' decision of the level of investment strongly depends on the expectation of investment (Result 7), but not on the expectation of sabotage. Finally, we do find strong evidence that the Low Type participants are more likely to sabotage when they expect a high level of sabotage (Result 8).

In Experiment 2 we ask participants to briefly explain their motivations of their choices of investment and sabotage. From this set of qualitative data we have collected at the end of the laboratory sessions we are able to bring some interesting discussion points on the individuals' behaviour and choices of investment and sabotage. The relative costs of each level of investment and sabotage play a crucial role in subjects' overall decisions. Many of the explanations submitted by subjects have reference to the cost and benefit they are considering in making their choice.

Specifically, costs and competitive feelings are important factors that frequently appeared overall in participants' explanations of their decision of sabotage. Moreover, we notice that when the bonus was fairly allocated (Meritocracy treatment) Low Type participants have more competitive attitude, and High Type participants expected this (the correlation coefficient between sabotage and expectation of sabotage is positive). However, it seems that Low Type subjects believed that the bonus allocation was objectively fair in the Meritocracy treatment. On the other hand, when the Low Type participants received the bonus in the Anti-Meritocracy treatment they felt more encouraged to sabotage so that they could win more. High Type participants expected this (correlation coefficient: 0.2805) and they generally were risk averse in their decision of sabotage (the correlation coefficient between sabotage and risk aversion is positive).

4.6 General Discussion and Conclusion

In our experiments (Experiment 1 and Experiment 2) participants took part to a 2-player contest game. They entered a competition for a fixed 'winner prize' by investing effort. They also had the possibility to engage in destructive effort, by reducing the level of investment of their opponent. Players had identical cost structure. Before entering the contest, they received some information about their performance of a previous (real effort) task, their ranking, and about their opponent's performance and ranking. Additionally, they were informed about the possibility to get a monetary reward that we called 'bonus' and how this was allocated among the two players. The allocation rule of the bonus differed across treatments: in Random Allocation 1 and Random Allocation 2 treatments the bonus was randomly allocated, in the Meritocracy treatment the bonus was given to participants that perform better the real effort task (High Type subgroup), whereas in the Anti-Meritocracy treatment the bonus was given to the subgroup of participants that performed worse the real effort task (Low Type subgroup).

The bonus was completely independent of the prize structure of the tournament. We expected, therefore, that the allocation of the bonus did not have any effects on participants behaviour in the tournament.

Results show that contestants choose on average a level of investment and sabotage significantly different from the optimal level and that there are differences in behaviour across treatments and participants' type. In Experiment 1 we notice how the allocation rule of the bonus affects the overall level of investment: participants are less likely to invest when the bonus is randomly allocated. They instead bid more aggressively when a rule of meritocracy (anti-meritocracy) is used. Specifically, we find that overall subjects are less likely to invest when they receive the bonus. However, if we consider participants' type, Low Type subjects invest more when they receive (unfairly) the bonus.

Another interesting result that attracted mostly our attention for future investigation is the Low Type participants' decision to sabotage the opponent. They sabotage more than the optimal level when they receive they bonus, that is when the anti-meritocracy rule is applied. We explain this behaviour with the role of expectation of sabotage. The higher their expectations of sabotage the higher their sabotage level. Probably, they expect to be victim of sabotage since they have received the bonus unfairly and hence they decide to sabotage more than the optimal level. Interestingly, the expectation of investment is not taken into account for sabotage decisions. In Experiment 2 we bring forward the elicitation of participants' expectations (and we introduce an incentive scheme for this) in order to test explicitly the role of expectations. We ask participants to guess their opponent's level of investment and sabotage. Results of Experiment 2 confirm the results of Experiment 1: beliefs of investments do not affect Low Type participants' sabotage decisions but beliefs of sabotage induce them to sabotage more when they expect to be sabotaged.

We support this argument with the results of the analysis of the set of qualitative data we collect at the end of the laboratory sessions of Experiment 2. It appears that Low Type participants feel more competitive after being assigned the bonus.

They also understand the objective fairness of the allocation rule in the Meritocracy treatment. The bonus unfairly given to them in the Anti-Meritocracy treatment induces them to engage more in destructive effort.

Mechanisms of (fair) allocation of outcomes, as for example monetary rewards, are factors external to the tournament game that we applied in our setting. They should not affect contestants' behaviour. However, we find that they do matter and significantly affect individuals' decisions in tournament. Our results suggest that a principal should consider effects of bonuses allocation and the fairness of the allocation procedures in pursuing his/her objectives. If the principal aims to increase overall productivity than a random allocation of bonuses is not effective. Allocating rewards among workers according to meritocracy rules increases competitive feelings and attitude, and finally overall agents' investment. If the principal aims, on the contrary, to promote workers' cooperative behaviour than the fairness of the allocation rule of bonuses matters. Not rewarding the best workers creates feelings of hostility and rewarded workers can engage in destructive behaviour which is not beneficial for the firm. The decision of unfairly rewarded workers of engaging in destructive effort comes from the expectation of sabotage they have. They in fact expected to be victim of sabotage. Expectations have, hence, a key role in understanding contestants' behaviour in cases where the fairness of a specific allocation procedure, applied before the contest, is controversial.

We add few final comments. We check for participants' risk aversion and we find that our result is consistent with the experimental literature (Millner and Pratt, 1991; Sheremeta and Wu, 2012): the higher risk aversion the lower the level of investment (and sabotage).

Moreover, we control for aggregate level of investment in tournaments when there are no opportunities of sabotage. In this case, subjects should put more costly resources on productive effort. Considering data of the Random Allocation 2 treatment as well (and comparing investment level across treatments) we find that the average level of investment is lower when sabotage is possible.

Finally, in the light of the results obtained and described in Chapter 3 of this thesis, we acknowledge that our results may depend on the effort task used. The implementation of a different effort task, as for example a real effort task, might (or might not) bring to the same results.

Chapter 5

Conclusion

This doctoral thesis aims first at investigating the effects of emotions on agents' performance. Secondly, it contributes to the scientific discussion about a methodological issue on the use of effort tasks in the economic laboratory experiments.

It is structured in three substantial essays. The first essay (Chapter 2) explores the motivating role of anxiety and its positive effects on agents' performance in an individual task under a piece rate incentive scheme. Contrary to the negative valence given to this emotion, anxiety can induce agents' to exert more effort in order to counter-balance the negative consequences of a poor outcome. I find that participants in my laboratory experiment, whose anxious mood was induced, achieve on average a higher level of performance than participants whose mood was positively manipulated.

In the second essay, Chapter 3, I point out the implications of the use of effort tasks in economic laboratory experiments. Participants are asked to carry out some tasks that allow the experimenter to measure for example participants' performance or decisions of investment. Effort tasks can be either real effort tasks, which require mental or physical effort to be performed, or induced value effort tasks, which asked participants to choose a costly level of desired effort or investment.

Experimenters use effort tasks in an equivalent way: they do not acknowledge the limitation of this tool in presenting their experimental results. Effort tasks might not lead in fact to similar results since they vary a lot in their nature.

My laboratory experiment compares the effect of a group of explanatory variables on agents' performance. Four different effort tasks have been used, three real effort tasks and an induced value effort task. Results show that there is no equivalence between tasks and that the experimental data might be significantly affected by the choice of the task.

In the last essay, Chapter 4, I considered the effects of different allocation rules of an extra monetary reward (bonus) on agents' behaviour in contests. We collected experimental data. Assuming profit maximizer agents, we expect that they would choose the optimal level of investment and sabotage in a tournament with fixed prizes. The tournament prize structure is in fact independent of the monetary reward. However, we find that contestants' behaviour is affected by the allocation of the bonus and that perceived fairness, expectations and competitive feelings drive contestants' choices of investment and sabotage.

An important qualification of our findings has referred to earlier but now is worth reiterating. I use a real effort task in my first essay and an induced value effort task in my third essay. I draw conclusions from the experimental data thus collected. However, I do acknowledge that my experimental results might depend on the effort tasks that I have used. The methodological lesson that comes from my second essay does suggest the limitation and the potential loss of external validity of the research when effort tasks are used in the economic laboratory. The reason why I have applied an induced value effort task in the last essay (Chapter 4) is that it assures full control over the cost of (productive and destructive) effort and guarantees very strong internal validity of the experiments run.

This doctoral thesis suggests interesting ideas for future research. Here just few examples. Monitoring mechanisms may be explored in order to check their joint effects with individual anxiety on performance. Moreover, other types of incentives can be studied rather than tournaments when it comes necessary in a firm to use particular allocation procedure. Finally, it would be very useful to strengthen the investigation on the limitations of the effort tasks implemented in the laboratory in order to provide complete evidence of the risks and potential of each (or the most used) type of task.

Appendix A

(appendix Chapter 2)

A.1 The State-Trait Anxiety Inventory (STAI) Questionnaire

[Instructions and statements for State Anxiety]

A number of statements which people have used to describe themselves are given below. Read each statement and than blacken in the appropriate circle to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

[Participants could select their answer by clicking one of the 5 buttons to the right of each statement where 1 indicates 'Not at all' and 4 indicates 'Very much so'. To the four-point Likert scale answers scheme another possible answer was added: "I did not understand the question", the fifth button to the right. Subjects that were not sure about the meaning of the question could select it. This type of answer was considered as not answering the question in the score counting. Scores were adjusted according to the number of not answered questions. If there were 3 or more questions not answered for each set of questions (Trait or State) the score for that set was considered missing.]

1. I feel calm
2. I feel secure
3. I am tense
4. I feel strained
5. I feel at ease
6. I feel upset
7. I am presently worrying over possible misfortunes
8. I feel satisfied
9. I feel frightened
10. I feel comfortable
11. I feel self-confident
12. I feel nervous
13. I am jiitery
14. I feel indecisive
15. I am relaxed
16. I feel content
17. I am worried
18. I feel confused
19. I feel steady
20. I feel pleasant

[Instructions and statements for Trait Anxiety]

A number of statements which people have used to describe themselves are given below. Read each statement and then blacken in the appropriate circle to the right of the statement to indicate how you *generally* feel. There are no right or wrong answer. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

21. I feel pleasant

22. I feel nervous and restless

23. I feel satisfied with myself

24. I wish I could be as happy as others seem to be

25. I feel like a failure

26. I feel rested

27. I am "calm, cool, and collected"

28. I feel that difficulties are piling up so that I cannot overcome them

29. I worry too much over something that really doesn't matter

30. I am happy

31. I have disturbing thoughts

32. I lack self-confidence

33. I feel secure

34. I make decision easily

35. I feel inadequate

36. I am content

37. Some unimportant thought runs through my mind and bothers me

38. I take disappointments so keenly that I can't put them out of my mind

39. I am a steady person

40. I get a state of tension or turmoil as I think over my recent concerns and interests

A.2 Experimental Instructions

[The experimenter reads these instructions at the beginning of the experiment].

You are participating in an experiment on decision making. The experiment is expected to last 1 hour and 15 minutes. Your earnings in this experiment will depend on your decisions and ability. A participation fee of £3 is guaranteed. During the experiment you are not allowed to communicate with other participants. If you have any questions at any moment please raise your hand and the experimenter will come to your desk.

The experiment consists of **THREE SECTIONS**:

FIRST SECTION

In the first section you are asked to fill in a questionnaire. The questionnaire consists of 2 parts: PART A and PART B. For each part you will receive specific instructions. Please read the instructions carefully.

After that all of you complete the questionnaire, you will receive the instructions for the second and third section of the experiment.

[The experimenter reads these instructions after that participants have completed the (STAI) questionnaire].

SECOND SECTION

In the second section you will be asked to answer some questions and to submit your answers.

THIRD SECTION

The task

In the third section you are asked to carry out a task for 10 rounds. In each round you will see 48 sliders on the screen. Each slider has integer values from 0 to 100. Each slider will appear on position 0. The sliders can be adjusted and readjusted an unlimited number of times and the current position is displayed to the right of each slider.

Your task is to position each slider at exactly 50 with the help of your mouse. You will have 2 minutes to carry out your task that is each round will last for 2 minutes. The remaining time is displayed in seconds in the top right corner of the screen.

Your score

You get one point per each slider you have positioned at 50.

Example:

Suppose you position five sliders at exactly 50 then your points score is 5.

At the beginning of each round your score is reset. This means that scores of previous rounds will not be added to scores of successive rounds.

The payment

At the end of the experiment the computer will randomly select one round and you will be paid based on your score of that round. Your earnings will be: your score x £0.30.

Example:

Suppose your score in round 1 is 5. If the computer selects round 1 then your earnings are: 5 x £0.30 = £1.50.

Example:

Suppose your score in round 2 is 10. If the computer selects round 2 then your earnings are: 10 x £0.30 = £3.

Example:

Suppose your score in round 9 is 20. If the computer selects round 9 then your earnings are: 20 x £0.30 = £6.

Your total earnings in this experiment will be £3 plus £0.30 x your score of the selected round.

STAI Questionnaire Instructions.

Instructions for the State Anxiety

You will be shown 20 statements. Read each statement and click on the appropriate answer to indicate how you feel *right now*, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe best your present feelings.

Instructions for the Trait Anxiety

You will be shown 20 statements. Read each statement and click on the appropriate answer to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe best how you generally feel.

Mood Manipulation Instructions.

Please think about two of your personal situations over the last year in which you felt very anxious. Please think about those events in details, remembering how you felt and what happened.

Please write down about the *first* personal situation you have previously thought about. Feel free to write whatever you want about that.

Please write down about the *second* personal situation you have previously thought about. Feel free to write whatever you want about that.

Please write down *some adjectives* that describe exactly your feelings in personal situations you have previously thought about. Feel free to write whatever you want about that.

A.3 Tables and Figures

Figure A.1: Slider Task

Table A.1: Aggregate average values of emotions elicited after the task

	Main Treatmens	**Control Treatments**
Anxiety	0.77 (1.22)	1.17 (1.56)
Happiness	1.58 (1.28)	1.82 (1.51)
Tiredness	1.87 (1.58)	2.12 (1.63)
Despondency	0.62 (1.02)	0.70 (1.17)
Sadness	0.48 (0.89)	0.67 (1.13)
Apprehension	0.83 (1.17)	1.04 (1.25)
Anger	0.75 (1.16)	0.43 (0.90)
Frustration	1.27 (1.38)	1.24 (1.49)

Note: Standard Deviation in parentheses

Table A.2: OLS Estimates with Error Clustering On Performance with Interaction Terms

	Model 3		Model 4	
Period	0.336***	(5.03)	0.341***	(5.03)
Main	0.616	(0.92)	0.860	(1.39)
Number of Sessions	-0.185	(-0.14)	-0.822	(-0.67)
Anxiety	-0.541	(-0.41)	0.835	(0.69)
Despondency	2.768***	(3.31)	2.247***	(2.95)
Sadness	-0.291	(-0.72)	0.103	(0.26)
Apprehension	-0.786	(-1.34)	-1.060*	(-1.91)
Tiredness	-0.681*	(-1.93)	-0.526	(-1.56)
Happiness	-0.516	(-1.62)	-0.147	(-0.49)
Anger	-0.162	(-0.24)	-0.173	(-0.25)
Frustration	-0.988	(-1.50)	-1.460**	(-2.22)
State Score	-0.157*	(-1.73)	-0.083	(-0.99)
Trait Score	0.028	(0.34)	0.028	(0.41)
Anxiety x Period	-0.011	(-0.40)	-0.013	(-0.46)
Anx x NS	-0.038	(-0.06)	0.052	(0.09)
Anx Square	-0.305	(-1.19)	-0.199	(-0.80)
Anx x Trait Anx	-0.014	(-0.38)	-0.019	(-0.64)
Anx x State Anx	0.063	(1.64)	0.039	(1.12)
Anx x Despondency	-1.194***	(-2.90)	-0.759*	(-1.97)
Anx x Apprehension	0.211	(0.82)	0.162	(0.70)
Anx x Tiredness	0.379**	(2.01)	0.196	(1.09)
Anx x Frustration	0.241	(0.73)	0.336	(1.09)
Anx x Anger	0.309	(0.98)	0.164	(0.49)
Age			-2.044***	(-4.06)
Gender			3.142***	(2.96)
British Nationality			1.650***	(2.65)
School of Economics			0.576	(0.50)
Anx x Gender			-0.542	(-1.03)
Constant	22.310***	(6.31)	17.499***	(5.24)
Observations	1480		1470	
R-Square	0.158		0.289	
F-test	5.351		5.109	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 3: Error clustering without demographic variables;

Model 4: Error clustering with demographic variables;

Anx x NS: Anxiety x Number of Sessions;

abs(Anx) and abs(Anx)Square have been omitted.

Table A.3: OLS Estimates with Error Clustering On Performance without Frustration or Despondency

	No Frustration		No Despondency	
	Model 5	Model 5a	Model 6	Model 6a
Period	0.322***	0.320***	0.322***	0.320***
Main	0.776	0.739	0.578	0.578
Number of Sessions	-0.877	-0.713	-1.004	-0.834
abs(Anx)	-0.137	-0.109	0.130	0.040
abs(Anx)Square	0.004	-0.055	-0.186	-0.183
Anxiety	0.749**	0.434	0.969***	0.588*
Apprehension	-0.916***	-0.599*	-0.795**	-0.509
Despondency	0.611	0.529		
Sadness	-0.079	-0.503	0.181	-0.316
Tiredness	-0.284	-0.180	-0.225	-0.138
Happiness	-0.008	-0.266	-0.067	-0.316
Anger	0.156	0.554	0.300	0.671*
Gender	2.355***		2.492***	
British Nationality	1.597**		1.645**	
School of Economics	0.172		0.284	
Age	-2.362***		-2.293***	
State Score	-0.004	-0.043	0.022	-0.020
Trait Score	-0.018	0.012	-0.025	0.004
Frustration			-0.424	-0.231
Constant	17.057***	18.951***	16.656***	18.860***
Observations	1470	1480	1470	1480
R-Square	0.251	0.090	0.249	0.086
F-test	6.166	6.239	6.079	6.185

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$;

Models 5a and 6a do not include demographic variables.

Table A.4: OLS Estimates with Error Clustering on Performance by Trait Anxiety Groups without Demographic Variables

	High Trait		Medium Trait		Low Trait	
Period	0.341***	(3.40)	0.279***	(6.18)	0.395***	(4.03)
Main	0.605	(0.36)	1.186	(1.49)	0.371	(0.11)
Number of Sessions	-0.624	(-0.47)	-2.864***	(-3.23)	2.518	(1.45)
abs(Anx)	-7.468	(-1.30)	-0.298	(-0.13)	5.017	(0.80)
abs(Anx)Square	1.669	(0.85)	0.212	(0.23)	-1.044	(-0.37)
Anxiety	2.413**	(2.57)	-0.362	(-0.81)	0.396	(0.45)
Despondency	-1.024	(-0.84)	0.676	(1.10)	2.303	(1.06)
Sadness	-0.555	(-0.58)	0.652	(1.23)	-1.969	(-1.16)
Apprehension	-1.142	(-1.47)	-0.810*	(-1.99)	-0.430	(-0.34)
Tiredness	-0.687	(-1.22)	-0.430	(-1.48)	-0.100	(-0.14)
Happiness	1.080	(1.40)	-0.771**	(-2.41)	-0.967	(-0.78)
Anger	0.231	(0.27)	0.649	(1.33)	-0.024	(-0.02)
Frustration	-0.376	(-0.60)	-0.973**	(-2.17)	2.935	(1.57)
State Score	0.047	(0.48)	-0.043	(-0.60)	-0.001	(-0.00)
Trait Score	0.331***	(3.03)	-0.080	(-0.69)	-0.046	(-0.13)
Constant	1.268	(0.22)	26.547***	(5.40)	14.404	(1.65)
Observations	380		780		320	
R-Square	0.266		0.249		0.245	
F-test	3.731		5.437		6.266	

t statistics in parentheses;

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.2: Trait Anxiety Score Distribution

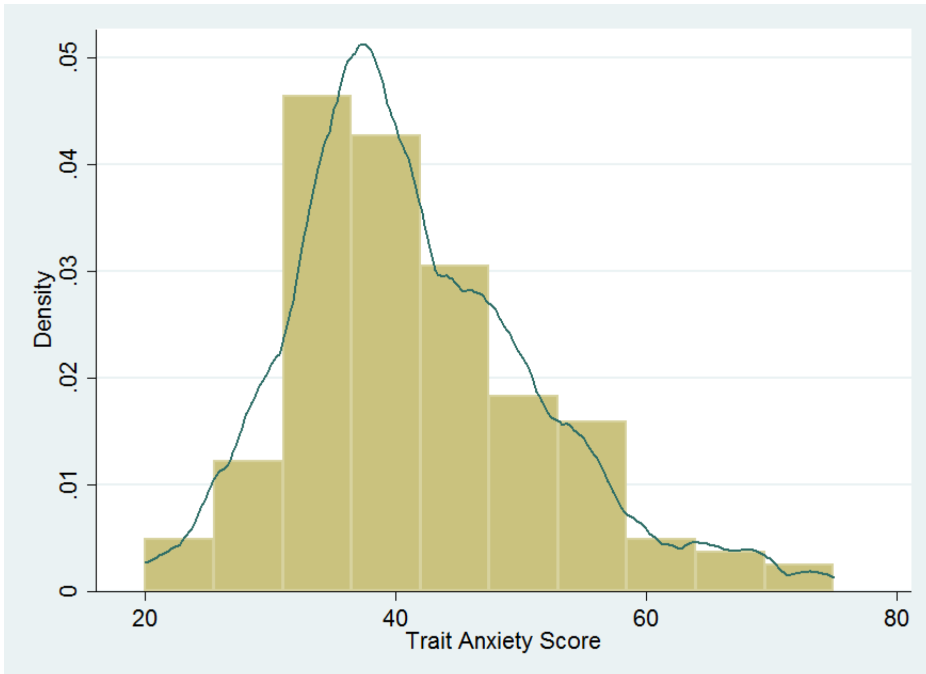
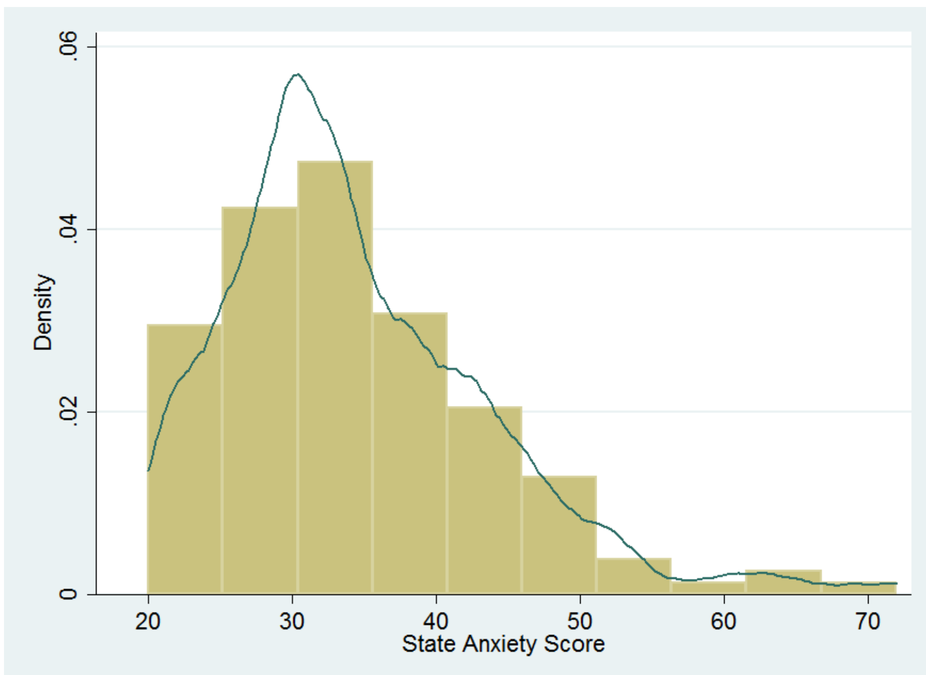


Table A.5: Aggregate average values of anxiety by Trait Anxiety Groups

Low Trait Anxiety	**Obs**	**Mean**	**St. Dev**	**Min**	**Max**
Anxiety	320	1.094	1.210	0	4
Medium Trait Anxiety	**Obs**	**Mean**	**St. Dev**	**Min**	**Max**
Anxiety	790	1.367	1.315	0	4
High Trait Anxiety	**Obs**	**Mean**	**St. Dev**	**Min**	**Max**
Anxiety	400	1.85	1.315	0	4

Figure A.3: State Anxiety Score Distribution



Appendix B

(appendix Chapter 3)

B.1 Experimental Instructions

Experimental Instructions - Treatment A - The Slider Task

[The experimenter reads this instructions at the beginning of the experiment].

You are participating in an experiment on decision making. The experiment is expected to last 1 hour.

During the experiment you are not allowed to communicate with other participants. If you have any questions at any moment please raise your hand and the experimenter will come to your desk.

The experiment consists of **THREE** SECTIONS:

FIRST SECTION

In the first section you are asked to fill in a questionnaire. The questionnaire consists of 2 parts: PART A and PART B. For each part you will receive specific instructions. Please read the instructions carefully.
After that all of you have completed the questionnaire, you will receive the instructions for the second and third section of the experiment.

[The experimenter reads this instructions after that participants have completed the (STAI) questionnaire].

SECOND SECTION

In the second section you are randomly matched with another participant in this room. You will not be told whom you are matched with. We will call him or her your “co-participant”.

In this section you and your co-participant are asked to perform a task that will be described shortly. In this task you get a score based on your performance.

The task

You will see 48 sliders on the screen. Each slider has integer values from 0 to 100.

Each slider will appear on position 0.

Your task is to position each slider at exactly 50 with the help of your mouse. The sliders can be adjusted and readjusted an unlimited number of times and the current position is displayed to the right of each slider.

Your score

You get a score of one point per each slider you have positioned at 50.

Example:

If you position five sliders at exactly 50, you get a score of 5 points.

You will have two minutes to correctly position as many sliders as you can. We will call these two minutes “round”.

In this experiment you are asked to carry out this task for 10 rounds. You will be matched with the same co-participant among the 10 rounds. All the sliders will appear at position 0 at the beginning of each round. Therefore, your score will be reset. This means that scores of previous rounds will not be added to scores of successive rounds.

Your earnings in the second section

At the end of the experiment the computer will randomly select one round out of 10. The selected round will be the same for you and your co-participant.

The computer will compare your score to the score of your co-participant obtained in the selected round. Three situations can occur:

- If your score in the selected round is higher than your co-participant's score, you earn a prize of £10;
- If your score in the selected round is lower than your co-participant's score, you earn £0;
- If your score in the selected round is equal to your co-participant's score, the computer will randomly select one of you; you have therefore a 50% chance of being selected and of winning the prize.

Example 1

Suppose the computer selects round 1. If your score in round 1 is higher than your co-participant's score, your earnings are £10.00.

Example 2

Suppose the computer selects round 5. If your score in round 5 is lower than your co-participant's score, your earnings are £0.00.

Example 3

Suppose the computer selects round 10. If your score in round 10 is equal to your co-participant's score, your earnings are £10 or £0 with a 50% chance.

THIRD SECTION

In the next screen you will be asked to make some decisions. Your decisions will affect your final earnings: what you will earn from this task will be added to your earnings of the previous task.

You now need to make 9 decisions for each of two successive computer screens. Each decision is a paired choice between two options (for example, “Option A” and “Option B”). You will make 9 decisions and record these in the final column, but only one of them from each computer screen will be used in the end to determine your earnings. You will only know which one at the end of this session.

Experimental Instructions - Treatment B - The Maths Task

[The experimenter reads this instructions at the beginning of the experiment].

You are participating in an experiment on decision making. The experiment is expected to last 1 hour.

During the experiment you are not allowed to communicate with other participants. If you have any questions at any moment please raise your hand and the experimenter will come to your desk.

The experiment consists of **THREE** SECTIONS:

FIRST SECTION

In the first section you are asked to fill in a questionnaire. The questionnaire consists of 2 parts: PART A and PART B. For each part you will receive specific instructions. Please read the instructions carefully.

After that all of you have completed the questionnaire, you will receive the instructions for the second and third section of the experiment.

[The experimenter reads this instructions after that participants have completed the (STAI) questionnaire].

SECOND SECTION

In the second section you are randomly matched with another participant in this room. You will not be told whom you are matched with. We will call him or her your "co-participant".

In this section you and your co-participant are asked to perform a task that will be described shortly. In this task you get a score based on your performance.

The task

The task consists of adding-up four numbers that are randomly generated and that are displayed on your screen.

For example, you will see on your screen:

$$50 + 20 + 10 + 33$$

Your task is to sum these numbers that is 50 + 20 + 10 + 33. The correct answer in this example is 113.

The use of paper, pencil or calculator is forbidden. You must make these calculations in your head.

Once your calculation has been done, you enter your answer in the specified area and you submit your answer by clicking the "submit" button. Having submitted your answer a new series of numbers appears automatically on your screen. Therefore, you are asked to sum these new numbers and to submit your answer.

You will have two minutes to submit as many correct answers as you can. We will call these two minutes "round". In this experiment you are going to perform the arithmetic task for 10 rounds. You will be matched with the same co-participant among the 10 rounds but the series of numbers will be different in each round.

Your score

You get a score of one point per each correct answer you submit. You will not be told whether your answers are correct or wrong when you submit them.

Example:

In the example: 50 + 20 + 10 + 33 the correct answer is 113.

If your answer is 113 you get a score of 1 point.

At the beginning of each round your score is reset. This means that scores of previous rounds will not be added to scores of successive rounds. Your score is independent in each round.

Your earnings in the second section

At the end of the experiment the computer will randomly select one round out of 10. The selected round will be the same for you and your co-participant.

The computer will compare your score to the score of your co-participant obtained in the selected round. Three situations can occur:

- If your score in the selected round is higher than your co-participant's score, you earn a prize of £10;
- If your score in the selected round is lower than your co-participant's score, you earn £0;
- If your score in the selected round is equal to your co-participant's score, the computer will randomly select one of you; you have therefore a 50% chance of being selected and of winning the prize.

Example 1

Suppose the computer selects round 1. If your score in round 1 is higher than your co-participant's score, your earnings are £10.00.

Example 2

Suppose the computer selects round 5. If your score in round 5 is lower than your co-participant's score, your earnings are £0.00.

Example 3

Suppose the computer selects round 10. If your score in round 10 is equal to your co-participant's score, your earnings are £10 or £0 with a 50% chance.

THIRD SECTION

In the next screen you will be asked to make some decisions. Your decisions will affect your final earnings: what you will earn from this task will be added to your earnings of the previous task.

You now need to make 9 decisions for each of two successive computer screens. Each decision is a paired choice between two options (for example, "Option A" and "Option B"). You will make 9 decisions and record these in the final column, but only one of them from each computer screen will be used in the end to determine your earnings. You will only know which one at the end of this session.

Experimental Instructions - Treatment C - The Grid Task

[The experimenter reads this instructions at the beginning of the experiment]. You are participating in an experiment on decision making. The experiment is expected to last 1 hour.

During the experiment you are not allowed to communicate with other participants. If you have any questions at any moment please raise your hand and the experimenter will come to your desk.

The experiment consists of **THREE** SECTIONS:

FIRST SECTION

In the first section you are asked to fill in a questionnaire. The questionnaire consists of 2 parts: PART A and PART B. For each part you will receive specific instructions. Please read the instructions carefully.

After that all of you have completed the questionnaire, you will receive the instructions for the second and third section of the experiment.

[The experimenter reads this instructions after that participants have completed the (STAI) questionnaire].

SECOND SECTION

In the second section you are randomly matched with another participant in this room. You will not be told whom you are matched with. We will call him or her your "co-participant".

In this section you and your co-participant are asked to perform a task that will be described shortly. In this task you get a score based on your performance.

The task

For the task 1 you will see one table on the screen. You will be asked to count the number of ones in each table. Then please write down your answer in the box at the bottom of the screen and submit it by clicking the "Submit" button. After you have submitted your answer a new table will be generated and displayed on the screen. Please notice that once you have submitted you cannot change your answer.

This task will last for 2 minutes. Try to solve as many tables as you can within the 2 minutes.

The remaining time is displayed in seconds in the top right corner of the screen.

Your score

You get a score of one point per each table you have correctly solved.

Example: If you correctly solve five tables, you get a score of 5 points.

You will have two minutes to correctly solve as many tables as you can. We will call these two minutes "round".

In this experiment you are asked to carry out this task for 10 rounds. You will be matched with the same co-participant among the 10 rounds. At the beginning of each round your score will be reset. This means that scores of previous rounds will not be added to scores of successive rounds.

Your earnings in the second section

At the end of the experiment the computer will randomly select one round out of 10. The selected round will be the same for you and your co-participant.

The computer will compare your score to the score of your co-participant obtained in the selected round. Three situations can occur:

- If your score in the selected round is higher than your co-participant's score, you earn a prize of £10;
- If your score in the selected round is lower than your co-participant's score, you earn £0;
- If your score in the selected round is equal to your co-participant's score, the computer will randomly select one of you; you have therefore a 50% chance of being selected and of winning the prize.

Example 1

Suppose the computer selects round 1. If your score in round 1 is higher than your co-participant's score, your earnings are £10.00.

Example 2

Suppose the computer selects round 5. If your score in round 5 is lower than your co-participant's score, your earnings are £0.00.

Example 3

Suppose the computer selects round 10. If your score in round 10 is equal to your co-participant's score, your earnings are £10 or £0 with a 50% chance.

THIRD SECTION

In the next screen you will be asked to make some decisions. Your decisions will affect your final earnings: what you will earn from this task will be added to your earnings of the previous task. You now need to make 9 decisions for each of two successive computer screens. Each decision is a paired choice between two options (for example, "Option A" and "Option B"). You will make 9 decisions and record these in the final column, but only one of them from each computer screen will be used in the end to determine your earnings. You will only know which one at the end of this session.

Experimental Instructions - Treatment D - The IV Effort Task

[The experimenter reads this instructions at the beginning of the experiment].

You are participating in an experiment on decision making. The experiment is expected to last 1 hour. During the experiment you are not allowed to communicate with other participants. If you have any questions at any moment please raise your hand and the experimenter will come to your desk.

The experiment consists of **THREE** SECTIONS:

FIRST SECTION

In the first section you are asked to fill in a questionnaire. The questionnaire consists of 2 parts: PART A and PART B. For each part you will receive specific instructions. Please read the instructions carefully.

After that all of you have completed the questionnaire, you will receive the instructions for the second and third section of the experiment.

[The experimenter reads this instructions after that participants have completed the (STAI) questionnaire].

SECOND SECTION

In the second section you are randomly matched with another participant in this room. You will not be told whom you are matched with. We will call him or her your "co-participant".

In this section you and your co-participant are asked to perform a task that will be described shortly. In this task you get a score based on your performance.

The task

You are given an endowment of 10 Experimental Points.

You are asked to decide how much you want to invest and how much you want to keep for yourself.

For example, if you decide to invest 5.55 experimental points your remaining endowment will be 10.00 - 5.55 = 4.45 experimental points.

Once you have decided, you enter your answer in the specified area and you submit your decision by clicking the "submit" button.

You will be asked to make this decision 10 times. We will call these times "rounds". Therefore this experiment will have 10 rounds in which you are matched with the same co-participant. You receive an endowment of ten experimental points at the beginning of each round. You will not be told about your co-participant's decisions.

Your earnings in the second section

At the end of the experiment the computer will randomly select one round out of 10. The selected round will be the same for you and your co-participant.

The computer will compare your decision to the decision of your co-participant made in the selected round. Three situations can occur:

- If your investment in the selected round is higher than your co-participant's investment, you earn a prize of £10;
- If your investment in the selected round is lower than your co-participant's investment, you earn £0;
- If your investment in the selected round is equal to your co-participant's investment, the computer will randomly select one of you; you have therefore a 50% chance of being selected and of winning the prize.

Your remaining endowment in the selected round will be converted in pounds with the following conversion rate:

0.01 Experimental Points = £0.01

For example, 1 Experimental Point will be converted in £1.

Therefore, your earnings in the second section will be:

- £10 + your remaining endowment in the selected round if your investment in the selected round is higher than your co-participant's investment;
- £0 + your remaining endowment in the selected round if your investment in the selected round is lower than your co-participant's investment;
- £10 + your remaining endowment in the selected round if your investment in the selected round is equal to your co-participant's investment and the computer selects you as the winner;
- £0 + your remaining endowment in the selected round if your investment in the selected round is equal to your co-participant's investment and the computer selects your co-participant as the winner.

Example 1

Suppose the computer selects round 1.

Suppose that your investment in round 1 is 4.49 experimental points.

Your remaining endowment is 10.00 - 4.49 = 5.51 experimental points.

Your remaining endowment converted in pounds is £5.51.

If in round 1 your investment is higher than your co-participant's investment, your earnings are £10.00 + £5.51 = £15.51.

Example 2

Suppose the computer selects round 4.

Suppose that your investment in round 4 is 3.40 experimental points.

Your remaining endowment is 10.00 - 3.40 = 6.60 experimental points.

Your remaining endowment converted in pounds is £6.60.

If in round 4 your investment is lower than your co-participant's investment, your earnings are £0.00 + £6.60 = £6.60.

Example 3

Suppose the computer selects round 10.

Suppose that your investment in round 10 is 5.00 experimental points.

Your remaining endowment is 10.00 - 5.00 = 5.00 experimental points.

Your remaining endowment converted in pounds is £5.00.

If in round 10 your investment is equal to your co-participant's investment, your earnings are £10.00 + £5.00 = £15.00 if the computer selects you as the winner OR £0.00 + £5.00 = £5.00 if the computer selects your co-participant as the winner.

THIRD SECTION

In the next screen you will be asked to make some decisions. Your decisions will affect your final earnings: what you will earn from this task will be added to your earnings of the previous task. You now need to make 9 decisions for each of two successive computer screens. Each decision is a paired choice between two options (for example, "Option A" and "Option B"). You will make 9 decisions and record these in the final column, but only one of them from each computer screen will be used in the end to determine your earnings. You will only know which one at the end of this session.

B.2 Tables and Figures

Tables

Table B.1: Ranksum and t-test on Performance - Comparison by Set of Rounds

	Obs	**z-value**	**p-value**	**t-value**		**p-value**	
					Pr(T < t)	*Pr(|T|>|t|)*	*Pr(T > t)*
Slider Task	10	-2.611	0.0090	-4.1019	0.0017	0.0034	0.9983
Maths Task	10	-2.611	0.0090	-4.9631	0.0006	0.0011	0.9994
Grid Task	10	-2.619	0.0088	-5.9538	0.0002	0.0003	0.9998
Ind. Value Effort Task	10	-0.731	0.4647	0.3628	0.6369	0.7261	0.3631

Note: z-value for a two-sample Wilcoxon-Mann-Whitney test; t-value for the t-test.

Table B.2: Sign-Rank and t-test on Anxiety - Comparison by set of Rounds

	Obs	**z-value**	**p-value**	**t-value**		**p-value**	
					Pr(T < t)	*Pr(|T|>|t|)*	*Pr(T > t)*
Slider Task	50	1.854	0.0637	1.1912	0.8803	0.2393	0.1197
Maths Task	54	-1.104	0.2697	-0.8485	0.2000	0.4000	0.8000
Grid Task	52	1.662	0.0964	1.9256	0.9701	0.0597	0.0299
Ind. Value Effort Task	50	-0.029	0.9767	0.4411	0.6695	0.6610	0.3305

Note: z-value for the Wilcoxon signed-rank test; t-value for the t-test.

Table B.3: Aggregate Descriptive Statistics of Anxiety - by Sets of Rounds

	Obs	**Mean**	**Std.Dev**	**Min**	**Max**
Rounds 1 - 5					
Slider Task	50	45.12	31.885	0	100
Maths Task	54	56.07	27.53	0	100
Grid Task	52	49.67	30.013	0	100
Ind. Value Effort Task	50	26.42	26.520	0	100
Rounds 6 - 10					
Slider Task	50	40.20	31.784	0	100
Maths Task	54	59.41	31.745	0	100
Grid Task	52	43.21	31.917	0	100
Ind. Value Effort Task	50	24.88	30.569	0	100

Table B.4: Summary Statistics on Risk Aversion

	Obs	**Mean**	**Std. Dev**	**Min**	**Max**
Slider Task	50	5.38	1.794436	1	9
Maths Task	54	5.91	2.226	0	9
Grid Task	52	5.5	2.227	0	9
Ind. Value Effort Task	50	6.12	2.067	1	9

Table B.5: Summary Statistics on Performance by Gender

		N	Mean	Std.Dev.	Min	Max
Rounds 1 - 5						
Slider Task	Male	120	19.508	3.902	10	28
	Female	130	16.685	4.274	6	28
Maths Task	Male	105	4.6	2.003	0	9
	Female	165	4.067	1.897	0	9
Grid Task	Male	125	10.928	5.966	0	23
	Female	135	9.963	6.128	0	25
Ind. Value Effort Task	Male	125	3.246	3.771	0	10
	Female	125	4.341	3.336	0	10
Rounds 6 - 10						
Slider Task	Male	120	21.7	4.351	11	34
	Female	130	18.323	4.083	9	27
Maths Task	Male	105	5.257	2.171	1	10
	Female	165	4.806	2.373	0	12
Grid Task	Male	125	12.168	6.149	0	24
	Female	135	11.681	5.933	0	25
Ind. Value Effort Task	Male	125	3.189	3.706	0	10
	Female	125	4.264	3.669	0	10

Table B.6: Spearman Correlation Coefficients

	Performance Slider Task	**Performance Maths Task**	**Performance Grid Task**	**Performance Ind. Value Effort Task**
Anxiety before round 1	0.0114	-0.1837	-0.0951	-0.0140
Anxiety before round 6	-0.0033	-0.2499*	-0.1145	0.1760
Risk Aversion	-0.0739	-0.1667	0.0812	-0.1281
State-Anxiety	-0.1437	0.0479	-0.1654	0.0554
Trait-Anxiety	-0.3418**	0.0025	-0.0427	0.1520

* $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$

Table B.7: Summary Statistics on Trait-Anxiety - Measured at the Beginning of the Experiment

	Obs	Mean	Std. Dev	Min	Max
Slider Task	48	39.54	11.622	23	68
Maths Task	53	42.68	9.058	27	62
Grid Task	52	41.16	8.964	21	61.052
Ind. Value Effort Task	50	42.15	10.782	20	65.263

Table B.8: Error Clustering on Performance Using the Square of Anxiety Levels

	Slider Task		Maths Task		Grid Task		Ind. Value Task	
Anxiety Round 1	0.109*		-0.034		-0.161		-0.015	
Anxiety Square Round 1	-0.001		0.000		0.001		0.000	
Risk Aversion	0.133	-0.151	-0.137	-0.096	0.914**	0.451	-0.139	-0.163
Trait-Anxiety	-0.089	-0.105	0.015	0.017	-0.100	-0.054	0.041	0.033
State-Anxiety	-0.045	0.020	0.027	0.039	0.080	-0.010	0.008	0.006
Male	2.355**	3.118***	0.297	0.156	-0.025	-0.939	-0.831	-0.705
British Nationality	-2.059**	-2.475**	0.350	1.007	4.248**	4.422**	-0.543	-0.651
School of Eco	-3.443	-0.631	-0.537	-0.826	4.309	5.225*	0.267	-0.899
Age	-0.148	0.033	0.027	0.163	-0.339	-0.131	0.098	0.102
Anxiety Round 6		0.043		-0.044		-0.034		-0.024
Anxiety Square Round 6		-0.000		0.000		-0.000		0.001
Constant	23.884***	22.919***	3.895	1.392	0.2460	0.1707	0.981	1.259
Observations	235	235	265	265	260	260	245	245
R-Square	0.2659	0.2502	0.0932	0.1289	0.1345	0.2817	0.2003	0.2811
F - test	4.38	3.57	0.79	1.72	2.64	1.18	0.0801	0.1296

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The first columns of each task refer to observations of the first set of 5 rounds.
The second columns of each task refer to observations of the second set of 5 rounds.

Table B.9: Error Clustering on Performance with Interaction Terms with Anxiety Before the Task

	Slider Task	Maths Task	Grid Task	Ind. Value Task
Anxiety Before Round 1	-0.016	-0.007	-0.037	0.110
Risk Aversion	-0.188	-0.197	1.226**	-0.293
Trait-Anxiety	-0.143	0.056	0.279	0.132
State-Anxiety	-0.002	-0.002	-0.629***	0.003
Male	2.874	0.589	2.134	-0.096
British Nationality	-2.261**	0.409	2.887	0.265
School of Eco	-3.911	-0.581	3.604	0.234
Age	-0.092	0.021	-0.356*	0.198
Anxiety*Risk Aversion	0.008	0.001	-0.009	0.005
Anxiety*Trait Anxiety	-0.000	-0.001	-0.008*	-0.002
Anxiety*State Anxiety	0.000	0.000	0.012***	-0.000
Anxiety*Male	-0.007	-0.006	-0.027	-0.028
Constant	25.719***	3.264	18.551	-4.913
Observations	235	265	260	245
R-Square	0.2452	0.0948	0.3331	0.1216
F - test	2.88	0.80	5.44	2.14

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.10: Error Clustering on Performance with Interaction Terms with Anxiety Before Round Six

	Slider Task	Maths Task	Grid Task	Ind. Value Task
Anxiety Before Round 6	-0.060	-0.011	0.120	0.069
Risk Aversion	-0.094	0.114	1.394***	-0.234
Trait-Anxiety	-0.173	-0.008	0.387***	0.055
State-Anxiety	0.024	0.027	-0.521***	0.016
Male	4.339**	1.885	0.458	-0.409
British Nationality	-2.434**	0.952	4.952*	-0.418
School of Eco	-1.226	-0.778	3.604	-0.646
Age	0.011	0.142	-0.270	0.155
Anxiety*Risk Aversion	0.001	-0.003	-0.017*	0.004
Anxiety*Trait Anxiety	0.001	0.001	-0.010***	-0.001
Anxiety*State Anxiety	0.001	-0.000	0.010***	-0.001
Anxiety*Male	-0.024	-0.027	-0.040	-0.024
Constant	25.272***	0.729	11.627	-1.406
Observations	235	265	260	245
R-Square	0.2813	0.1511	0.3517	0.1316
F - test	3.39	0.80	3.33	1.49

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.11: OLS and GLM Estimates with Errors Clustering on Performance in the Slider Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	0.027	0.027		
	(1.51)	(1.54)		
Risk Aversion	0.147	0.147	-0.219	-0.219
	(0.63)	(0.64)	(-0.64)	(-0.65)
State-Anxiety	-0.006	-0.006	0.005	0.005
	(-0.06)	(-0.07)	(0.04)	(0.04)
Trait-Anxiety	-0.146**	-0.146**	-0.090	-0.090
	(-2.26)	(-2.30)	(-1.18)	(-1.20)
Male	2.540**	2.540***	3.268***	3.268***
	(2.59)	(2.63)	(3.11)	(3.16)
Age	-0.087	-0.087	0.020	0.020
	(-0.74)	(-0.75)	(0.16)	(0.17)
British Nationality	-2.164**	-2.164**	-2.630**	-2.630**
	(-2.41)	(-2.45)	(-2.50)	(-2.54)
School of ECO	-3.457	-3.457	-0.742	-0.742
	(-1.56)	(-1.58)	(-0.31)	(-0.32)
Anxiety Before Round 6			-0.000	-0.000
			(-0.01)	(-0.01)
Constant	24.029***	24.029***	24.031***	24.031***
	(5.65)	(5.75)	(5.48)	(5.58)
Observations	235	235	235	235
R-squared	0.2381		0.2438	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: GLM on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: GLM on effort of the second 5 set of rounds.

Table B.12: OLS and GLM Estimates with Errors Clustering on Performance in the Maths Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	-0.017	-0.017*		
	(-1.63)	(-1.66)		
Risk Aversion	-0.136	-0.136	-0.092	-0.092
	(-1.18)	(-1.20)	(-0.85)	(-0.87)
State-Anxiety	0.028	0.028	0.033	0.033
	(0.92)	(0.93)	(0.98)	(1.00)
Trait-Anxiety	0.013	0.013	0.015	0.015
	(0.47)	(0.48)	(0.46)	(0.47)
Male	0.298	0.298	0.202	0.202
	(0.70)	(0.71)	(0.44)	(0.44)
Age	0.029	0.029	0.144	0.144
	(0.38)	(0.38)	(1.57)	(1.59)
British Nationality	0.423	0.423	0.948	0.948
	(0.89)	(0.91)	(1.61)	(1.64)
School of ECO	-0.548	-0.548	-0.828	-0.828
	(-0.87)	(-0.88)	(-1.41)	(-1.44)
Anxiety Before Round 6			-0.026***	-0.026***
			(-3.02)	(-3.06)
Constant	3.511	3.511	1.686	1.686
	(1.52)	(1.54)	(0.64)	(0.65)
Observations	265	265	265	265
R-squared	0.0891		0.1264	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: GLM on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: GLM on effort of the second 5 set of rounds.

Table B.13: OLS and GLM Estimates with Errors Clustering on Performance in the Grid Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	-0.061**	-0.061**		
	(-2.12)	(-2.15)		
Risk Aversion	0.840**	0.840**	0.459	0.459
	(2.07)	(2.10)	(1.17)	(1.19)
State-Anxiety	0.083	0.083	-0.011	-0.011
	(0.85)	(0.86)	(-0.11)	(-0.11)
Trait-Anxiety	-0.096	-0.096	-0.051	-0.051
	(-1.00)	(-1.01)	(-0.51)	(-0.52)
Male	-0.112	-0.112	-0.914	-0.914
	(-0.07)	(-0.07)	(-0.55)	(-0.56)
Age	-0.246	-0.246	-0.137	-0.137
	(-1.12)	(-1.14)	(-0.57)	(-0.58)
British Nationality	4.618**	4.618**	4.447**	4.447**
	(2.40)	(2.44)	(2.21)	(2.25)
School of ECO	3.552	3.552	5.263*	5.263*
	(1.27)	(1.29)	(1.87)	(1.90)
Anxiety Before Round 6			-0.046*	-0.046*
			(-1.72)	(-1.74)
Constant	11.855	11.855	13.794*	13.794*
	(1.58)	(1.60)	(1.70)	(1.73)
Observations	260	260	260	260
R-squared	0.2321		0.1704	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: GLM on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: GLM on effort of the second 5 set of rounds.

Table B.14: OLS and GLM Estimates with Errors Clustering on Performance in the Induced Value Effort Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	0.007	0.007		
	(0.35)	(0.35)		
Risk Aversion	-0.133	-0.133	-0.171	-0.171
	(-0.69)	(-0.71)	(-0.75)	(-0.76)
State-Anxiety	0.006	0.006	-0.004	-0.004
	(0.13)	(0.13)	(-0.06)	(-0.07)
Trait-Anxiety	0.044	0.044	0.044	0.044
	(0.73)	(0.74)	(0.73)	(0.74)
Male	-0.777	-0.777	-0.691	-0.691
	(-0.86)	(-0.87)	(-0.78)	(-0.79)
Age	0.113	0.113	0.134	0.134
	(1.04)	(1.06)	(1.25)	(1.27)
British Nationality	-0.511	-0.511	-0.634	-0.634
	(-0.48)	(-0.49)	(-0.52)	(-0.53)
School of ECO	0.302	0.302	-0.719	-0.719
	(0.22)	(0.23)	(-0.68)	(-0.69)
Anxiety Before Round 6			0.023	0.023
			(1.47)	(1.50)
Constant	0.308	0.308	0.101	0.101
	(0.07)	(0.07)	(0.02)	(0.02)
Observations	245	245	245	245
R-squared	0.0773		0.1180	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: GLM on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: GLM on effort of the second 5 set of rounds.

Table B.15: OLS with Errors Clustering and Random Effects Estimates on Performance in the Slider Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	0.027	0.027		
	(1.51)	(1.51)		
Risk Aversion	0.147	0.147	-0.219	-0.219
	(0.63)	(0.63)	(-0.64)	(-0.64)
State-Anxiety	-0.006	-0.006	0.005	0.005
	(-0.06)	(-0.06)	(0.04)	(0.04)
Trait-Anxiety	-0.146**	-0.146**	-0.090	-0.090
	(-2.26)	(-2.26)	(-1.18)	(-1.18)
Male	2.540**	2.540***	3.268***	3.268***
	(2.59)	(2.59)	(3.11)	(3.11)
Age	-0.087	-0.087	0.020	0.020
	(-0.74)	(-0.74)	(0.16)	(0.16)
British Nationality	-2.164**	-2.164**	-2.630**	-2.630**
	(-2.41)	(-2.41)	(-2.50)	(-2.50)
School of ECO	-3.457	-3.457	-0.742	-0.742
	(-1.56)	(-1.56)	(-0.31)	(-0.31)
Anxiety Before Round 6			-0.000	-0.000
			(-0.01)	(-0.01)
Constant	24.029***	24.029***	24.031***	24.031***
	(5.65)	(5.65)	(5.48)	(5.48)
Observations	235	235	235	235
R-squared	0.2381		0.2438	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: Random Effects on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: Random Effects on effort of the second 5 set of rounds.

Table B.16: OLS with Errors Clustering and Random Effects Estimates on Performance in the Maths Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 6	-0.017	-0.017		
	(-1.63)	(-1.63)		
Risk Aversion	-0.136	-0.136	-0.092	-0.092
	(-1.18)	(-1.18)	(-0.85)	(-0.85)
State-Anxiety	0.028	0.028	0.033	0.033
	(0.92)	(0.92)	(0.98)	(0.98)
Trait-Anxiety	0.013	0.013	0.015	0.015
	(0.47)	(0.47)	(0.46)	(0.46)
Male	0.298	0.298	0.202	0.202
	(0.70)	(0.70)	(0.44)	(0.44)
Age	0.029	0.029	0.144	0.144
	(0.38)	(0.38)	(1.57)	(1.57)
British Nationality	0.423	0.423	0.948	0.948
	(0.89)	(0.89)	(1.61)	(1.61)
School of ECO	-0.548	-0.548	-0.828	-0.828
	(-0.87)	(-0.87)	(-1.41)	(-1.41)
Anxiety Before Round 6			-0.026***	-0.026***
			(-3.02)	(-3.02)
Constant	3.511	3.511	1.686	1.686
	(1.52)	(1.52)	(0.64)	(0.64)
Observations	265	265	265	265
R-squared	0.0891		0.1264	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: Random Effects on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: Random Effects on effort of the second 5 set of rounds.

Table B.17: OLS with Errors Clustering and Random Effects Estimates on Performance in the Grid Task

	Model 1	Model 2	Model 3	Model 4
Anxious Period 1	-0.061**	-0.061**		
	(-2.12)	(-2.12)		
Risk Aversion	0.840**	0.840**	0.459	0.459
	(2.07)	(2.07)	(1.17)	(1.17)
State-Anxiety	0.083	0.083	-0.011	-0.011
	(0.85)	(0.85)	(-0.11)	(-0.11)
Trait-Anxiety	-0.096	-0.096	-0.051	-0.051
	(-1.00)	(-1.00)	(-0.51)	(-0.51)
Male	-0.112	-0.112	-0.914	-0.914
	(-0.07)	(-0.07)	(-0.55)	(-0.55)
Age	-0.246	-0.246	-0.137	-0.137
	(-1.12)	(-1.12)	(-0.57)	(-0.57)
British Nationality	4.618**	4.618**	4.447**	4.447**
	(2.40)	(2.40)	(2.21)	(2.21)
School of ECO	3.552	3.552	5.263*	5.263*
	(1.27)	(1.27)	(1.87)	(1.87)
Anxiety Before Round 6			-0.046*	-0.046*
			(-1.72)	(-1.72)
Constant	11.855	11.855	13.794*	13.794*
	(1.58)	(1.58)	(1.70)	(1.70)
Observations	260	260	260	260
R-squared	0.2321		0.1704	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: Random Effects on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: Random Effects on effort of the second 5 set of rounds.

Table B.18: OLS with Errors Clustering and Random Effects Estimates on Performance in the Induced Value Effort Task

	Model 1	Model 2	Model 3	Model 4
Anxiety Before Round 1	0.007	0.007		
	(0.35)	(0.35)		
Risk Aversion	-0.133	-0.133	-0.171	-0.171
	(-0.69)	(-0.69)	(-0.75)	(-0.75)
State-Anxiety	0.006	0.006	-0.004	-0.004
	(0.13)	(0.13)	(-0.06)	(-0.06)
Trait-Anxiety	0.044	0.044	0.044	0.044
	(0.73)	(0.73)	(0.73)	(0.73)
Male	-0.777	-0.777	-0.691	-0.691
	(-0.86)	(-0.86)	(-0.78)	(-0.78)
Age	0.113	0.113	0.134	0.134
	(1.04)	(1.04)	(1.25)	(1.25)
British Nationality	-0.511	-0.511	-0.634	-0.634
	(-0.48)	(-0.48)	(-0.52)	(-0.52)
School of ECO	0.302	0.302	-0.719	-0.719
	(0.22)	(0.22)	(-0.68)	(-0.68)
Anxiety Before Round 6			0.023	0.023
			(1.47)	(1.47)
Constant	0.308	0.308	0.101	0.101
	(0.07)	(0.07)	(0.02)	(0.02)
Observations	245	245	245	245
R-squared	0.0773		0.1180	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model 1: OLS on effort of the first 5 set of rounds.

Model 2: Random Effects on effort of the first 5 set of rounds.

Model 3: OLS on effort of the second 5 set of rounds.

Model 4: Random Effects on effort of the second 5 set of rounds.

Table B.19: Summary Statistics on State-Anxiety - Measured at the Beginning of the Experiment

	Obs	**Mean**	**Std. Dev**	**Min**	**Max**
Slider Task	47	34.26	8.889	20	65
Maths Task	54	37.31	9.355	22	62
Grid Task	52	33.69	9.336	21	61.052
Ind. Value Effort Task	49	35.76	10.071	20	70

Figures

Figure B.1: Slider Task

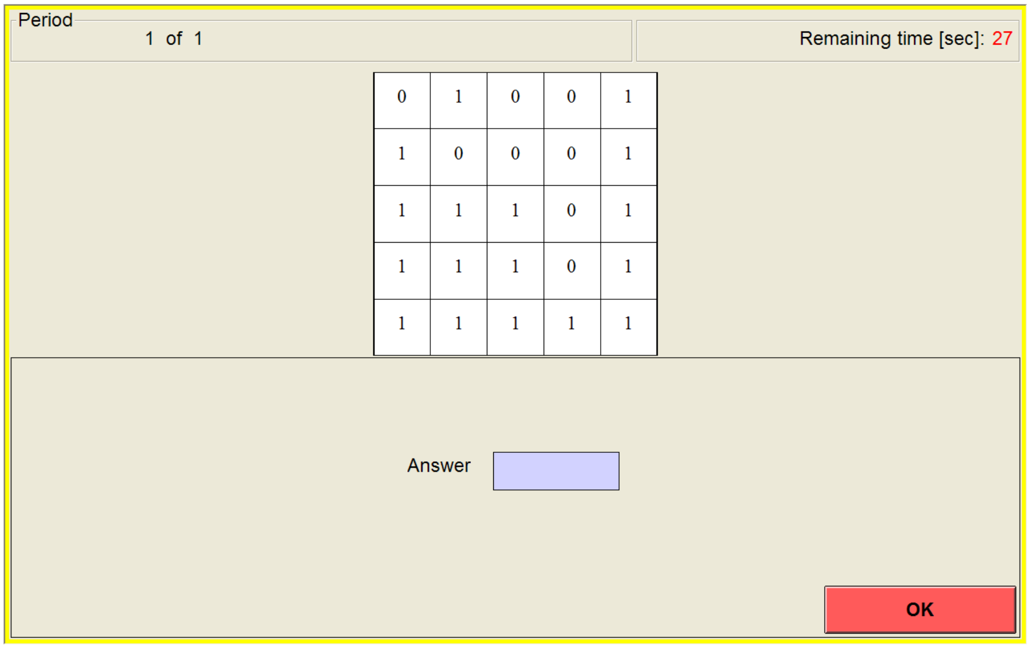


Figure B.2: Grid Task

Figure B.3: Performance Distributions

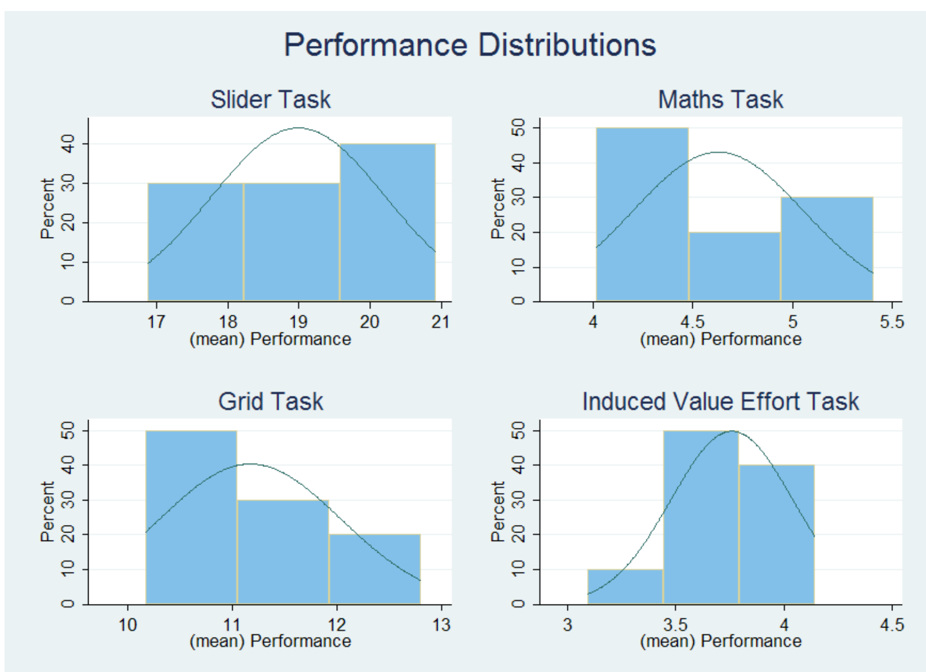


Figure B.4: Performance Box Plot

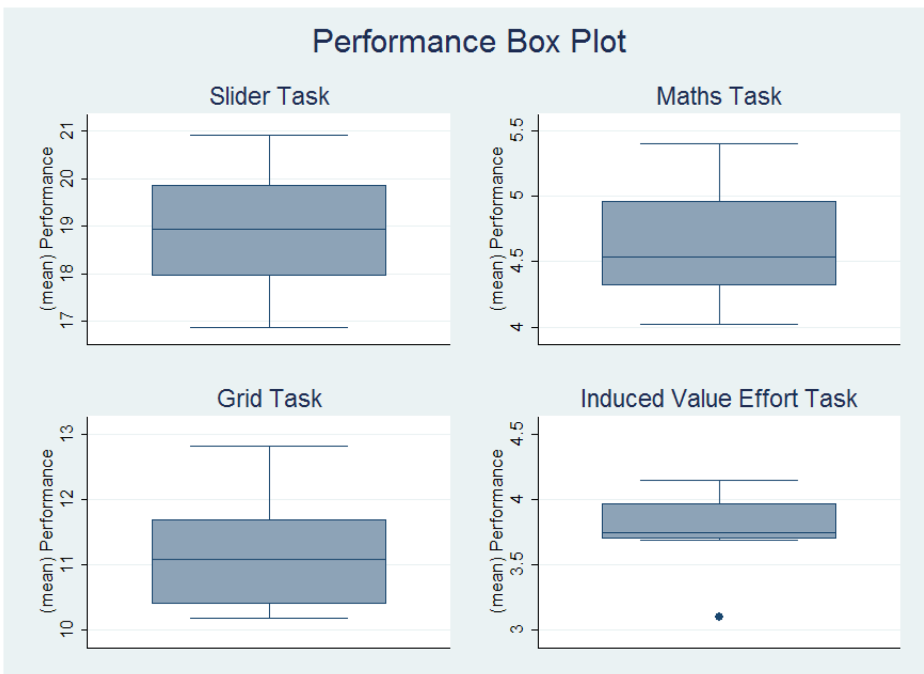


Figure B.5: State-Anxiety Score Distributions

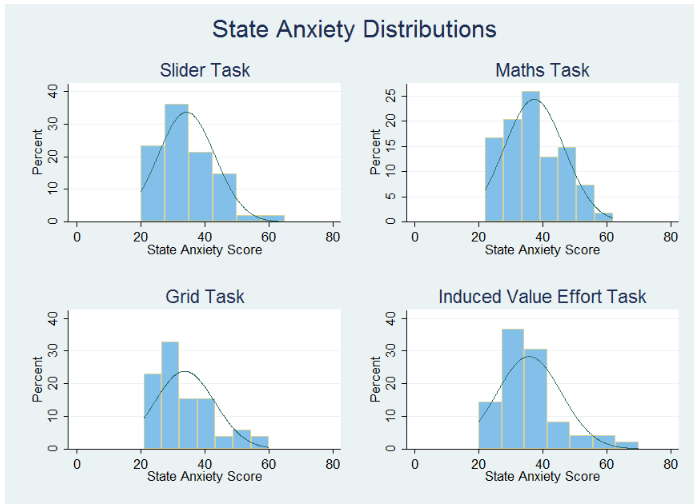


Figure B.6: Emotions Elicited at the End of Round 10 - Values in Percentage

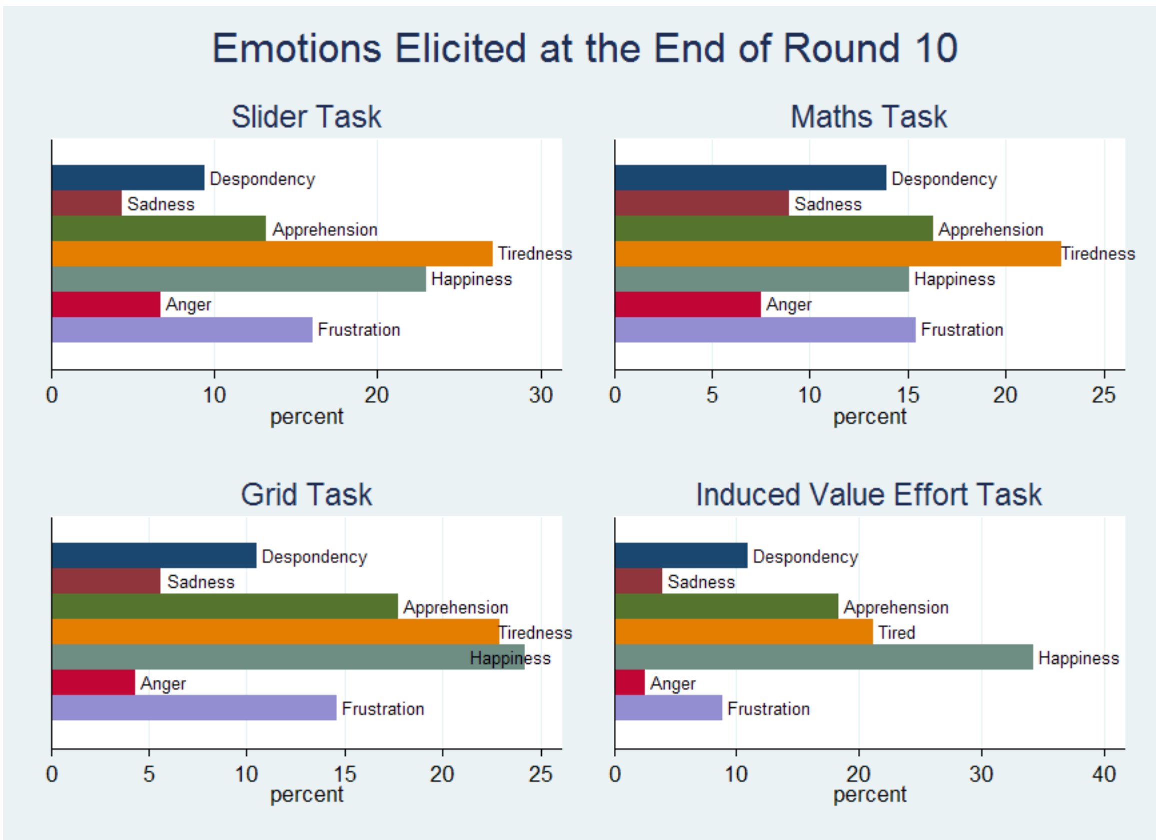


Figure B.7: Emotions Elicited Before Round 1 - Values in Percentage

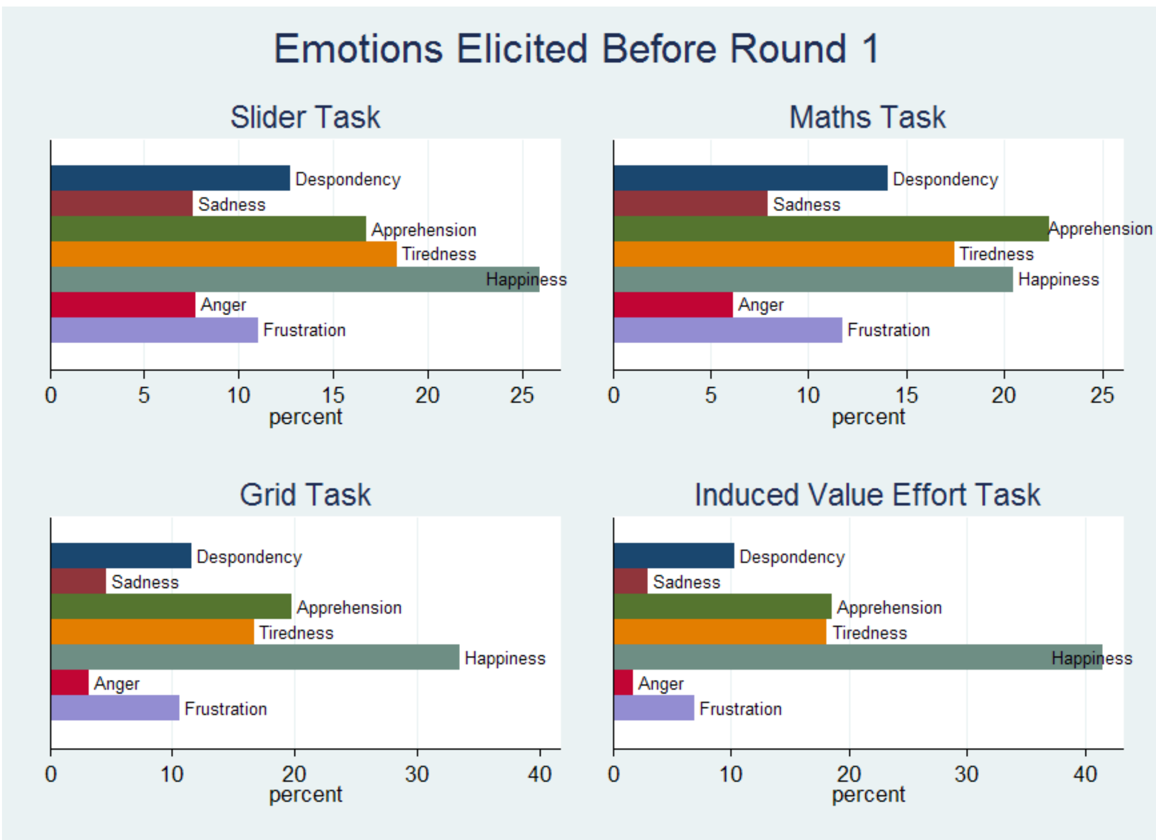


Figure B.8: Emotions Elicited Before Round 6 - Values in Percentage

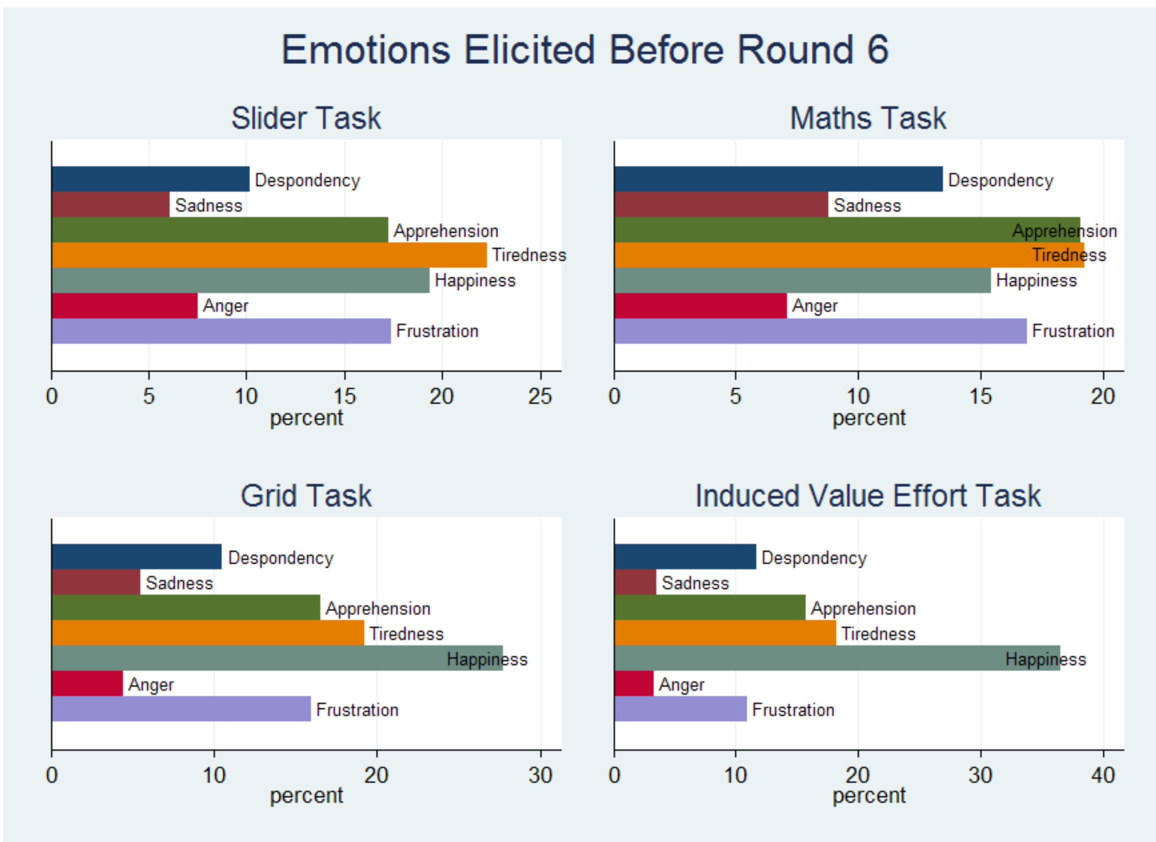
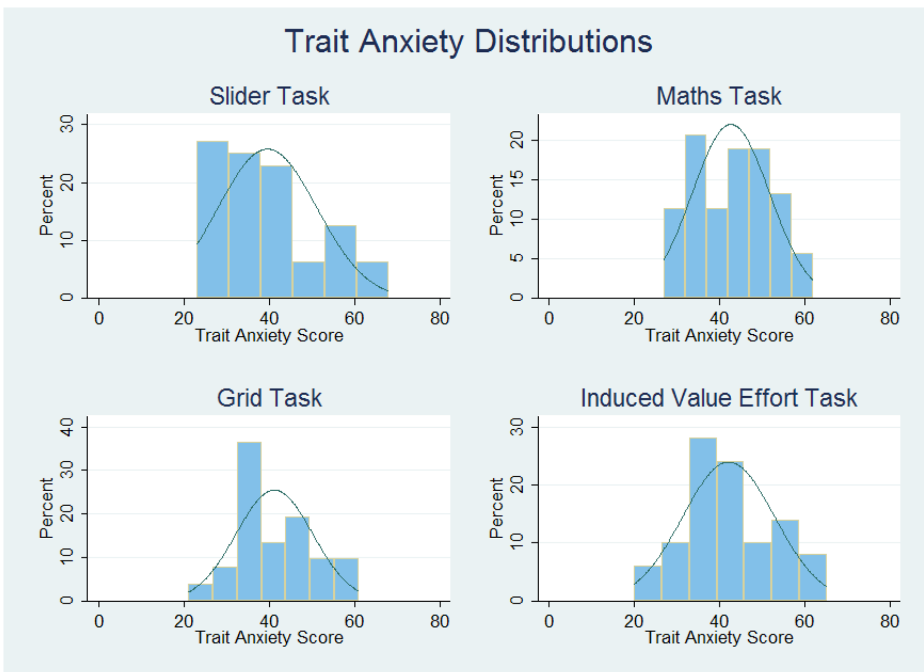


Figure B.9: Trait-Anxiety Score Distributions



Appendix C

(appendix Chapter 4)

C.1 Rank-order Tournaments

Effort in Rank-Order Tournaments

In rank-order tournaments an agent's payment depends only on the rank of his or her performance and not on either the absolute level of performance or the size of the differences in performance across agents[1].

Lazear and Rosen (1981) will consider a two-player symmetric tournament. The tournament is symmetric when players are identical (*homogeneous*), that is when they have the same utility function and same cost of effort. Moreover, they assume agents to be risk-neutral.

Therefore agents i and j have the following **utility function *U* separable** in the payment received and the effort exerted:

$$U_i(p, e) = U_j(p, e) = u(p) - c(e)$$

where p is the non-negative payment to the agent and e is the agent's non-negative effort. $u(.)$ and $c(.)$ are both positive and increasing functions, **concave** and **convex** respectively.

Agent j provides a level of effort that is *not observable* by the principal and that is converted in output according to the following production function:

[1]Rank-order tournaments have the structure of a contest as contestants expend effort (we will call the agents' observable effort performance or output) in trying to win a prize, and the prize is awarded as a function of these efforts.

$$q_j = e_j + \epsilon_j$$

where e_j is the level of investment or effort, a measure of skill or average output of agent j, and ϵ_j is a random component or luck drawn out of a known distribution with zero mean and variance σ^2. Agent i has a similar technology to agent j and simultaneously makes a decision on her/his level of investment[2].

Let us assume that there are two fixed prizes W_1 and W_2 where W_1 and $W_2 > 0$ and $W_1 > W_2$. The winner of the contest is determined by the largest production q. W_1 is the prize that goes to the winner and W_2 is the prize that goes to the loser of the contest.

A contestant k's expected utility is:

$$EU_k = P\left[W_1 - C\left(e_k\right)\right] + (1-P)\left[W_2 - C\left(e_k\right)\right] = PW_1 + (1-P)\,W_2 - C\left(e_k\right) \quad \text{(C.1)}$$

with $k = i, j$ and where P is the probability of winning.

The probability that j wins the winning prize is

$$\begin{aligned} P = prob\left(q_j > q_i\right) &= prob\left(e_j + \epsilon_j > e_i + \epsilon_i\right) = prob\left(e_j - e_i > \epsilon_i - \epsilon_j\right) = \\ prob\left(e_j - e_i > \epsilon\right) &= G\left(e_j - e_i\right) \end{aligned} \quad \text{(C.2)}$$

where $\epsilon \equiv \epsilon_i - \epsilon_j$ where ϵ_i and ϵ_j are i.i.d. The random variable ϵ follows a pdf g, $\epsilon \sim g\left(\epsilon\right)$, with mean zero (E($\epsilon$)=0) and variance $2\sigma^2$ ($E(\epsilon^2 = 2\sigma^2)$. $G\left(\cdot\right)$ is the cdf of ϵ.

[2]The agent's (lifetime) output is a random variable whose distribution is controlled by the agent himself. In particular, the agent is allowed to control the mean of the distribution by investing in costly skills prior to entering the competition. However, a given productivity realization also depends on a random factor which is beyond anyone's control. Contestants pre-commit their investments knowing the tournament prizes and the rules of the game, but they do not communicate each other or collude. Moreover, productivity risk is non-diversifiable by the agent himself. This is one reason for choosing a long period (agent's lifetime is considered a single period in the tournament game): ability effect cannot be diversified when it is not possible to discover it quickly (managerial talent, for example). If the period were short and the random factor was independently distributed across periods, the agent could diversify per period to balance off good and bad years. However, it is assumed that ϵ is i.i.d. across individuals, so that owners of firms can diversify risk by pooling workers together or by holding a portfolio.

Each contestant chooses e_k to maximize C.1. Assuming interior solutions, this imply the following first and second order conditions (F.O.C. and S.O.C.):

F.O.C.

$$\frac{\partial EU}{\partial e_k} = 0 \Rightarrow (W_1 - W_2)\frac{\partial P}{\partial e_k} - C'(e_k) = 0 \tag{C.3}$$

S.O.C.

$$\frac{\partial^2 EU}{\partial e_k^2} < 0 \Rightarrow (W_1 - W_2)\frac{\partial^2 P}{\partial e_k^2} - C''(e_k) < 0 \tag{C.4}$$

with $k = i, j$.

Player j takes e_i as given in determining his effort and conversely for i (Nash-Cournot assumptions). Moreover, we know that:

$$\frac{\partial P}{\partial e_j} = \partial G(e_j - e_i) / \partial e_j = g(e_j - e_i) \tag{C.5}$$

Therefore, the FOC can be written as follow:

$$(W_1 - W_2)\, g(e_j - e_i) - C'(e_j) = 0 \tag{C.6}$$

Equation C.5 is the j's reaction function. Player i's reaction function is symmetrical to C.5.

Symmetry implies that when the Nash equilibrium exists:

$$e^* = e_j = e_i$$

and P = G(0) = $\frac{1}{2}$. The outcome is purely random in equilibrium; players whose investment is identical, have the same winning probability. The winner will be randomly determined in equilibrium. However, each player affects his probability of winning by investing.

At the Nash equilibrium equation C.6 reduces to

$$C'(e_i) = (W_1 - W_2)\, g(0) \tag{C.7}$$

because $e_j = e_i$. Players' investment depend on the spread between winning and loosing prizes: the higher the spread, the more they will invest. Levels of the prizes only influence the decision to enter the game, which requires non-negative expected wealth.

However, it is important to remember that it is not necessarily true that there is a solution because with arbitrary density functions the objective function may not be concave in the relevant range. A pure strategy solution exists if σ^2 is sufficiently large. Contests are feasible only when chance is a significant factor.

Specification of the Model

With suitable restrictions on the distribution of the random term and the utility functions, a unique pure strategy Nash equilibrium will exist for the game. Testing the theory in the lab requires the specification of the utility function, the production function, the distribution of ϵ, and the prizes W_1 and W_2.

One simple simplification is the following:

$$U_k(p_k, e_k) = p_k - \frac{e_k^2}{c}$$

and

$$q_k = e_k + \epsilon_k$$

where $k = i, j$, $c > 0$, and ϵ_k is distributed uniformly over the interval $[-a, a]$ where $a > 0$, and independently across the agents. We, moreover, restrict agents' effort to lie in the interval $[b, d]$.

The agent k' expected utility, as previously shown, is

$$EU_k = P\left[W_1 - C\left(e_k\right)\right] + (1 - P)\left[W_2 - C\left(e_k\right)\right]$$

that can be written as

$$W_2 + P\left(W_1 - W_2\right) - C\left(e_k\right)$$

with $k = i, j$. If a pure strategy Nash equilibrium exists, it will be symmetric $e_i = e_j = e^*$. If the equilibrium is interior of $[b, d]$, each agent's first and second order conditions must be fulfilled:

$$\frac{\partial EU}{\partial e_k} = 0 \Rightarrow (W_1 - W_2) \frac{\partial P}{\partial e_k} - C'(e_k) = 0 \quad \text{(C.8)}$$

with $k = i, j$. Substituting the first order derivative of the cost of effort function and C.5 into C.8 we have:

$$\frac{\partial EU}{\partial e_k} = (W_1 - W_2) g(e_j - e_k) - \frac{2e^*}{c} = 0 \quad \text{(C.9)}$$

Given the distributional assumptions on the random term the condition C.9 can be rewritten as

$$(W_1 - W_2) \left(\frac{1}{2a} \right) - \frac{2e^*}{c} = 0 \quad \text{(C.10)}$$

where a is the size of the interval of the random term. We can thus write explicitly the expression for the equilibrium value

$$e^* = \frac{(W_1 - W_2) c}{4a} \quad \text{(C.11)}$$

The concavity of the agent's utility function ensures that C.8 is sufficient for a maximum. Moreover, the second order condition is fulfilled:

$$\frac{\partial^2 EU}{\partial e_k^2} < 0 \Rightarrow (W_1 - W_2) \frac{\partial^2 P}{\partial e_k^2} - C''(e_k) < 0 \Rightarrow 0 - \frac{2}{c} < 0 \quad \text{(C.12)}$$

This holds because c is a strictly positive scalar.

Therefore, the efforts chosen by agents in the rank-order tournaments will proportionally increase with any increase in the spread between the prizes and in the cost of effort c, while they will move inversely with both the size of the uniform distribution of the random term[3].

[3]The more important luck is in determining the outcome of the contest, the lower is effort.

Effort and Sabotage in Rank-Order Tournaments

Let us assume now that the agents' production functions are as follows:

$$q_i = f(e_i, s_j) + \epsilon_i$$

$$q_j = f(e_j, s_i) + \epsilon_j$$

The term s_j is the j's '**sabotage**' inflicted on the other player[4]. The term sabotage is used to indicate any costly action that one agent takes and that negatively affects the output of another agent. The aggressive behaviour has therefore the direct effect of output reduction. It is not clear, however, that the final effect is output reduction. The outcome of this relative performance contest is indeed affected by choices of effort and sabotage; thus it is possible that agents' effort is sufficiently larger to make up for the lost output through sabotage[5]. This is the so-called *identification issue*: the observation of agents' final outcome does not allow to clearly identify between the agent's effort and the sabotage choice[6].

In his model Lazear (1989) considers a game which consists of 2 symmetric agents i and j; agent i wishes to maximize his expected utility:

$$max_{e_i, s_i} EU = P(e_i, s_i; e_j, s_j) W_1 + [1 - P(.)] W_2 - C(e_i, s_i) \tag{C.13}$$

where $P(e_i, s_i; e_j, s_j)$ is the probability that agent i wins conditional on his choice of effort and sabotage. Hence, i wins the winning prize if $q_i > q_j$ and the winning probability is given by

$$P(e_i, s_i; e_j, s_j) = prob(q_i > q_j) = prob[f(e_i, s_j) - f(e_j, s_i) > \epsilon_j - \epsilon_i] = G[f(e_i, s_j) - f(e_j, s_i)] \tag{C.14}$$

[4]The other terms of the production function have the same characteristics described in the previous paragraph.

[5]It is clear however that output always decreases when the cost of sabotage falls.

[6]For this reason an experiment which consists of treatments that do not allow for sabotage would help in the analysis of agents' behaviour.

where $G(.)$ is the distribution function of the random variable $\epsilon_j - \epsilon_i$.

The first order conditions (F.O.C.) to agent i's maximization problem (C.13) are:

$$(W_1 - W_2) \tfrac{\partial P}{\partial e_i} = C_1(e_i, s_i)$$

and

$$(W_1 - W_2) \frac{\partial P}{\partial s_i} = C_2(e_i, s_i) \tag{C.15}$$

or

$$(W_1 - W_2)\, g\left[f(e_i, s_j) - f(e_j, s_i)\right] f_1(e_i, s_j) = C_1(e_i, s_i) \tag{C.16}$$

$$(W_1 - W_2)\, g\left[f(e_i, s_j) - f(e_j, s_i)\right] f_2(e_j, s_i) = -C_2(e_i, s_i) \tag{C.17}$$

Agent j solves the corresponding problem. Since players are identical, in equilibrium, $e_i = e_j$ and $s_i = s_j$. Therefore, the solution is given by the following four[7] first order conditions:

$$\begin{gathered}
(W_1 - W_2)\, g(0)\, f_1(e_i, s_j) = C_1(e_i, s_i) \\
(W_1 - W_2)\, g(0)\, (-)\, f_2(e_j, s_i) = C_2(e_i, s_i) \\
(W_1 - W_2)\, g(0)\, f_1(e_j, s_i) = C_1(e_j, s_j) \\
(W_1 - W_2)\, g(0)\, (-)\, f_2(e_i, s_j) = C_2(e_j, s_j)
\end{gathered}$$

The existence of a unique interior solution is not guaranteed unless the second order conditions[8] are fulfilled as well:

$$(W_1 - W_2)\, g(0)\, f_{11} - C_{11} < 0$$

and

[7] Two per each player.
[8] Both per each agent.

$$(W_1 - W_2)\, g\,(0)\, f_{22} - C_{22} < 0$$

Moreover, to ensure a unique interior solution there must be enough dispersion of the random term.

Finally, the agent has also a participation constraint. Under the symmetrical Nash equilibrium, each agent has a probability of winning that match equal to $G(0) = \frac{1}{2}$. This implies that in equilibrium an agent's expected utility is

$$EU = \frac{W_1+W_2}{2} - C(e^*, s^*)$$

The agent will participate in the contest (assuming that the outside option is 0) as long as his expected utility is positive

$$\frac{W_1+W_2}{2} >= C(e^*, s^*)$$

This guarantees that an interior solution exists and that agents have no incentive to deviate to zero.

Specification of the Model with Sabotage

Let us assume that the random variable is uniformly distributed over the interval $[-b, b]$ and that the effort cost function and sabotage cost function are respectively
$C_e = \frac{e^2}{c_e}$
$C_s = \frac{s^2}{c_s}$

We know that[9]

$$\frac{\partial P}{\partial e_i} = \frac{1}{4b}$$

and

$$\frac{\partial P}{\partial s_i} = \frac{1}{4b}$$

[9]See the proof in the appendix in Harbring and Irlenbusch (2008).

Thus our FOC (C.1 and C.15) for agent i will reduce to

$(W_1 - W_2)\,\frac{1}{2b} = \frac{2e_i}{c_e}$

and

$(W_1 - W_2)\,\frac{1}{2b} = \frac{2s_i}{c_s}$

from which we obtain the effort and sabotage level in equilibrium

$$e^* = (W_1 - W_2)\,\frac{c_e}{4b}$$

and

$$s^* = (W_1 - W_2)\,\frac{c_s}{4b}$$

Note that an increase in the prize spread leads to higher effort as well as sabotage activities and that the marginal cost of effort and sabotage are equal in equilibrium (Harbring and Irlenbusch, 2008, page 7).

Parametrization

Experimental parameters' choice is restricted by the above simplifications:

$e \in [0, 100]$

$s \in [0, 50]$

The range for sabotage value follows the experimental literature in sabotage in contests (Harbring and Irlenbusch, 2005, 2003, 2008).

We will use the same range of the random term in treatments with and without sabotage. $\epsilon \in [-60, 60]$

The constant c_e and c_s are respectively 3 and 2. Therefore we have the following total cost:

$$C\,(e, s) = e^2/76.8 + s^2/21.6$$

Sabotaging is more expensive. An aggressive behaviour could be punished and the agent actually exposes herself to this risk. Therefore, imposing higher costs of sabotaging is an indicator of potential risk and related negative consequences.

We impose for convenience the following prizes values:

$$W_1 = 350$$
$$W_2 = 250$$

The spread of 100 points is convenient for calculations. Beside, the relative proportion of prizes follows the literature (Moldovanu and Sela, 2001; Bull et al., 1987).

The effort level in equilibrium, assuming maximizing agents, is 32 and the sabotage level in equilibrium is 9.

C.2 Experimental Instructions

TASK 2 - INSTRUCTIONS

In this task you are matched with another participant in this room. We will call him or her **participant X**.

You and participant X will be asked to make some decisions. Your decisions are made anonymously and you will have no information identifying who you are matched with.

Task 2 consists of a **PRIZE COMPETITION**: one of you will get the **Winner Prize** and the other will get the **Loser Prize**.

The Winner Prize is **350 Experimental Points**.

The Loser Prize is **250 Experimental Points**.

In this task you and participant X will gain experimental points. At the end of the experiment your experimental points will be converted by the computer into pounds at the rate of **2 pence for each point**.

For example, if you have 100 Experimental Points it means that you have earned 2 pounds.

HOW CAN I WIN THE WINNER PRIZE?

To win the Winner Prize **your Total Experimental Points must be higher than participant X's Total Experimental Points**.
Otherwise, you will get the Loser Prize.

Therefore:

If your Total Experimental Points are **HIGHER** than participant X' Total Experimental Points then you will get the **Winner Prize**.

If your Total Experimental Points are **LOWER** than participant X' Total Experimental Points then you will get the **Loser Prize**.

If your Total Experimental Points are **EQUAL TO** participant X' Total Experimental Points then you will get either the **Winner Prize** or the **Loser Prize** with a 50% chance.

HOW DO I GET MY TOTAL EXPERIMENTAL POINTS?

Your Total Experimental Points and participant X's Total Experimental Points are determined as follows:

Your Total Experimental Points = Your choice of Number A - Number B chosen by participant X + your Random Number

Participant X's Total Experimental Points = his/her choice of Number A - your choice of Number B + her/his Random Number

We are going now to explain each component of your Total Experimental Points and participant X's Total Experimental Points.

Number A

The first component of your Total Experimental Points is a number that we will call **Number A**.

The Number A can be any whole number **between 0 and 100**. **You will be asked to choose one number between 0 and 100 for Number A**.

Each Number A is associated with a cost. The higher the number the higher the cost is. You can find the list of costs associated with each Number A in table 1 *"Number A and Costs"*.

The cost associated with the Number A you will choose will be deducted from your prize.

Participant X will be simultaneously asked to choose her/his Number A. She/he will be provided with the same list of numbers and related costs.

[Here Table 1. "Number A and Costs" was inserted]

Number B

The second component of your Total Experimental Points is a number that we will call **Number B**.

The Number B can be any whole number **between 0 and 50**.

Participant X can reduce your Total Experimental Points by a Number B. **Participant X will be asked to choose one number between 0 and 50 for Number B**

.

Each Number B is associated with a cost. The higher the number the higher the cost is.

You can reduce participant X's Total Experimental Points as well by a Number B. You will be provided with the same list of numbers and related costs.

You can find the list of costs associated with each Number B in table 2 *"Number B and Costs"*.

The cost of your choice of Number B will be deducted from your prize.

[Here Table 2. "Number B and Costs" was inserted]

The Random Number

The third component of your Total Experimental Points is a **Random Number**. After you have chosen your A and B numbers the computer will select a personal random number.

This number can be any whole number **between - 60 and + 60**.

- 60, - 59, - 58, - 57, - 56, ..., -2, -1, 0, + 1, + 2, ..., + 56, + 57, + 58, + 59, + 60.

Each number is equally likely to be drawn.

There is one separate and independent random draw for you and one for participant X. This means that either you can be assigned the same random number as participant X or a different one.

Please note that you will be asked to choose your A and B numbers

simultaneously with participant X. Once you both have made your decisions, the computer will draw the random numbers.. You will know your Total Experimental Points and participant X's Total Experimental Points only at the end of the experiment.

Here is a reminder of how your Total Experimental Points and participant X's Total Experimental Points are determined:

Your Total Experimental Points = Your choice of Number A - Number B chosen by participant X + your Random Number

Participant X's Total Experimental Points = his/her choice of Number A - your choice of Number B + her/his Random Number

And a reminder of how you can win the prize:

If your Total Experimental Points are **HIGHER** than participant X' Total Experimental Points then you will get the **Winner Prize.**

If your Total Experimental Points are **LOWER** than participant X' Total Experimental Points then you will get the **Loser Prize.**

If your Total Experimental Points are **EQUAL TO** participant X' Total Experimental Points then you will get either the **Winner Prize** or the **Loser Prize** with a 50% chance.

MY EARNINGS FROM TASK 2

Your earnings in experimental points from Task 2 are determined as follows:

Your earnings = the Prize you won - cost of your Number A - costs of your Number B

C.3 Task 6 - Questionnaire

The following 48 items have been used in Experiment 1:

Nine items from the original scale have been used (Cohen-Charash, 2009). Items have been adapted.

1. After I found out the allocation of the bonus I felt some hatred.
2. After I found out the allocation of the bonus I felt some resentment against participant X.
3. After I found out the allocation of the bonus I felt some rancour against participant X.
4. After I found out the allocation of the bonus I felt bitter.
5. After I found out the allocation of the bonus I felt irritated and annoyed.
6. I desired the bonus that participant X has.
7. After I found out the allocation of the bonus I felt I lacked some of the things that participant X has.
8. Task 2 (the task in which I chose my A and B numbers) seems to be going better for participant X.
9. After I found out the allocation of the bonus I felt envious.

Items have been created ad hoc to measure the Model by (Fiske et al., 2002).

10. I think participant X is competent.
11. After I found out the allocation of the bonus I felt warm towards participant X.
12. After I found out the allocation of the bonus I felt positive towards participant X.
13. I think participant X's success to Task 1 (the task in which I counted the number of ones) is due to luck.
14. I think participant X's success to Task 1 (the task in which I counted the number

of ones) is due to skill.

Three items for competitive feelings toward the other have been adapted (Cohen-Charash, 2009).

15. Task 2 (the task in which I chose my A and B numbers) between participant X and I was competitive.
16. I was concerned to outperform participant X in Task 2 (the task in which I chose my A and B numbers).
17. I was concerned to maximize my own earnings relative to X's earnings in Task 2 (the task in which I chose my A and B numbers).

Five items out of six on subjective injustice beliefs have been used. Items have been adapted (Smith et al., 1994).

18. After I found out the allocation of the bonus I felt unfairly treated.
19. After I found out the allocation of the bonus I felt resentment over the unfairness of my treatment.
20. It seemed unfair that participant X started Task 2 (the task in which I chose my A and B numbers) with certain advantages over me.
21. It seemed unfair that the good fortune of participant X came naturally to him/her.
22. It seemed unfair that participant X had an advantage over me because of luck.

Three items out of four on inferior beliefs have been used. Items have been adapted (Smith et al., 1994).

23. Participant X made me feel inferior.
24. In Task 1 (the task in which I counted the number of ones) I was aware of my

inferior qualities.

25. The discrepancy in performance in Task 1 (the task in which I counted the number of ones) between participant X and I was due to my own inferior qualities.

Seven items on hostile and depressive feelings. All the items of the original scale have been used. Items have been adapted (Smith et al., 1994).

26. After I found out the allocation of the bonus I felt anger against participant X.

27. After I found out the allocation of the bonus I felt hostile feelings against participant X.

28. After I found out the allocation of the bonus I hated participant X.

29. After I found out the allocation of the bonus I disliked participant X.

30. After I found out the allocation of the bonus I felt in low spirits.

31. After I found out the allocation of the bonus I felt that I lacked energy.

32. Participant X made me feel depressed in Task 2 (the task in which I chose my A and B numbers).

All three items of the original scale on objective injustice have been used. Items have been adapted (Smith et al., 1994).

33. An objective judge would agree that participant X did not deserve the bonus.

34. Anyone would agree that participant X's advantage was unfairly obtained.

35. Participant X achieved his/her advantage through undeniably unjust procedures.

Item created ad hoc for the measure of competitive feelings.

36. I did enjoy competing with participant X in Task 2 (the task in which I chose my A and B numbers).

Twelve items created ad hoc to measure emotion arousal. The list (anxiety, anger, envy, irritation, jealousy, surprise, happiness, contempt, sadness, fear, joy, shame) is used by economists (Bosman and Van Winden, 2002; Bolle et al., 2010).

37. I was happy because of the assignment of the bonus.
38. I was angry because of the assignment of the bonus.
39. I enjoyed the assignment of the bonus.
40. I was sad because of the assignment of the bonus.
41. I was anxious because of the assignment of the bonus.
42. I was envious because of the assignment of the bonus.
43. I was irritated because of the assignment of the bonus.
44. I was jealous because of the assignment of the bonus.
45. I was surprised because of the assignment of the bonus.
46. I felt contempt because of the assignment of the bonus.
47. I felt fear because of the assignment of the bonus.
48. I felt shame because of the assignment of the bonus.

The following 14 items have been added to the questionnaire for Experiment 2.

Six items created ad hoc to measure participant's expectation of hostile feelings.

49. I think that after participant X found out the allocation of the bonus she/he felt anger against me.
50. I think that after participant X found out the allocation of the bonus she/he felt hostile feelings against me.
51. I think that after participant X found out the allocation of the bonus she/he hated me.
52. I think that after participant X found out the allocation of the bonus she/he disliked me.

53. I think that after participant X found out the allocation of the bonus she/he felt in low spirits.
54. I think that after participant X found out the allocation of the bonus she/he felt that she/he lacked energy.

Two items created ad hoc to be added to the objective injustice beliefs items.

55. An objective judge would agree that I did not deserve the bonus.
56. Anyone would agree that my advantage was unfairly obtained.

Two items created ad hoc to be added to the 'Warmth and Competence Model' items.

57. I think my success to Task 1 (the task in which I counted the number of ones) is due to luck.
58. I think my success to Task 1 (the task in which I counted the number of ones) is due to skill.

Two items created ad hoc to be added to the subjective injustice beliefs items.

59. The allocation rule of the bonus was fair.
60. I believe that participant X thinks that the allocation rule of the bonus was fair.

One item created ad hoc to measure participant's expectation of investment.

61. After I found out the allocation of the bonus I thought that participant X would have chosen a high Number A.

One item created ad hoc to measure participant's expectation of sabotage.

62. After I found out the allocation of the bonus I thought that participant X would have chosen a high Number B.

Two items created ad hoc to measure participant's response to the bonus allocation.

63. After I found out the allocation of the bonus I have chosen a high Number A.
64. After I found out the allocation of the bonus I have chosen a high Number B.

C.4 Task 6 - Results of the Questionnaire

For Experiment 1 the questionnaire consists of 48 items[10].

For the data analysis we have divided our observations into four groups (per each experiment):

1. High Type participants that received the bonus;
2. High Type participants that have not received the bonus;
3. Low Type participants that have received the bonus;
4. Low Type participants that have not received the bonus.

Per each group (and per each experiment) we carry out a principal component analysis, a technique for identifying groups or clusters of variables able to give as much information as possible on an underlying variable. In our case we aim at understanding the (group of) reasons and motivations that guided participants towards their decisions of investment and sabotage in the tournament game.

We grouped the items of the questionnaire. In Experiment 1 we have 19 variables, 7 of them are groups of items[11] while 12 of them are represented by only one item. Here our variables for Experiment 1:

1 Episodic Envy: (items 1-9);

2 Warmth towards the other and competence: items 10-14;

3 Competitive feelings toward the other: items 15-17 and item 36;

4 Subjective injustice beliefs: items 18-22;

5 Inferior beliefs: items 23-25;

6 Hostile and depressive feelings: items 26-32;

[10] We have added 14 items to the questionnaire for Experiment 2.

[11] We calculate the average of the participant's answers to the items that belongs to a group.

7 Objective injustice: items 33-35;

8 Happiness: item 37;

9 Anger: item 38;

10 Joy: item 39;

11 Sadness: item 40;

12 Anxiety: item 41;

13 Envy: item 42;

14 Irritation: item 43;

15 Jealousy: item 44;

16 Surprise: item 45;

17 Contempt: item 46;

18 Fear: item 47;

19 Shame: item 48;

For High Type participants that have received the bonus we retain only factors whose eigenvalue is equal or greater than one according to the Kaiser criterion. Hence we retain five factors. By the way the first three factors explain more the half of the total variance (the first factor explains the 39% of the variance, the second one the 13% and the third one the 10%). For the first factor anger and sadness have nearly the same weight followed by envy, jealousy and fear. Factor 2 is defined by beliefs of subjective and objective injustice (variable 4 and 7), and anxiety. Happiness and joy define Factor 3.
For High Type participants that have not received the bonus we retain 6 factors but only 3 of them combined explain nearly the half of the total variance (49%) with the same proportion. Sadness and surprise have the highest weight together

with episodic envy and anger in Factor 1. The variables that define Factor 2 are contempt, irritation and subjective injustice beliefs, while we have inferior beliefs, joy and happiness for Factor 3.

The factor that we retain in the principal component analysis carried out for Low Type participants that have received the bonus are five but already the first two can explain the 50% of the total variance. Factor 1 is defined mostly by irritation, episodic envy and hostile feelings while the variables with the highest weights for Factor 2 are fear and objective injustice beliefs. For Low Type participants that have not received the bonus the variables with the highest weights overall are joy, jealousy and subjective injustice beliefs which define Factor 1, Factor 2, and Factor 3 respectively. These three factors explain the 53% of the total variance.

We discuss now the results of the principal component analysis performed for the four groups of participants (indicated above) of Experiment 2. For this study we have 23 variables, 9 of them are groups of items while 14 of them are represented by only one item. Variables are identical to those used in Experiment 1 except for the following: 1) we added two items to variable 2 (items 57 and 58); 2) we added two items to variable 4 (items 59 and 60); and 3) we added two items to variable 7 (items 55 and 56). Finally, we have four more variables compared to Study I:

20 Expectation of other's hostile feelings: items 49-54;

21 Expectation of investment: item 61;

22 Expectation of sabotage: item 62;

23 Response to the bonus allocation: items 63-64;

We consider three factors in the principal component analysis for High Type participants that have received the bonus. They explain the 53% of the total variance. We observe a high weight of jealousy, irritation and hostile feelings for Factor 1. Competitive feelings towards the other and shame are significant variables for

Factor 2 and finally we notice that expectations of investment and sabotage define Factor 3. For High Type participants that have not received the bonus we retain two factors only that evenly explain the 52% of the total variance. The first group of variables (Factor 1) refers to emotions of sadness and irritation, and subjective and objective injustice beliefs, while inferior beliefs and shame define Factor 2.

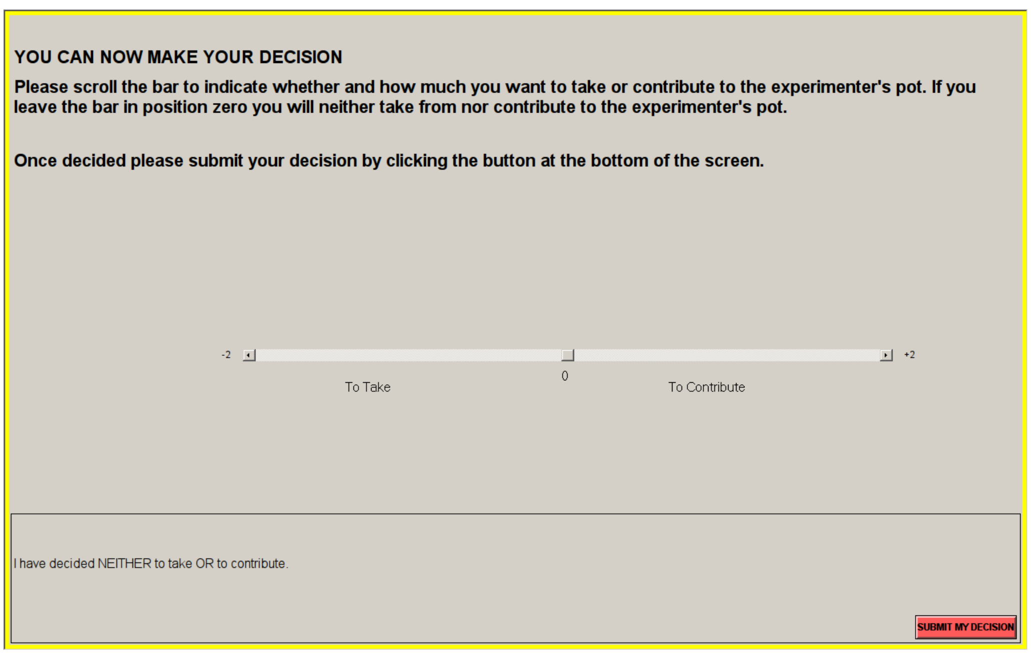
Seven factor have eigenvalues equal or higher than one in the principal component analysis for Low Type participants that have received the bonus and the first three of them explain the 55% of the total variance. Three variables have even weight in defining Factor 1: episodic envy, envy, and irritation. Variables in Factor 2 clearly refers to expectations of sabotage and investment while the group of items of objective injustice beliefs have the highest weight for Factor 3.

Two factors finally seems to explain most of the total variance (48%) in the analysis for Low Type participants that have not received the bonus. The first consists of episodic envy, jealousy, hostile feelings and subjective injustice beliefs, while for the second factor significant variables anxiety, fear and high expectation of investment.

C.5 Task 7: "Sabotaging the Experimenter" Task

In this task subjects read on their screens that the experimenter is responsible for the decision about the allocation rule of the bonus. Moreover, they are told that the experimenter had an extra pot of money for future research, and that they are given the opportunity either to take up to two pounds from the experimenter's pot or to contribute up to two pounds to the experimenter's pot[12]. Amounts of money are expressed in 2 decimal places. If the subject decides to contribute to the experimenter's pot, the amount of money indicated by the subject is deducted by the her/his final earnings. If the subject decides to take from the experimenter's pot, the amount of money indicated by the subject is added to her/his final earnings. Figure C.1 shows a screen shot of the task.

Figure C.1: Task 7. Sabotaging the experimenter



We introduce this task in our experiment because we understand the possibility that participants might have hostile feelings towards the experimenter who decided the

[12]Subjects can also choose to neither take from or to contribute to the experimenter's pot.

allocation rule to be applied in each experimental session. Because of the explorative nature of this task we do not make any predictions on the direction of the effects of this information on individuals' behaviour in this task.

We will refer to the variable "Contribution", which can take any value up to two decimal places from 0 to + 2, when the subject contributes to the experimenter's pot. Negative values up to - 2 indicate that participants take from the experimenter's pot; we call this decision "Sabotage". We pooled data of Experiment 1 and Experiment 2.

Overall, 18% of participants decided to contribute to the experimenter's pot of money. 76% took from the experimenter's pot. On average, participants contributed 1.26 to and took 1.78 from the experimenter' pot. Table C.1 shows the average contribution and sabotage by treatment and by type. In Experiment 1, High Type participants sabotage the experimenter more than Low Type participants when a bonus is allocated according to a meritocracy or anti-meritocracy rule. In Experiment 2, on average participants that did not receive the bonus (Low Type in the Meritocracy and High Type in the Anti-Meritocracy) took more from the experimenter's pot.

Table C.1: Contributing to and Sabotaging the Experimenter - Average values

	Contribution		Sabotage	
	High Type	Low Type	High Type	Low Type
Experiment 1				
Random Allocation 1	0	1.33	-1.92	-2
Meritocracy	0	0.91	-1.8	-1.42
Anti-Meritocracy	2	0.5	-1.91	-1.57
Random Allocation 2	1.67	1.49	-1.56	-1.71
Experiment 2				
Meritocracy	2	0.89	-1.82	-1.96
Anti-Meritocracy	1.41	1.01	-1.86	-1.54

35% of participants that decided to contribute to the experimenter's pot contributed

£2 and 74% of participants that decided to take from the experimenter's pot took £2. Since data are censored we apply a Tobit regression model in our regression analysis to investigate the effects of a group of explanatory variables on individual decisions to contribute to or to take from the experimenter's research pot. Table C.2 shows the estimates of Model 13-14 for contribution and Model 15-16 for sabotage decisions. We include in the models participants' type (Low Type = 1 if the subject belongs to the Low Type subgroup), a bonus dummy (Bonus = 1 if the subject has received the bonus), the score of risk aversion, the social desirability score, treatment dummies, an interaction term between participants' type and bonus, and some demographic information.

Table C.2: Tobit Estimates on Subjects' Contribution and Sabotage

	Contribution Model 13	Contribution Model 14	Sabotage Model 15	Sabotage Model 16
Low Type	-0.347	-0.150	-0.356	-0.406
Meritocracy	0.086	0.204	0.944**	0.947**
Anti-Meritocracy	0.110	0.411	0.299	0.372
Random Allocation 2	0.2796	0.611*	0.726*	0.712*
Bonus	1.042	1.077**	-0.261	-0.262
Risk Aversion	-0.107	-0.041	-0.030	-0.028
Social Desirability Score	0.111*	0.058	0.062	0.048
Low Type x Bonus	-1.222	-0.822	0.856	0.867
Male		0.483*		-0.152
Age		0.198***		0.004
British Nationality		0.682*		-0.282
School of ECO		0.356		-0.224
Constant	0.907	-4.192***	-3.346***	-3.072**
sigma				
Constant	0.696***	0.457***	1.165***	1.150***
Observations	30	30	126	126
Left-censored observations	0	0	93	93
Right-censored observations	11	11	0	0
Likelihood ratio test (LR)	15.44*	34.66***	10.19	11.53

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

There is evidence that in the Meritocracy treatment and in the Random Allocation

2 treatments participants were more likely to take from the experimenter's pot than in the Random Allocation 1 treatment. Individuals' social desirability level and demographic variables are significant only for contribution decisions (Model 13-16).

C.6 The Guessing Stage

YOUR GUESSES

Before choosing your Numbers A and B you will be asked to make **TWO** guesses, one on participant X's choice of Number A and the other on participant X's choice of Number B.

At the end of the experiment the computer will select ONE of your guesses at random. The Experimental Points you have earned for that guess will be converted into pounds at the rate of **2 pence for each point**.

Your guess on participant X's Number A

You will see the following sentence on the screen: "Please guess participant X's choice of Number A".

You will be asked to make a guess of the Number A, from 0 to 100, that participant X has chosen.

You can make the higher earnings by choosing your best guess. If your guess turns to be exactly right, you will earn 200 Experimental Points. If not, for every percentage point your guess is 'out', a deduction of 20 points is made from the 200 Experimental Points. But you cannot lose more than 200 Experimental Points.
Example 1

Suppose your guess of participant X's choice of Number A is 50. Suppose participant X's choice of Number A is 50. In this case you will earn 200 Experimental Points.

Table C.3: Your guess and your earnings in Experimental Points

If your guess is	**You will earn**
exactly right	200
1% out	200 - 20 = 180
2% out	200 - (20x2) = 160
3% out	200 - (20x3) = 140
4% out	200 - (20x4) = 120
5% out	200 - (20x5) = 100
6% out	200 - (20x6) = 80
7% out	200 - (20x7) = 60
8% out	200 - (20x8) = 40
9% out	200 - (20x9) = 20
10% out	200 - (20x10) = 0
More than 10% out	0

Example 2

Suppose your guess of participant X's choice of Number A is 50. Suppose participant X's choice of Number A is 53. In this case you will earn (200 - 20 x (53-50)) = (200 - (20 x 3)) = 140 Experimental Points.

Example 3

Suppose your guess of participant X's choice of Number A is 50. Suppose participant X's choice of Number A is 15. In this case you will earn 0 Experimental Points.

Your guess on participant X's Number B

You will see the following sentence on the screen: "Please guess participant X's choice of Number B".

You will be asked to make a guess of the number B, from 0 to 50, that participant X has chosen.

You can make the higher earnings by choosing your best guess. If your guess turns to be exactly right, you will earn 200 Experimental Points. If not, for every point your guess is 'out', a deduction of 40 Experimental Points is made from the 200 Experimental Points. But you cannot lose more than 200 Experimental Points.

Table C.4: Your guess and your earnings in Experimental Points

If your guess is	**You will earn**
exactly right	200
1 point out	200 - 40 = 160
2 point out	200 - (40x2) = 120
3 point out	200 - (40x3) = 80
4 point out	200 - (40x4) = 40
5 point out	200 - (40x5) = 0
More than 5 point out	0

Example 1

Suppose your guess of participant X's choice of Number B is 25. Suppose participant Xs' choice of Number B is 25. In this case you will earn 200 Experimental Points.

Example 2

Suppose your guess of participant X's choice of Number B is 25. Suppose participant X's choice of Number B is 23. In this case you will earn (200 - 40 x (25-23)) = (200 - (40 x 2)) = 120 Experimental Points.

Example 3

Suppose your guess of participant X's choice of Number B is 25. Suppose participant X's choice of Number B is 40. In this case you will earn 0 Experimental Points.

C.7 Tables and Figures

Table C.5: Descriptive Statistics of Investment by Type - Experiment 1

		N	Mean	Std.Dev.	Min	Max
Overall	High Type	48	32.791	27.517	0	88
	Low Type	48	33.625	28.196	0	90
Baseline	High Type	13	28	33.687	0	88
	Low Type	13	22.153	23.744	0	65
Treatment A	High Type	11	24.818	25.922	0	70
	Low Type	11	49.363	30.220	0	90
Treatment B	High Type	12	37.916	26.314	0	76
	Low Type	12	33.333	28.750	0	80
Treatment C	High Type	12	40.166	22.702	0	77
	Low Type	12	28.916	26.262	0	62

Table C.6: Descriptive Statistics of Sabotage by Type - Experiment 1

		N	Mean	Std.Dev.	Min	Max
Overall	High Type	36	10.972	13.028	0	50
	Low Type	36	14.416	14.422	0	50
Baseline	High Type	13	8.461	12.971	0	36
	Low Type	13	13.846	15.186	0	40
Treatment A	High Type	11	12.909	15.584	0	50
	Low Type	11	10.636	12.339	0	36
Treatment B	High Type	12	11.916	11.114	0	30
	Low Type	12	18.5	15.459	0	50

Table C.7: Bivariate Tobit Estimates on Investment and Sabotage with Demographic Variables

	Model 13	Model 14	Model 15
Investment			
Low Type	-12.482	-11.276	-17.605
Meritocracy	36.133***	33.527***	34.496***
Anti-Meritocracy	24.746*	24.343**	25.297**
Bonus	-38.902**	-33.921**	-34.821**
Risk Aversion	-6.574***	-5.977***	-6.364***
Low Type x Bonus	37.255	31.703	32.686
Male	3.562	4.841	4.398
Age	0.270	0.470	0.486
British Nationality	-10.765	-8.262	-6.625
School of ECO	-18.697**	-16.073*	-17.240**
Exp. Inv.		0.357***	0.208
Exp. Sab.		-0.034	0.194
Low Type x Exp. Sab.			-0.253
Low Type x Exp. Inv.			0.238
Constant	65.014**	37.774	41.672
Sabotage			
Low Type	-10.423	-10.920	-30.277**
Meritocracy	15.349*	16.804*	19.142**
Anti-Meritocracy	10.195	10.792	11.804*
Bonus	-11.871	-13.813	-15.768
Risk Aversion	-3.689***	-3.647***	-4.093***
Low Type x Bonus	28.724	28.224	30.050*
Male	5.601	4.068	3.496
Age	-0.090	-0.448	-0.478
British Nationality	-0.543	-0.977	0.138
School of ECO	-6.638	-8.102	-8.074
Exp. Inv.		-0.130	-0.284**
Exp. Sab.		0.354**	0.368
Low Type x Exp. Sab.			0.101
Low Type x Exp. Inv.			0.319*
Constant	26.466*	34.368**	44.284***
sigma1			
Constant	30.118***	27.392***	27.331***
sigma2			
Constant	17.569***	16.892***	15.860***
rho			
Constant	0.936***	1.225***	1.192***
Observations	71	71	71
Wald test	27.51**	40.73***	41.66***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.8: Bivariate Tobit Estimates on Investment and Sabotage with Social Desirability Measure

	Model 17	Model 18	Model 19
Investment			
Low Type	6.472	-14.276	-14.121
Meritocracy	24.609**	37.684***	39.405*
Anti-Meritocracy	29.154***	20.628*	16.889
Bonus	-17.056**	-39.006**	-41.897*
Risk Aversion	-7.061***	-6.354***	-6.295***
Social Desirability Score	-0.288	-0.099	-0.068
Low Type x Bonus		42.714	41.001
Low Type x Anti-Meritocracy			7.233
Low Type x Meritocracy			-5.424
Constant	57.955***	60.294***	61.996***
Sabotage			
Low Type	3.296	-12.184	2.114
Meritocracy	5.554	15.945*	92.272
Anti-Meritocracy	15.713***	10.172	13.335
Bonus	4.265	-11.854	-82.076
Risk Aversion	-4.628***	-4.256***	-4.411***
Social Desirability Score	-1.893**	-1.821**	-2.123**
Low Type x Bonus		31.542*	91.334
Low Type x Anti-Meritocracy			-5.803
Low Type x Meritocracy			-93.624
Constant	40.260***	42.865***	44.594***
sigma1			
Constant	31.612***	31.232***	31.013***
sigma2			
Constant	17.993***	17.701***	16.996***
rho			
Constant	1.106***	1.044***	1.057***
Observations	71	71	71
Wald test	19.96**	21.50**	21.53**

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure C.2: Expectation of Investment Distribution

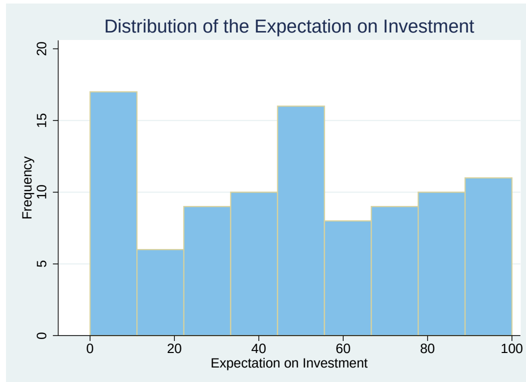


Figure C.3: Expectation of Investment - Low Type Participants

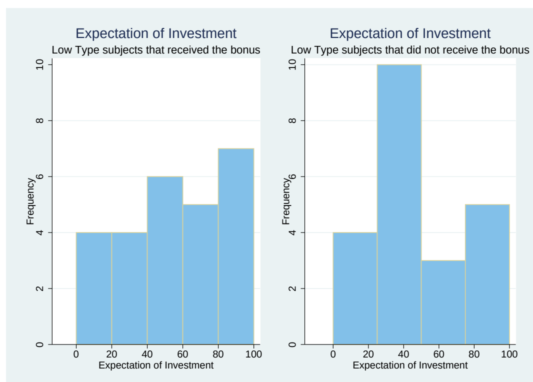


Figure C.4: Expectation of Investment - High Type Participants

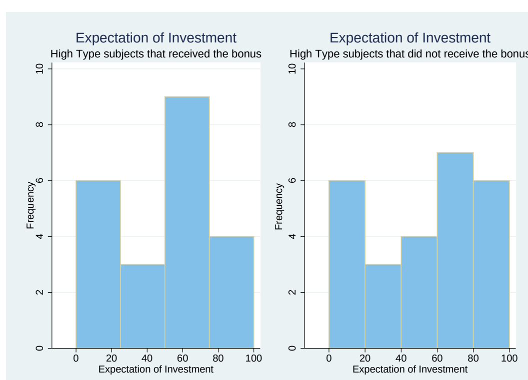


Figure C.5: Expectation of Sabotage Distribution

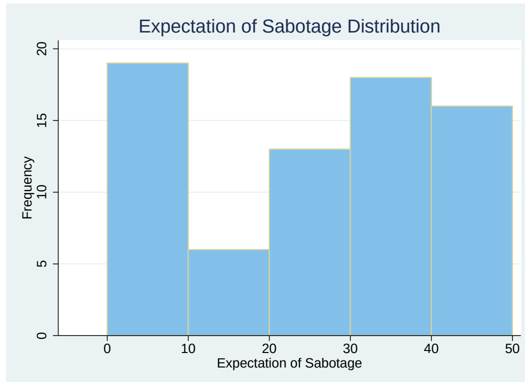


Figure C.6: Expectation of Sabotage - Low Type Participants

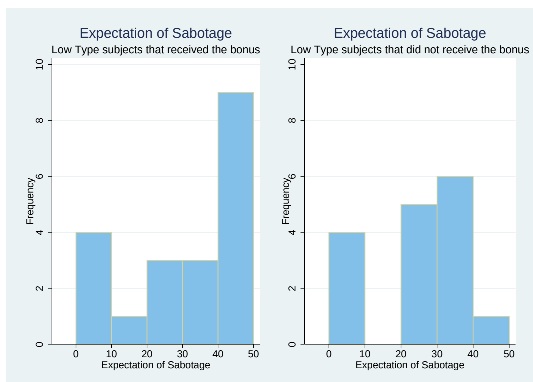


Figure C.7: Expectation of Sabotage - High Type Participants

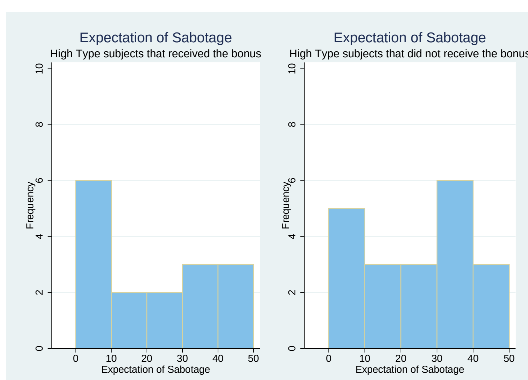


Table C.9: Correlation between variables - Experiment 1

	Inv.	Sab.	Exp. Inv.	Exp. Sab.	Bonus	Risk-Av.	Social Des.	High Type	Low Type	Contribution
Inv.	1.000									
Sab.	0.542	1.000								
	0.000									
Exp. Inv.	0.373	0.028	1.000							
	0.001	0.812								
Exp. Sab.	0.189	0.285	0.519	1.000						
	0.113	0.015	0.000							
Bonus	-0.209	0.044	-0.035	0.100	1.000					
	0.080	0.713	0.770	0.4034						
Risk-Av.	-0.277	-0.292	-0.101	-0.080	-0.079	1.000				
	0.019	0.013	0.399	0.506	0.510					
Social Des.	0.010	-0.034	0.213	0.228	0.206	-0.222	1.000			
	0.933	0.776	0.074	0.055	0.084	0.062				
High Type	-0.092	-0.113	-0.069	-0.176	-0.127	0.137	-0.100	1.000		
	0.443	0.347	0.567	0.141	0.291	0.251	0.405			
Low Type	0.092	0.113	0.069	0.176	0.127	-0.137	0.100	-1.000	1.000	
	0.443	0.347	0.567	0.141	0.291	0.251	0.405	1.000		
Contribution	0.234	0.127	0.129	0.100	-0.043	-0.186	0.153	-0.289	0.289	
	0.049	0.287	0.282	0.405	0.720	0.118	0.202	0.014	0.014	

Figure C.8: Decisions of Investment and Sabotage - by Type

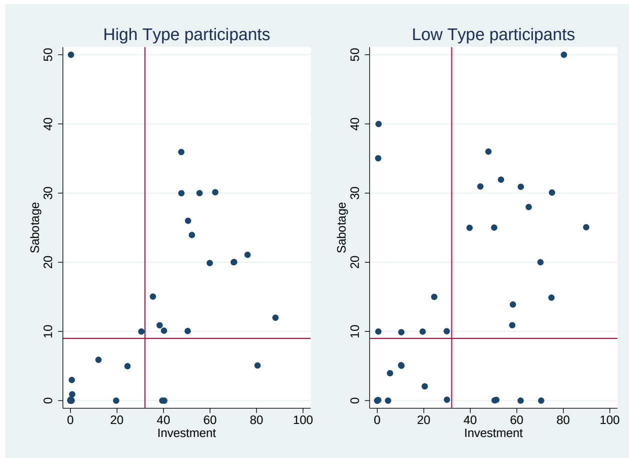


Figure C.9: Decisions of Investment and Sabotage - Meritocracy

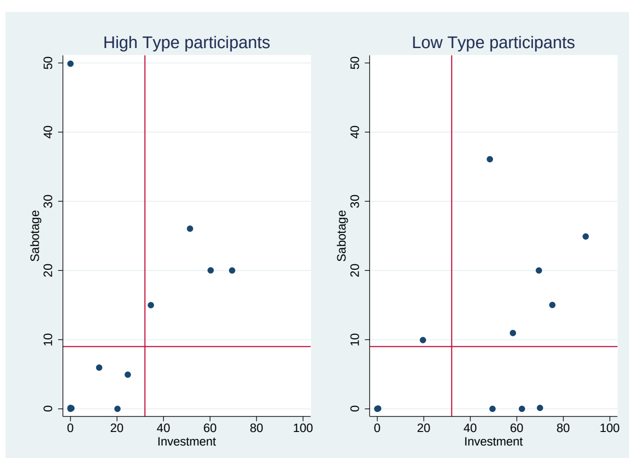


Figure C.10: Decisions of Investment and Sabotage - Anti-Meritocracy

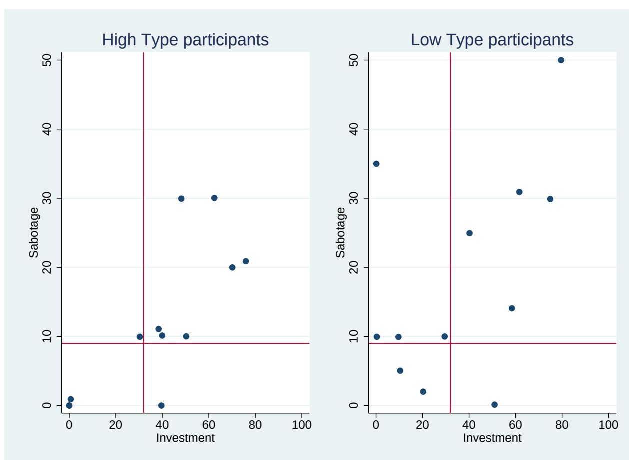


Table C.10: Descriptive Statistics of Investment by Treatments - Experiment 2

	N	Mean	Std.Dev.	Min	Max
Treatment A2	36	39.5	25.905	0	100
Treatment B2	36	41.083	30.152	0	100

Table C.11: Descriptive Statistics of Investment by Type - Experiment 2

		N	Mean	Std.Dev.	Min	Max
Overall	High Type	36	30.666	23.099	0	90
	Low Type	36	49.916	29.281	0	100
Treatment A2	High Type	18	31.388	19.189	0	65
	Low Type	18	47.611	29.565	10	100
Treatment B2	High Type	18	29.944	27.004	0	90
	Low Type	18	52.222	29.661	0	100

Table C.12: Descriptive Statistics of Sabotage by Treatments - Experiment 2

	N	Mean	Std.Dev.	Min	Max
Treatment A2	36	18.139	12.613	0	50
Treatment B2	36	15.944	12.954	0	40

Table C.13: Descriptive Statistics of Sabotage by Types - Experiment 2

		N	Mean	Std.Dev.	Min	Max
Overall	High Type	36	14.111	10.208	0	35
	Low Type	36	19.972	14.405	0	50
Treatment A2	High Type	18	15.5	7.920	0	30
	Low Type	18	20.777	15.813	0	50
Treatment B2	High Type	18	12.722	12.154	0	35
	Low Type	18	19.166	13.258	0	40

Figure C.11: Expectation of Investment Distribution

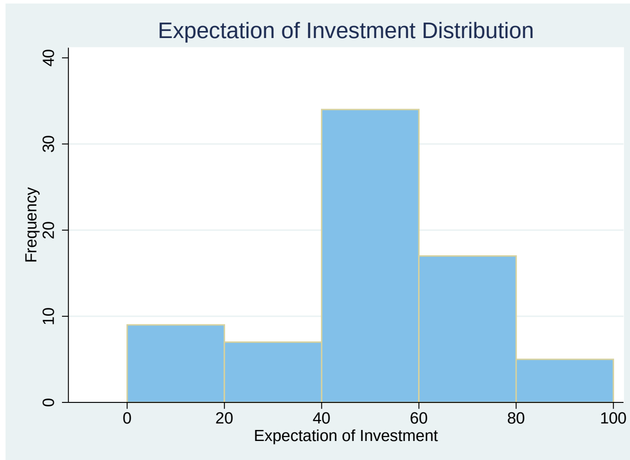


Figure C.12: Expectation of Investment - Low Type Participants

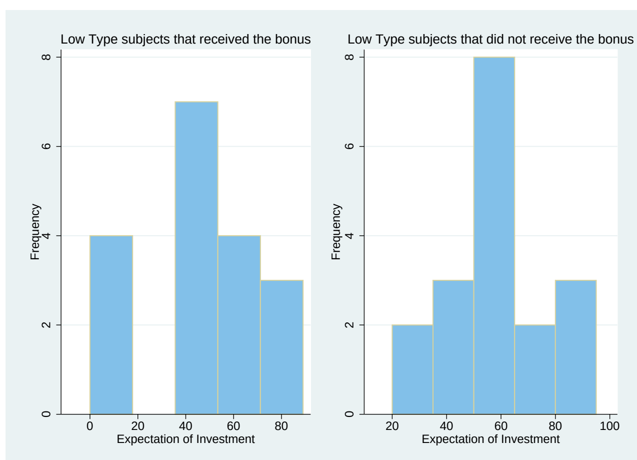


Table C.14: Bivariate Tobit Estimates on Investment and Sabotage with Demographic Variables - Experiment 2

	Model 19	Model 20	Model 21
Investment			
Low Type	20.811***	16.722***	-23.036
Anti-Meritocracy	-3.449	-0.760	-1.403
Bonus	5.737	8.195	10.166*
Risk Aversion	-2.742	-1.287	-0.474
Male	8.011	5.402	1.266
Age	-2.583	-2.993**	-2.628*
British Nationality	-8.299	-10.123	-5.183
School of ECO	1.563	4.548	6.949
Exp. Inv.		0.742***	0.572**
Exp. Sab.		0.013	-0.532
Low Type x Exp. Sab.			0.971*
Low Type x Exp. Inv.			0.371
Constant	97.868**	62.254*	67.718**
Sabotage			
Low Type	5.697	3.398	-12.060
Anti-Meritocracy	-4.208	-1.043	-3.138
Bonus	1.982	-0.880	1.352
Risk Aversion	-0.212	-0.826	-0.367
Male	-1.262	1.907	1.372
Age	-0.413	-0.969	-1.063*
British Nationality	-6.073	-8.644**	-7.702**
School of ECO	1.479	3.648	3.905
Exp. Inv.		-0.022	0.088
Exp. Sab.		0.741***	0.163
Low Type x Exp. Sab.			0.994***
Low Type x Exp. Inv.			-0.129
Constant	25.670	26.185	32.552**
sigma1			
Constant	28.407***	23.971***	22.482***
sigma2			
Constant	14.421***	11.439***	9.988***
rho			
Constant	1.180***	1.701***	1.623***
Observations	71	71	71
Wald test	15.91**	60.17***	62.96***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Exp. Inv.' = Expectation of Investment;

'Exp. Sab.' = Expectation of Sabotage.

Table C.15: Correlation between variables - Experiment 2

	Inv.	Sab.	Exp. Inv.	Exp. Sab.	Bonus	Risk-Av.	Social Des.	High Type	Low Type	Contribution
Inv.	1.000									
Sab.	0.479	1.000								
	0.000									
Exp. Inv.	0.524	0.094	1.000							
	0.000	0.435								
Exp. Sab.	0.130	0.554	0.202	1.000						
	0.281	0.000	0.092							
Bonus	0.049	0.044	-0.123	0.168	1.000					
	0.683	0.713	0.310	0.164						
Risk-Av.	-0.210	-0.051	-0.134	0.124	0.061	1.000				
	0.080	0.671	0.267	0.306	0.613					
Social Des.	0.108	0.231	-0.127	0.148	-0.064	-0.070	1.000			
	0.370	0.053	0.292	0.220	0.597	0.560				
High Type	-0.323	-0.166	-0.145	-0.096	0.000	0.213	-0.153	1.000		
	0.006	0.167	0.228	0.4265	1.000	0.075	0.204			
Low Type	0.323	0.166	0.145	0.096	0.000	-0.213	0.153	-1.000	1.000	
	0.006	0.167	0.228	0.426	1.000	0.075	0.204	1.000		
Contribution	0.089	-0.025	-0.209	-0.071	0.184	-0.203	0.214	-0.130	0.1302	
	0.462	0.832	0.082	0.558	0.126	0.091	0.074	0.282	0.282	

Figure C.13: Expectation of Investment - High Type Participants

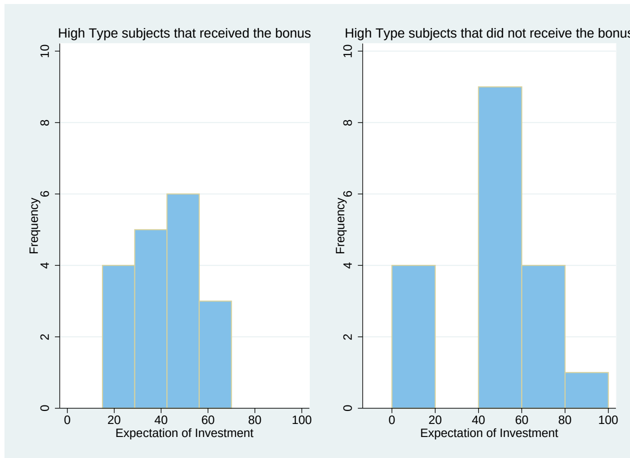


Figure C.14: Expectation of Sabotage Distribution

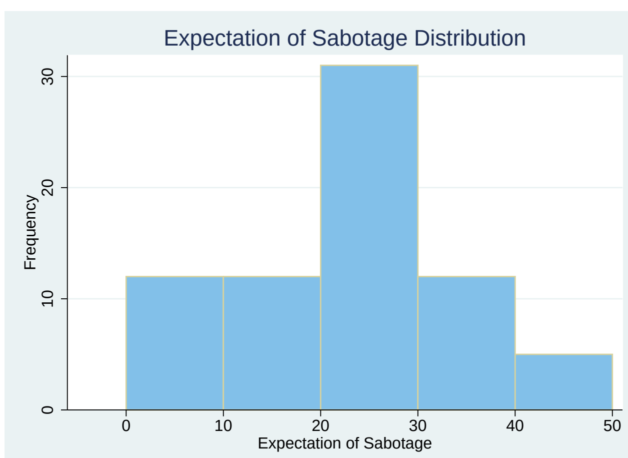


Figure C.15: Expectation of Sabotage - Low Type Participants

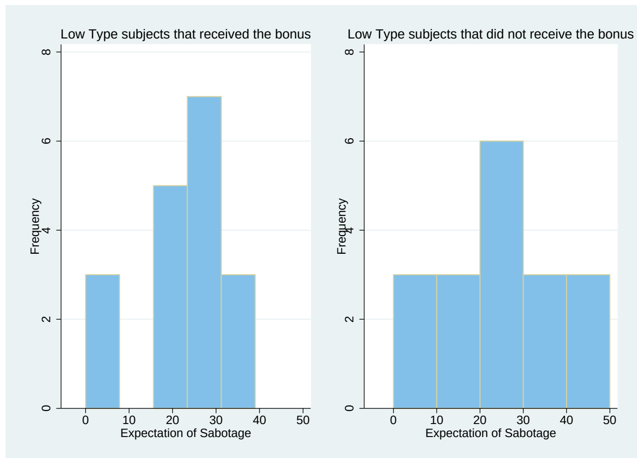


Figure C.16: Expectation of Sabotage - High Type Participants

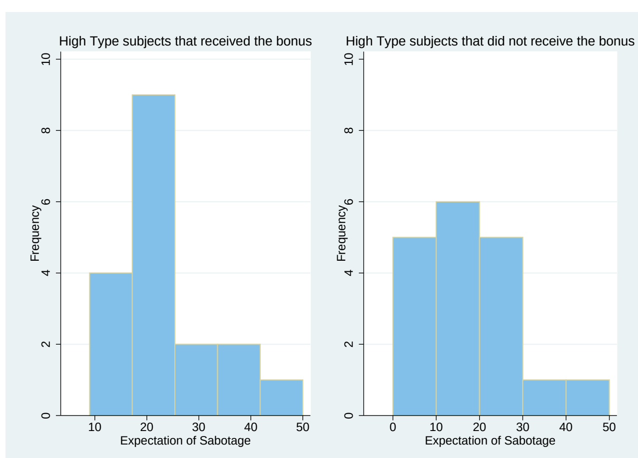


Figure C.17: Decisions of Investment and Sabotage - by Type

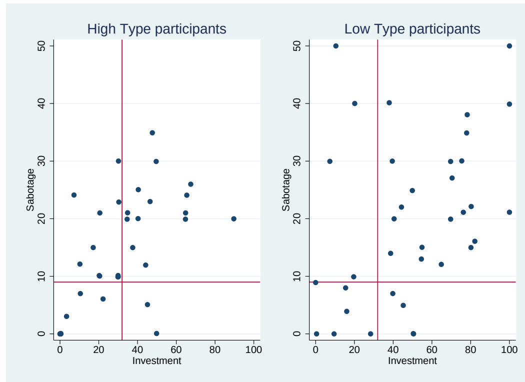


Figure C.18: Decisions of Investment and Sabotage - Meritocracy

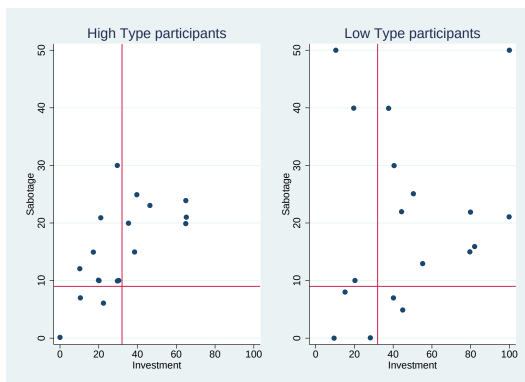
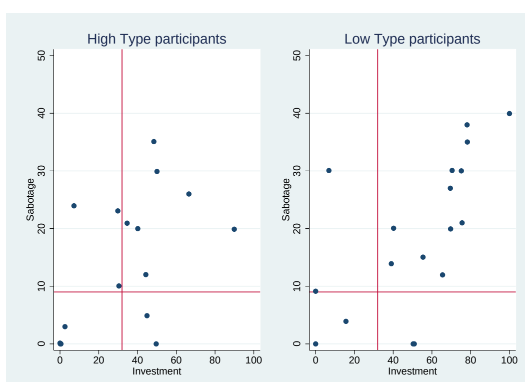


Figure C.19: Decisions of Investment and Sabotage - Anti-Meritocracy



Bibliography

Abeler, J., A. Falk, L. Götte, and D. Huffman (2011). Reference points and effort provision. *The American Economic Review 101*(2), 470–492.

Agster, K. L., N. J. Fortin, and H. Eichenbaum (2002). The hippocampus and disambiguation of overlapping sequences. *The Journal of neuroscience 22*(13), 5760–5768.

Akerlof, G. and J. Yellen (1986). *Efficiency Wage Models of the Labor Market.* Cambridge University Press.

Ambler, T. and T. Burne (1999). The impact of affect on memory of advertising. *Journal of Advertising Research 39*(2), 25–34.

Amegashie, J. A. and M. Runkel (2006). Sabotaging Potential Rivals. *Social Choice and Welfare 28*(1), 143–162.

Apesteguia, J. and I. Palacios-Huerta (2010). Psychological pressure in competitive environments: evidence from a randomized natural experiment. *American Economic Review 100*, 2548–2564.

Arkes, H. R., L. T. Herren, and A. M. Isen (1988). The role of potential loss in the influence of affect on risk-taking behavior. *Organizational Behavior and Human Decision Processes 42*(2), 181–193.

Ashcraft, M. H. and E. P. Kirk (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology 130*(2), 224.

Auriol, E., G. Friebel, and L. Pechlivanos (2002). Career concerns in teams. *Journal of Labor Economics 20*(2), 289–307.

Azar, O. H. (2009). Does relative thinking exist in mixed compensation schemes. Technical report, Working paper, Ben-Gurion University of the Negev.

Baker, G. (1992). Incentive contracts and performance measurement. *Journal of Political Economy 100*, 598–614.

Baker, G., M. C. Jensen, and K. J. Murphy (1987). Compensation and incentives: Practice vs theory. *The Journal of Finance 43*(3), 593–616.

Balafoutas, L. and M. Sutter (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science 335*(6068), 579–582.

Bannerman, D., J. Rawlins, S. McHugh, R. Deacon, B. Yee, T. Bast, W.-N. Zhang, H. Pothuizen, and J. Feldon (2004). Regional dissociations within the hippocampus?memory and anxiety. *Neuroscience & Biobehavioral Reviews 28*(3), 273–283.

Ben-Shakhar, G., G. Bornstein, A. Hopfensitz, and F. van Winden (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of Economic Psychology 28*, 314–323.

Bolle, F., J. H. Tan, and D. Zizzo (2010). Vendettas. *CeDEx Discussion Paper Number 2010-02*.

Bolton, G. E., J. Brandts, and A. Ockenfels (2005). Fair procedures: Evidence from games involving lotteries. *The Economic Journal 115*(506), 1054–1076.

Bortolotti, S. (2004). *Incentives, Group Pride, and Real Effort in the Weak-link Game: An Experimental Analysis*. Ph. D. thesis, University of Trento.

Bose, A., D. Pal, and D. E. Sappington (2010). Equal pay for unequal work: Limiting sabotage in teams. *Journal of Economics & Management Strategy 19*(1), 25–53.

Bosman, R., M. Sutter, and F. Van Winden (2005). The impact of real effort and emotions in the power-to-take game. *Journal of Economic Psychology 26*(3), 407–429.

Bosman, R. and F. Van Winden (2002). Emotional hazard in a power-to-take experiment. *The Economic Journal 112*(476), 147–169.

Bosman, R. and F. van Winden (2010). Global risk, investment and emotions. *Economica 77*(307), 451–471.

Brandts, J. and G. Charness (2004). Do labour market conditions affect gift exchange? some experimental evidence. *The Economic Journal 114*(497), 684–708.

Breitmoser, Y., J. Tan, and D. Zizzo (2010). Understanding Perpetual R&D Races. *Economic Theory 44*(3), 445–467.

Brown, C. (1992). Wage levels and methods of pay. *Rand Journal of Economics 23*, 366–375.

Brown, M. and P. Philips (1986). The decline of the piece-rate system in california canning: Technological innovation, labor management, and union pressure, 1890-1947. *The Business History Review 60*(4), 564–601.

Brüggen, A. and M. Strobel (2007). Real effort versus chosen effort in experiments. *Economics Letters 96*(2), 232–236.

Bull, C., A. Schotter, and K. Weigelt (1987). Tournaments and piece rates: an experimental study. *Journal of Political Economy 95*(1), 1–33.

Burrows, P. and G. Loomes (1994). The impact of fairness on bargaining. *Empirical Economics 19*(2), 201–221.

Camerer, C., G. Loewenstein, and D. Prelec (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of economic Literature*, 9–64.

Camerer, C. F. and R. M. Hogarth (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty 19*(1), 7–42.

Camille, N., G. Coricelli, J. Sallet, P. Pradat-Diehl, J.-R. Duhamel, and A. Sirigu (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science 304*(5674), 1167–1170.

Caplin, A. and J. Leahy (2001). Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics 116*(1), 55–79.

Carpenter, J., P. Matthews, and J. Schirm (forthcoming). Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*.

Cassady, J. and R. Johnson (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology 27*(2), 270–295.

Chan, W. (1996). External recruitment versus internal promotion. *Journal of Labor Economics*, 555–570.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics 22*(3), 665–688.

Charness, G., G. R. Frechette, and J. H. Kagel (2004). How robust is laboratory gift exchange? *Experimental Economics 7*(2), 189–205.

Charness, G. and M. Villeval (2009). Cooperation and competition in intergenerational experiments in the field and the laboratory. *The American Economic Review 99*(3), 956–978.

Chen, K.-P. (2003). Sabotage in promotion tournaments. *Journal of Law, Economics, and Organization 19*(1), 119–140.

Chen, K.-P. (2005). External recruitment as an incentive device. *Journal of Labor Economics 23*(2), 259–277.

Chowdhury, S. and O. Gurtler (2013). Sabotage in contests: A survey. *Working paper*.

Cohen-Charash, Y. (2009). Episodic envy. *Journal of Applied Social Psychology 39*(9), 2128–2173.

Cornes, R. and R. Hartley (2003). Risk aversion, heterogeneity and contests. *Public Choice 117*(1-2), 1–25.

Cornes, R. and R. Hartley (2012). Risk aversion in symmetric and asymmetric contests. *Economic Theory 51*(2), 247–275.

Corr, P. (2008). *The Reinforcement Sensitivity Theory of Personality.* Cambridge Press.

Corr, P., A. Pickering, and J. Gray (1997). Personality, punishment, and procedural learning: A test of Gray's Anxiety Theory. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology 73*(2), 337.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature*, 448–474.

Dabbs, J. M. and R. L. Helmreich (1972). Fear, anxiety, and affilation following a role-played accident. *The Journal of Social Psychology 86*(2), 269–278.

Damasio, A. (2008). *Descartes' error: Emotion, reason and the human brain.* Random House.

Deci, E. (1972). Intrinsic motivation, extrinsic reinforcement and equity. *Journal of Personality and Social Psychology 22*(1), 113–120.

Dickinson, D. (2001). The carrot vs. the stick in work team motivation. *Experimental Economics 4*(1), 107–124.

Dickinson, D. and M. Villeval (2004). Does monitoring decrease work effort? the complementarity between agency and crowding-out theories. *Working Paper*.

Digman, J. (1990). Personality structure: Emergence of the five-factor model. *Annual review of psychology 41*(1), 417–440.

Dohmen, T. and A. Falk (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review 101*(2), 556–590.

Drago, R. and J. S. Hyewood (1995). The choice of payment schemes: Australian establishment data. *Industrial Relations: A Journal of Economy and Society 34*(4), 507–531.

Drago, R. and G. K. Turnbull (1991). Competition and cooperation in the workplace. *Journal of Economic Behavior & Organization 15*(3), 347–364.

Dreger, R. (1985). Real and random p-technique analyses of the state-trait anxiety inventory and their relation to r-technique analyses. *Southern Psychologist*.

Eichenbaum, H. and N. Fortin (2003). Episodic memory and the hippocampus it's about time. *Current Directions in Psychological Science 12*(2), 53–57.

Ekman, P. (1992). *Are There Basic Emotions?* American Psychological Association.

Ellis, H. C., R. L. Thomas, A. D. McFarland, and J. W. Lane (1985). Emotional mood states and retrieval in episodic memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition 11*(2), 363.

Ellis, H. C., R. L. Thomas, and I. A. Rodriguez (1984). Emotional mood states and memory: Elaborative encoding semantic processing, and cognitive effort. *Journal of Experimental Psychology. Learning, Memory, and Cognition 10*(3), 470–482.

Endler, N. (1997). Stress, anxiety and coping: The multidimensional interaction model. *Canadian Psychology 38*(3), 136.

Eriksson, T., A. Poulsen, and M. Villeval (2008). Feedback and incentives: Experimental evidence. *Working Paper*.

Eysenck, M. (1985). Anxiety and cognitive-task performance. *Personality and Individual differences 6*(5), 579–586.

Eysenck, M. (1997). *Anxiety and Cognition: A Unified Theory.* Psychology Press.

Eysenck, M. and M. Calvo (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion 6*, 409–434.

Fahr, R. and B. Irlenbusch (2010). Fairness as a constraint on trust and reciprocity. *Economics Letters 66*, 275–282.

Falk, A., E. Fehr, and D. Huffman (2008). The power and limits of tournament incentives. *Work in progress*.

Falk, A., D. Huffman, and U. Sunde (2006). Self-confidence and search. *Working Paper*.

Falk, A. and A. Ichino (2006). Clean evidence on peer effects. *Journal of Labor Economics 24*(1), 39–57.

Fama, E. F. (1991). Time, salary, and incentive payoffs in labor contracts. *Journal of Labor Economics 9*(1), 25–44.

Fehr, E. and S. Gächter (2000). Fairness and retaliation: The economics of reciprocity. *The Journal of Economic Perspectives 14*(3), 159–181.

Fehr, E., S. Gächter, and G. Kirchsteiger (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica*, 833–860.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics 108*(2), 437–459.

Fehr, E. and K. M. Schmidt (2004). Fairness and incentives in a multi-task principal–agent model. *The Scandinavian Journal of Economics 106*(3), 453–474.

Felix-Ortiz, A. C., A. Beyeler, C. Seo, C. A. Leppla, C. P. Wildes, and K. M. Tye (2013). Bla to vhpc inputs modulate anxiety-related behaviors. *Neuron 79*(4), 658–664.

First, M., A. Frances, and P. Harold (1987). *Diagnostic and Statistical Manual of Mental Disorders.* American Psychiatric Association.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economics experiments. *Experimental Economics 10*(2), 171–178.

Fiske, S. T., A. J. C. Cuddy, P. Glick, and J. Xu (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology 82*(6), 878–902.

Folger, R. and J. Greenberg (1985). *Procedural Justice: an Interpretative Analysis of Personnel System.* Addison-Wesley.

Fortin, N. J., K. L. Agster, and H. B. Eichenbaum (2002). Critical role of the hippocampus in memory for sequences of events. *Nature neuroscience 5*(5), 458–462.

Frey, B. (1993). Does Monitoring Increase Work Effort? The Rivalry With Trust and Loyalty. *Economic Inquiry 31*, 663–670.

Frey, B. S. and F. Oberholzer-Gee (1997). The cost of price incentives: An empirical analysis of motivation crowding- out. *The American Economic Review 87*(4), pp. 746–755.

Frijda, N. (1986). *The Emotions.* Cambridge: Cambridge University Press.

Frijda, N. H. (1993). Moods, emotion episodes, and emotions. In M. Lewis and J. M. Haviland (Eds.), *Handbook of Emotions*, pp. 381–404. Guilford Press, New York.

Gächter, S. and A. Falk (2002). Reputation and reciprocity: Consequences for the labour relation. *The Scandinavian Journal of Economics 104*(1), 1–26.

Gächter, S. and C. Thöni (2010). Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis. *Journal of Economic Behavior & Organization 76*(3), 531–543.

Gibbons, R. and M. Waldman (1999). Careers in organizations: Theory and evidence. *Handbook of labor economics 3*, 2373–2437.

Gill, D. and V. Prowse (2011). A novel computerized real effort task based on sliders. *Working Paper*.

Glaser, B. G. and A. L. Strauss (1967). *The discovery of grounded theory: Strategies for qualitative research.* Aldine de Gruyter.

Gneezy, U. (2002). Does high wage lead to high profits: An experimental study of reciprocity using real efforts. *Graduate School of Business, University of Chicago*.

Gneezy, U. and J. List (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica 74*(5), 1365–1384.

Gneezy, U., M. Niederle, and A. Rustichini (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics 118*(3), 1049–1074.

Gneezy, U. and A. Rustichini (2004). Gender and competition at a young age. *The American Economic Review 94*(2), 377–381.

Goldin, C. (1986). Monotoring costs and occupational segregation by sex: A historical analysis. *Journal of Labour Economics*, 1–27.

Gray, J. (1985a). *Issues in the Neuropsychology of Anxiety.* Lawrence Erlbaum Associates, Inc.

Gray, J. (1985b). The whole and its parts: Behaviour, the brain, cognition and emotion. *Bullettin of the British Psychological Society 38*, 99–112.

Gray, J. and N. McNaughton (2000). The neuropsychology of anxiety.

Gray, J. A. (1982). On mapping anxiety. *Behavioral and Brain Sciences 5*(03), 506–534.

Green, J. R. and N. Stokey (1982). A comparison of tournaments and contracts. *Working Paper*.

Greenberg, J. (1988). Equity and workplace status: A field experiment. *Journal of applied Psychology 73*(4), 606–613.

Greenberg, J. (1989). Research notes. cognitive reevaluation of outcomes in response to underpayment inequity. *Academy of Management Journal 32*(1), 174–184.

Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of management 16*(2), 399–432.

Gürtler, O. and J. Münster (2010). Sabotage in dynamic tournaments. *Journal of Mathematical Economics 46*(2), 179–190.

Hackman, R. J., G. O. (1980). *Work Redesign.* Reading: MA: Addison-Wesley.

Hackman, J. and E. Lawler (1971). Employee reactions to job characteristics. *Journal of applied psychology 55*(3), 259.

Harbring, C. and B. Irlenbusch (2003). An experimental study on tournament design. *Labour Economics 10*(4), 443–464.

Harbring, C. and B. Irlenbusch (2005). Incentives in Tournaments with Endogenous Prize Selection. *Journal of Institutional and Theoretical Economics JITE 161*(4), 636–663.

Harbring, C. and B. Irlenbusch (2008). How many winners are good to have?: On tournaments with sabotage. *Journal of Economic Behavior & Organization 65*(3), 682–702.

Harbring, C. and B. Irlenbusch (2011). Sabotage in Tournaments: Evidence from a Laboratory Experiment. *Management Science 57*(4), 611–627.

Harbring, C., B. Irlenbusch, M. Kräkel, and R. Selten (2007). Sabotage in corporate contests: An experimental analysis. *International Journal of the Economics of Business 14*(3), 367–392.

Harbring, C. and G. K. Lünser (2008). On the competition of asymmetric agents. *German Economic Review 9*(3), 373–395.

Hebb, D. O. (1970). *Textbook of Psychology.* Philadelphia: Saunders.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for research in mathematics education*, 33–46.

Hennig-Schmidt, H., B. Rockenbach, and A. Sadrieh (2005). In search of workers' real effort reciprocity: A field and a laboratory experiment. (55).

Hertwig, R. and O. Andreas (2008). Deception in experiments: Revisting the arguments in its defence. *Ethics and Behaviour 18*, 59–92.

Heyman, J. and D. Ariely (2004). Effort for payment: A tale of two markets. *Psychological Science 15*(11), 787–793.

Hillman, A. L. and E. Katz (1984). Risk-averse rent seekers and the social cost of monopoly power. *The Economic Journal 94*(373), 104–110.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior 7*(3), 346–380.

Hoffman, E. and M. Spitzer (1985). Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *Journal of Legal Studies 14*, 259.

Holt, C. and S. Laury (2002). Risk aversion and incentive effects. *The American economic review 92*(5), 1644–1655.

Hopfensitz, A. and F. van Winden (2008). Dynamic choice, independence and emotions. *Theory and Decision 64*, 249–300.

Huff, N. C. and J. W. Rudy (2004). The amygdala modulates hippocampus-dependent context memory formation and stores cue-shock associations. *Behavioral neuroscience 118*(1), 53.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. *Handbook of Social Cognition 3*, 179–236.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowits (Ed.), *Advances in Experimental Socialpsychology*, pp. 203–253. Academic Press, San Diego.

Isen, A. M. and N. Geva (1987). The influence of positive affect on acceptable level of risk: The person with a large canoe has a large worry. *Organizational Behavior and Human Decision Processes 39*(2), 145–154.

Isen, A. M. and B. Means (1983). The influence of positive affect on decision-making strategy. *Social Cognition 2*(1), 18–31.

Isen, A. M. and R. Patrick (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance 31*(2), 194–202.

Isen, A. M., A. S. Rosenzweig, and M. J. Young (1991). The influence of positive affect on clinical problem solving. *Medical Decision Making 11*(3), 221–227.

Isen, A. M., T. E. Shalker, M. Clark, and L. Karp (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology 36*(1), 1.

Ivanova-Stenzel, R. and D. Kübler (2005). Courtesy and idleness: Gender differences in team work and team competition. *Working Paper*.

Izard, C. E. (1977). *Human Emotions.* New York: Plenum Press.

Kelman, H. (1966). Deception in social research. *Trans-action 3*, 20–24.

Keogh, E. and C. C. French (2001). Test anxiety, evaluative stress, and susceptibility to distraction from rhreat. *European Journal of Personality 15*(2), 123–141.

Kesner, R. P., P. E. Gilbert, and L. A. Barua (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behavioral neuroscience 116*(2), 286.

Koivula, N., P. Hassmen, and J. Fallby (2002). Self-esteem and perfectionism in elite athletes: effects on competitive anxiety and self-confidence. *Personality and individual differences 32*(5), 865–875.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *The American Economic Review 90*(4), 1072–1091.

Konow, J., T. Saijo, and K. Akai (2009). Morals and mores: Experimental evidence on equity and equality. *Working Paper*.

Konrad, K. A. and H. Schlesinger (1997). Risk aversion in rent-seeking and rent-augmenting games. *The Economic Journal 107*(445), pp. 1671–1683.

Kräkel, M. (2005). Helping and sabotaging in tournaments. *International Game Theory Review 7*(02), 211–228.

Krakel, M. and P. Nieken (2012). Relative performance pay in the shadow of crisis.

Ku, H. and T. Salmon (2012). Procedural fairness and the tolerance for income inequality. *Working paper*.

Kube, S., M. Maréchal, and C. Puppe (2006). Putting reciprocity to work: Positive versus negative responses in the field. *University of St. Gallen Economics Discussion Paper* (2006-27).

Kuhnen, C. M. and A. Tymula (2008). Rank expectation, feedback and social hierarchies. *Working Paper*.

Laffont, J. J. and D. Martimort (2002). *The Theory of Incentives*. Princeton University Press.

Lau, C. M. and S. L. Tan (2006). The effects of procedural fairness and interpersonal trust on job tension in budgeting. *Management Accounting Research 17*(2), 171–186.

Lazarus, R. (1966). *Psychological stress and the coping process.* New York, NY, US: McGraw-Hill.

Lazear, E. (1999). Personnel economics: Past lessons and future directions. Technical report, National Bureau of Economic Research.

Lazear, E. (2000a). The future of personnel economics. *The Economic Journal 110*(467), 611–639.

Lazear, E. (2000b). Performance pay and productivity. *Working Paper*.

Lazear, E. and P. Oyer (2007). Personnel economics. Technical report, National Bureau of Economic Research.

Lazear, E. P. (1986). Salaries and piece rates. *The Journal of Business 59*(3), 405–431.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economics 97*(3), 561–580.

Lazear, E. P. and P. Oyer (2000). Personnel economics. Working Paper 13480, National Bureau of Economic Research.

Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–864.

LeDoux, J. (1996). *The Emotional Brain.* New York: Simon and Schuster.

Levitt, E. E. (1967). *The Psychology of Anxiety.* The Bobbs-Merril Company, Inc.

Lévy-Garboua, L., D. Masclet, and C. Montmarquette (2009). A behavioral laffer curve: Emergence of a social norm of fairness in a real effort experiment. *Journal of Economic Psychology 30*(2), 147–161.

Lind, E. A. and T. R. Tyler (1988). *The social psychology of procedural justice.* Plenum Publishing Corporation.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal 97*(387), 666–684.

Mackie, D. M. and L. T. Worth (1991). Feeling good, but not thinking straight: The impact of positive mood on persuasion. *Emotion and Social Judgments 23*, 210–219.

Martens, R., B. D. V. R. B. L. and D. Smith (1990). Development and validation of the competitive state anxiety inventory-2. In R. V. Martens, R. and D. Burton (Eds.), *Competitive Anxiety in Sport*, pp. 117?–190. Champaign, IL: Human Kinetics.

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review 10*(6), 669–697.

Milgrom, P. and J. Roberts (1992). *Economics, Organization, and Management.* New Jersey:Prentice-Hall: Englewood Cliffs.

Millner, E. L. and M. D. Pratt (1989). An experimental investigation of efficient rent-seeking. *Public Choice 62*(2), 139 – 151.

Millner, E. L. and M. D. Pratt (1991). Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice 69*(1), 81–92.

Montague, E. K. (1953). The role of anxiety in serial rote learning. *Journal of Experimental Psychology 45*, 91–96.

Montmarquette, C., J. Rullière, M. Villeval, and R. Zeiliger (2004). Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Science 50*(10), 1379–1389.

Morgan, J., H. Orzen, and M. Sefton (2012). Endogenous entry in contests. *Economic Theory 51*(2), 435–463.

Morris, L., D. D. and C. Hutchings (1981). Cognitive and emotional components of anxiety: literature review and revised worry-emotionality scale. *Journal of Educational Psychology 75*, 541?–555.

Morris, R. G., E. Moser, G. Riedel, S. Martin, J. Sandin, M. Day, and C. O'Carroll (2003). Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358*(1432), 773–786.

Morris, W. N. and P. P. Schnurr (1989). *Mood: The Frame of Mind*. Springer-Verlag New York.

Muchinsky, P. (2002). *Personnel Psychology*. Wiley Online Library.

Münster, J. (2007). Selection tournaments, sabotage, and participation. *Journal of Economics & Management Strategy 16*(4), 943–970.

Murdock, K. (2002). Intrinsic motivation and optimal incentive contracts. *RAND Journal of Economics*, 650–671.

Murray, N., H. Sujan, E. R. Hirt, and M. Sujan (1990). The influence of mood on categorization: A cognitive flexibility interpretation. *Journal of Personality and Social Psychology 59*(3), 411.

Nalbantian, H. R. and A. Schotter (1997). Productivity under group incentives: An experimental study. *The American Economic Review*, 314–341.

Nalebuff, B. J. and J. E. Stiglitz (1983). Prizes and incentives: towards a general theory of compensation and competition. *The Bell Journal of Economics*, 21–43.

Nasby, W. and R. Yando (1982). Selective encoding and retrieval of affectively valent information: Two cognitive consequences of children's mood states. *Journal of Personality and Social Psychology 43*(6), 1244–1255.

Niederle, M. and L. Vesterlund (2011). Gender and competition. *Annu. Rev. Econ. 3*(1), 601–630.

Oatley, K. and P. Johnson-Laird (1987). Towards a cognitive theory of emotions. *Cognition and Emotion 1(1)*, 29–50.

Ong, D. and Z. C. Chen (2012). Tiger women: An all-pay auction experiment on gender signaling of desire to win. *Available at SSRN 1976782*.

Orrison, A., A. Schotter, and K. Weigelt (2004). Multiperson tournaments: An experimental examination. *Management Science 50*(2), 268–279.

Ortony, A. and T. Turner (1990). What's basic about basic emotions? *Psychological review 97*(3), 315.

Ottone, S. and F. Ponzano (2008). How People Perceive the Welfare State. A Real Effort Experiment. *Working Paper*.

Paré, D. (2003). Role of the basolateral amygdala in memory consolidation. *Progress in neurobiology 70*(5), 409–420.

Parfitt, G. and J. Pates (1999). The effects of cognitive and somatic anxiety and self-confidence on components of performance during competition. *Journal of sports sciences 17*(5), 351–356.

Parrott, W. G. and J. Sabini (1990). Mood and memory under natural conditions: Evidence for mood incongruent recall. *Journal of Personality and Social Psychology 59*(2), 321.

Paul R., M. and J. Roberts (1992). *Economics, Organization and Management*. Prentice-Hall International.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik and Hellerman (Eds.), *Emotions: Theory, Research and Experience*. New York Academic.

Pokorny, K. (2008). Pay but do not pay too much: An experimental study on the impact of incentives. *Journal of Economic Behavior & Organization 66*(2), 251–264.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature 37*(1), 7–63.

Price, C. and R. Sheremeta (2012). Endowment origin, demographic effects and individual preferences in contests. *Demographic Effects and Individual Preferences in Contests (March 28, 2012)*.

Rauh, M. and G. Seccia (2006). Anxiety and performance: An endogenous learning by doing model. *International Economic Review 47*(2), 583–609.

Reuben, E. and F. van Winden (2005). Negative reciprocity and the interaction of emotions and fairness norms. *Working paper*.

Roozendaal, B., B. S. McEwen, and S. Chattarji (2009). Stress, memory and the amygdala. *Nature Reviews Neuroscience 10*(6), 423–433.

Rosen, S. (1986). Prizes and incentives in elimination tournaments. *The American Economic Review*, 701–715.

Ruckmick, C. A. (1936). The psychology of feeling and emotion.

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavour 23*(2), 247–265.

Sappington, D. E. M. (1991). Incentives in principal-agent relationship. *Journal of Economic Perpectives*, 45–66.

Sarason, I. (1973). Test anxiety and cognitive modeling. *Journal of Personality and Social Psychology 28*(1), 58–61.

Sarason, I. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and social Psychology 46*(4), 929.

Sarason, I. and E. Palola (1960). The relationship of test and general anxiety, difficulty of task, and experimental instructions to performance. *Journal of Experimental Psychology 59*(3), 185.

Sarason, I., B. Sarason, and G. Pierce (1990). Anxiety, cognitive interference, and performance. *Journal of Social Behavior & Personality*.

Sashegyi, A. I., K. S. Brown, and P. J. Farrel (2000). Application of a generalized random effects regression model for cluster-correlated longitudinal data to a school-based smoking prevention trial. *American Journal of Epistemology 152*(12), 1192–1200.

Schotter, A. and K. Weigelt (1992). Asymmetric tournaments, equal opportunities laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics 107*(2), 511–539.

Schwarz, N., H. Bless, and G. Bohner (1991). Mood and persuasion: Affective states influence the processing of persuasive communications. *Advances in Experimental Social Psychology 24*, 161–199.

Seiler, E. (1984). Piece rate vs time rate: The effect of incentives on earnings. *Review of Econimcs and Statistics 66*, 363–375.

Sheremeta, R. and S. Wu (2011). Optimal tournament design and incentive response: an experimental investigation of canonical tournament theory. *Chapman University, ESI Working Paper*.

Sheremeta, R. and S. Wu (2012). Testing canonical tournament theory: On the impact of risk, social preferences and utility structure. *Working paper*.

Shostak, B. B. and R. A. Peterson (1990). Effects of anxiety sensitivity on emotional response to a stress task. *Behaviour Research and Therapy 28*(6), 513–521.

Shubik, M. (1954). Does the fittest necessarily survive? *M. Shubik, Readings in game theory and related behavior, New York: Doubleday*, 43–46.

Skaperdas, S. and L. Gan (1995). Risk aversion in contests. *The Economic Journal*, 951–962.

Sloof, R. and C. van Praag (2008). Performance measurement, expectancy and agency theory: An experimental study. *Journal of Economic Behavior & Organization 67*(3), 794–809.

Slyker, J. P. and R. McNally (1991). Experimental induction of anxious and depressed mood: Are velten and musical procedures necessary? *Cognitive Therapy and Research 15*(1), 33–45.

Smith, R. H., W. G. Parrott, E. F. Diener, R. H. Hoyle, and S. H. Kim (1999). Dispositional envy. *Personality and Social Psychology Bulletin 25*(8), 1007–1020.

Smith, R. H., W. G. Parrott, D. Ozer, and a. Moniz (1994). Subjective Injustice and Inferiority as Predictors of Hostile and Depressive Feelings in Envy. *Personality and Social Psychology Bulletin 20*(6), 705–711.

Smith, V. (1976). Experimental economics: Induced value theory. *American Economic Review 66*(2), 274–279.

Smith, V. (1979). Incentive compatible experimental processes for the provision of public goods. *Research in experimental economics*.

Spence, J. F. and K. W. Spence (1966). The motivational components of manifest anxiety: Drive and drive. In C. D. Spielberger (Ed.), *Anxiety and Behaviour.* London: Academic Press.

Spence, K. W., I. E. Farber, and H. H. McFann (1956). The relation of anxiety (drive) level to performance in competitional and non-competitional paired-associates learning. *Journal of Experimental Psychology 52(5)*, 296–305.

Spence, K. W., J. Taylor, and R. Ketchel (1957). Anxiety (drive) level and degree of competition in paired-associates learning. *Journal of Experimental Psychology 52*, 306–310.

Spielberger, C. (1966). *Anxiety and Behavior.* Academic Press.

Spielberger, C. (1985). *Anxiety, Cognition and Affect: A State-Trait Perspective.* Lawrence Erlbaum Associates, Inc.

Spielberger, C. D. (1972). *State-Trait Anxiety Inventory.* John Wiley and Sons, Inc.

Spielberger, C. D. (1976). The nature and measurement of anxiety. In C. D. Spielberger and R. Diaz-Guerrero (Eds.), *Cross-cultural Anxiety.* DC: Hemisphere: Washington.

Spielberger, C. D., R. Gorsuch, and R. Lushene (1970). *Manual for the State-Trait Anxiety Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press, Inc.

Spielberger, C. D., P. R. Vagg, L. R. Barker, G. W. Donham, and L. G. Westberry (1980). The factor structure of the state-trait anxiety inventory. In I. G. Sarason and C. D. Spielberger (Eds.), *Stress and Anxiety*, Volume 7. DC: Hemisphere: Washington.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist 52*(6), 613.

Steele, C. M. and J. Aronson (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of personality and social psychology 69*(5), 797.

Sutter, M. and H. Weck-Hannemann (2003). Taxation and the veil of ignorance–a real effort experiment on the laffer curve. *Public Choice 115*(1), 217–240.

Taylor, J. (1953). A personality scale of manifest anxiety. *Journal of Abnormal Social Psychology 48*, 285–290.

Treich, N. (2010). Risk-aversion and prudence in rent-seeking games. *Public Choice 145*(3-4), 339–349.

Tye, K. M., R. Prakash, S.-Y. Kim, L. E. Fenno, L. Grosenick, H. Zarabi, K. R. Thompson, V. Gradinaru, C. Ramakrishnan, and K. Deisseroth (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature 471*(7338), 358–362.

Uske, T. (2009). Tournament fever and the perception of uncertainty in performance contests. *Working paper*.

van Dijk, F., J. Sonnemans, and F. Van Winden (2001). Incentive systems in a real effort experiment. *European Economic Review 45*(2), 187–214.

van Long, N. and N. Vousden (1987). Risk-averse rent seeking with shared rents. *The Economic Journal 97*(388), 971–985.

van Winden, F., M. Krawczyk, and A. Hopfensitz (2010). Investment, resolution of risk, and the role of affect. *CESifo Working Paper Series No. 2975.*

Vesterlund, L. and M. Nierderle (2007). Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics 122*(3), 1067–1101.

Weigelt, K., J. Dukerich, and A. Schotter (1989). Reactions to discrimination in an incentive pay compensation scheme: A game-theoretic approach. *Organizational Behavior and Human Decision Processes 44*(1), 26–44.

Woodman, T. and L. Hardy (2001). Stress and anxiety. In R. N. Singer and C. M. Janelle (Eds.), *Handbook of Sport Psychology* (2nd ed.). New York: Wiley.

Woodman, T. and L. Hardy (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of sports sciences 21*(6), 443–457.

Wu, G. (1999). Anxiety and decision making with delayed resolution of uncertainty. *Theory and Decision 46*, 159–199.

Yamazaki, T. (2008). On the existence and uniqueness of pure-strategy nash equilibrium in asymmetric rent-seeking contests. *Journal of Public Economic Theory 10*(2), 317–327.

Zaichkowsky, L. D. and A. Baltzell (2001). Arousal and performance. In R. N. Singer and C. M. Janelle (Eds.), *Handbook of Sport Psychology* (2nd ed.). New York: Wiley.

Zizzo, D. J. (2004a). *Inequality and procedural fairness in a money-burning and stealing experiment*, Volume 11. Emerald Group Publishing Limited.

Zizzo, D. J. (2004b). The neuroeconomics of anger. *Homo Oeconomicus 21*, 473–494.

Zizzo, D. J. (2008). Anger and economic rationality. *Journal of Economic Methodology 15*(2), 147–167.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics 13(1)*, 75–98.